The Bayesian Approach to the Philosophy of Science

Michael Strevens

For the *Macmillan Encyclopedia of Philosophy*, second edition

The posthumous publication, in 1763, of Thomas Bayes' "Essay Towards Solving a Problem in the Doctrine of Chances" inaugurated a revolution in the understanding of the confirmation of scientific hypotheses—two hundred years later. Such a long period of neglect, followed by such a sweeping revival, ensured that it was the inhabitants of the latter half of the twentieth century above all who determined what it was to take a "Bayesian approach" to scientific reasoning.

Like most confirmation theorists, Bayesians alternate between a descriptive and a prescriptive tone in their teachings: they aim both to describe how scientific evidence is assessed, and to prescribe how it ought to be assessed. This double message will be made explicit at some points, but passed over quietly elsewhere.

Subjective Probability

The first of the three fundamental tenets of Bayesianism is that the scientist's epistemic attitude to any scientifically significant proposition is, or ought to be, exhausted by the *subjective probability* the scientist assigns to the proposition. A subjective probability is a number between zero and one that re-

flects in some sense the scientist's confidence that the proposition is true. (Subjective probabilities are sometimes called *degrees of belief* or *credences*.)

A scientist's subjective probability for a proposition is, then, more a psychological fact about the scientist than an observer-independent fact about the proposition. Very roughly, it is not a matter of how likely the truth of the proposition actually is, but about how likely the scientist thinks it to be. Thus *subjective*—though with hindsight, *psychological* might have been a better term.

Unlike every other approach to confirmation theory, Bayesianism has no use for the notion of theory acceptance: there is no amount of evidence sufficient to induce a qualitative shift in a Bayesian's epistemic attitude from not accepting to accepting a theory. Learning from the evidence is always a matter of a quantitative adjustment, of changing your subjective probability for a hypothesis to reflect the latest evidence. At any time, the most favored theories are simply those with the highest subjective probabilities.

To found its first tenet, Bayesianism must establish that it is plausible to suppose, or reasonable to require, that scientists have a subjective probability for every proposition that figures in their inquiry. Ramsey (1931) proposed that to have a subjective probability for a proposition is to have a certain complex disposition to act, a disposition that can be measured at least tolerably well in many cases by assessing betting behavior, as follows. The higher your subjective probability for a proposition, the lower the odds, all other things being equal, you will be prepared to accept in betting on the truth of that proposition. To be precise, given a subjective probability p for the proposition, you will accept odds of up to $p : (1 - p)$ on its truth—you will avoid just those bets, in other words, where you have to pay in more than p for every dollar you stand to win, so that, for example, if your subjective probability for the proposition is 0.3 then you will pay no more than $3 to play a game in which you win $10 just in case the proposition is true. Ramsey thought it very likely that we have appropriately stable behavioral

dispositions of this sort, accessible to measurement using the betting test, with respect to just about any proposition we understand, and so that we have subjective probabilities for all these propositions.

The Bayesian's principal tool is mathematical argument, and the mathematics in question is the probability calculus—the standard mathematics of probability—to which all subjective probabilities are assumed to conform. Conformance to the axioms is Bayesianism's second fundamental tenet.

Here the Bayesian argument tends to take a prescriptive turn. Having established that scientists have, as a matter of psychological fact, subjective probabilities for all propositions that matter, the next step is to show that scientists ought to—whether they do or not—arrange their probabilities so as to satisfy the axioms of the probability calculus.

Typically this is done by way of a Dutch Book argument, an argument that shows that, if you do not adhere to the calculus, there is a certain set of bets on the truth of various propositions that you are committed in principle to accepting, but that will lead to a certain loss however things turn out.

Objections to the Dutch Book argument typically turn on the vagueness of the idea that you are "committed in principle" to accepting the bets in question; replies to these objections attempt to make the nature of the commitment more precise without leavening its evident undesirability.

For more on Ramsey's subjectivism, the axioms of probability, and the Dutch book argument, see the entry on Probability and Chance.

Bayesian Conditionalization

The third of Bayesianism's three fundamental tenets is Bayes' conditionalization rule, which instructs you on how to update your subjective probabilities as the evidence arrives.

There are four steps to Bayes' rule. The first step is to define prior and posterior subjective probability. These notions are relative to your receipt of

a piece of evidence: your prior probability for a hypothesis is your subjective probability for the hypothesis immediately before the evidence comes in; your posterior probability for the hypothesis is your subjective probability immediately after the evidence (and nothing else) comes in. Bayes' rule gives you a formula for calculating your posterior probabilities for every hypothesis given your prior probabilities and the nature of the evidence. In so doing, it offers itself as the complete story as to how to take evidence into account. In what follows, prior subjective probabilities are written as $C(\cdot)$, and posterior subjective probabilities as $C^+(\cdot)$.

The second step towards Bayes' rule is the introduction of the notion of conditional probability, a standard notion in probability mathematics. An example of a conditional probability is the probability of obtaining a four on a die roll, given that an even number is obtained. This probability is 1/3, since there are three equally probable ways for a die roll to be even, one of which is a four. Formally, the probability of a proposition h conditional on another proposition g is written $C(h|g)$; it is usually defined to be $C(hg)/C(g)$. (Alternatively, conditional probability may be taken as a primitive, as explained in the entry on Probability and Chance.)

The third step is to make the following simple posit about conditionalization: when you receive a piece of evidence e, you should update your probability for any given hypothesis h so that it is equal to your prior probability for h given e. That is, on learning that e is true, you should set your posterior probability $C^+(h)$ equal to your prior probability $C(h|e)$. This *is* Bayes' rule in its simplest form, but one further step will produce a more familiar, and revealing, version of the rule.

The fourth and final step is to notice a simple mathematical consequence of the definition of conditional probability, confusingly called Bayes' theorem (confusing because Bayes' theorem and Bayes' rule are two very different propositions). According to Bayes' theorem,

$$C(h|e) = \frac{C(e|h)}{C(e)}C(h).$$

Combine Bayes' theorem and the simple form of Bayes' rule and you obtain the more familiar version of Bayes' rule:

$$C^+(h) = \frac{C(e|h)}{C(e)}C(h).$$

The effect of the application of Bayes' rule, then—or as philosophers usually say, the effect of *Bayesian conditionalization*—is, on receipt of e, to multiply the old probability for h by the factor $C(e|h)/C(e)$. Call this factor the *Bayesian multiplier*.

What justification can be offered for Bayesian conditionalization? Since the notion of conditional probability is introduced by definition, and Bayes' theorem is a simple consequence of the definition, this amounts to the question why you ought, on learning e, to set your posterior probability for a hypothesis h equal to the prior probability $C(h|e)$.

Various arguments for conditionalizing in this way exist in the literature, often based on Dutch book considerations that invoke the notion of a conditional bet. The consensus is that none is entirely convincing. It is important to note that mathematics alone cannot settle the question: the probability calculus relates only different probabilities that are part of the same overall distribution, whereas Bayes' rule relates probabilities from two quite different distributions, the prior and posterior distributions.

Two further remarks on Bayesian conditionalization. First, Bayes' rule assumes that the subjective probability of the evidence e goes to one when it is acquired, therefore that when evidence arrives, its content is exhausted by a proposition that comes to be known for sure. A natural extension of the rule, called Jeffrey conditionalization, relaxes this assumption (Jeffrey 1983).

Second, you may wonder whether background knowledge must be taken into account when conditionalizing. In fact, it is automatically taken into account: background knowledge has subjective probability one, and for any proposition k with probability one, $C(h|k) = C(h)$; thus, your subjective probability distribution always has your background knowledge in every re-

spect "built in".

Now to discuss the implications of Bayesianism for confirmation. (Further implications will be considered in the next section.)

The impact of evidence e on a hypothesis h is determined, recall, by the Bayesian multiplier, $C(e|h)/C(e)$, which when multiplied by the prior for h yields its posterior. You do not need any great mathematical expertise to see that, when $C(e|h)$ is greater than $C(e)$, the probability of h will increase on receipt of e, while when it is $C(e)$ that is greater, the probability of h will decrease.

When the receipt of e causes the probability of h to increase, e is said to *confirm h*. When it causes the probability of h to decrease, it is said to *disconfirm h*. This may look like a definition, but it is in fact a substantive philosophical thesis: the Bayesian claims that the preexisting notions of confirmation and disconfirmation can be given a satisfactory Bayesian analysis. (Or at least, the Bayesian usually makes this claim: they also have the option of interpreting their definition as a piece of revisionism, not intended to capture our actual notion of confirmation but to replace it with something better.)

Two remarks. First, to say that a hypothesis is confirmed is only to say that its probability has received some kind of upward bump. The bump may be very small, and the resulting posterior probability, though higher than the prior, may be almost as small. The term *confirmed* has, in philosophical usage, a very different sense from a term such as *verified*.

Second, since whether or not a piece of evidence confirms a hypothesis depends on a subjective probability distribution, confirmation is in the first instance a relative matter. More on this in *The Subjectivity of Bayesian Confirmation* below.

One further definition: the quantity $C(e|h)$ is called a *likelihood*, specifically the likelihood of h on e (not to be confused with the probability of h given e, though there is a close relationship between the two, spelled out by

6

Bayes' theorem).

The significance of the Bayesian multiplier can now be stated in natural language: a piece of evidence confirms a hypothesis relative to a particular subjective probability distribution just in case the likelihood of the hypothesis on the evidence is greater than the subjective probability for the evidence.

Consider a special case, that in which a hypothesis h entails the evidence e. By a theorem of the probability calculus the likelihood of h on e, that is, $C(e|h)$, is in any such case equal to one. Suppose that e is observed to be true. Assuming that $C(e)$ is less than one (which will be true unless all viable hypotheses predict e), then the likelihood will be greater than $C(e)$, and so h will be confirmed. Ignoring the parenthetical qualification, a hypothesis is always confirmed by its predictions. Further, the more surprising the prediction, in a sense—the lower the prior probability of e—the more h will be confirmed if e is in fact observed.

The significance of this observation is limited in two ways. First, some hypotheses predict evidence only with a certain probability less than one. Second, hypotheses tend to make observable predictions only in conjunction with other, "auxiliary" hypotheses. The Bayesian response will be considered in the next section.

The Bayesian Machine

Suppose you want to know whether a certain coin is fair, that is, biased neither towards heads nor tails. You toss the coin ten times, obtaining exactly five heads and five tails. How to conditionalize on this evidence? You will need three subjective probabilities: the prior probability for the hypothesis h that the coin is fair, the prior probability for the evidence e, and the likelihood of h on e. A good Bayesian is committed to adopting definite values for these subjective probabilities one way or another. If necessary, they will be set "by hand", that is, by some sort of reflective process that is constrained

only by the axioms of the probability calculus. But a great part of the appeal of Bayesianism is that the vast majority of subjective probabilities can be set "mechanically", that is, that they will have their values fully determined once a few special probabilities are set by hand. In the case of the coin, once the prior probability for h and its rivals is set by hand, a little philosophy and mathematics of probability will take care of everything else, mechanically fixing the likelihood and the probability for the evidence.

Begin with the likelihood, the probability of getting exactly five heads in ten tosses given that the coin is fair. Since the fairness of the coin entails (suppose) both a physical probability for heads of 0.5 and the independence of the tosses, the hypothesis that the coin is fair assigns a definite physical probability to your observed outcome of five heads—a probability of about 0.25, as it happens. Intuitively, it seems right to take this as the likelihood—to set your subjective probability $C(e|h)$, that is, equal to the physical probability that h assigns to e. In its sophisticated form, this intuition is what is sometimes known as Miller's Principle or the Principal Principle; call it the Probability Coordination Principle or PCP for short (see the entry on Probability and Chance). Bayesians normally take PCP on board, thus relieving you of the effort of setting a value by hand for the likelihood in a case such as this.

Now consider the probability of the evidence. A theorem of the probability calculus, the total probability theorem, looks (in one of its forms) like this:

$$C(e) = C(e|h_1)C(h_1) + C(e|h_2)C(h_2) + \cdots$$

where the hypotheses $h_1, h_2, \ldots$ form a mutually exclusive, exhaustive set, in the sense that one and only one of them must be true. In many cases, the set of hypotheses among which you are trying, with the help of e, to decide form such a set (though see below). Thus, if you have set values for the likelihoods $C(e|h_i)$ and prior probabilities $C(h_i)$ for all your rival hypotheses, the probability calculus gives you a unique correct subjective

probability to assign to *e*.

To sum up: if your rival hypotheses assign definite physical probabilities to the evidence *e* and form a mutually exclusive, exhaustive set, then by an independent principle of rationality, PCP, and a theorem of the probability calculus, total probability, the Bayesian multipliers for all of the hypotheses are completely determined once their prior probabilities are fixed.

As a consequence, you need only assign subjective probabilities by hand to a relatively small set of propositions, and only once in your life: at the very beginning, before any evidence comes in, you will assign subjective probabilities to every possible scientific hypothesis. These assignments made, everything you need for Bayesian conditionalization is decided for you by PCP and the probability axioms. In this sense, Bayesian confirmation runs like a well-conditioned machine: you flip the *on* switch, by assigning initial prior probabilities to the different hypotheses that interest you, and then sit back and enjoy the evidential ride. (Conditionalization is also machine-like without PCP and total probability, but in that case flipping the *on* switch involves assigning values to $C(e|h_i)$ and $C(e)$ for every possible piece of evidence *e*.)

There are two obstacles to the smooth functioning of the Bayesian machine. First, it may be that some or all of the rival hypotheses do not, on their own, assign a determinate physical probability to the evidence. In such cases, the likelihood must either be fixed by hand, without the help of PCP, or (more usually in the quantitative sciences) by supplementing the hypothesis with an auxiliary hypothesis in conjunction with which it does fix a physical probability for the evidence. In the latter case, PCP can be applied but complications arise when, as is typical, the truth of the auxiliary hypothesis is itself not known for sure. The conjunction of original and auxiliary hypothesis may be confirmed or disconfirmed mechanically, but the implication for the original hypothesis on its own—whether it is confirmed, and if so by how much—will continue to depend on handcrafted likelihoods such as $C(e|h)$. This is the Bayesian's version of confirmation theory's Quine-

Duhem problem (cross-reference). Strevens (2001) offers a partial solution to the problem. (The application of PCP will also fall through if the evidence is "inadmissible"; see the discussion of probability coordination in the entry on Probability and Chance.)

Second, even when the likelihoods are fixed mechanically, the theorem of total probability may not apply if the rival hypotheses are either not mutually exclusive or not exhaustive. Lack of exhaustiveness is the more pressing worry, as it would seem to be the norm: exhaustiveness implies that you have thought of every possible theory that predicts e to any extent—an unlikely feat. A simple fix is to include a residual hypothesis in your set to the effect that none of the other hypotheses is correct. Such a hypothesis will not, however, determine a definite physical probability for the evidence, so its likelihood, and therefore the probability for the evidence, will after all have to be fixed by hand.

Bayesianism and the Problem of Induction

Does the Bayesian theory of confirmation solve the problem of induction? The case for an affirmative answer: adherence to the tenets of Bayesianism can be justified a priori (by Dutch book arguments and the like, or so some philosophers believe). And this adherence alone is sufficient to turn you into an inductive reasoner: once you have settled on priors for all the hypotheses, the Bayesian machinery tells you what sort of things to expect in the future given your experience of the past.

Suppose, for example, that you wish to predict the color of the next raven. You have various theses about raven color: all ravens are blue; ravens are green with 50% probability, otherwise black; all ravens are black, and so on. In your life to date, you have observed a number of ravens, all of them black. This evidence rules out altogether some of the raven color theses, such as the thesis that all ravens are blue. (The likelihood of the blue the-

sis on this evidence is zero, so the multiplier is zero: observation of a black raven therefore causes your subjective probability for the blue thesis to drop to zero.)

Other theses have their probability shifted around by the evidence in other ways. The more they probabilify the evidence, the greater their likelihoods on the evidence and so the higher their Bayesian multipliers. Observing many black ravens has the effect, then, of moving your subjective probability away from hypotheses that do not probabilify blackness and towards theses that do. As a result, the observation of many black ravens in the past increases your subjective probability that the next raven will be black. Thus you have an a priori argument—the argument for accepting Bayesianism— that justifies inductive behavior.

The case for a negative answer as to whether Bayesianism solves the problem of induction can be made in two ways: by arguing that the a priori arguments for adopting the Bayesian apparatus fall through, or by arguing that Bayesianism does not, after all, underwrite inductive behavior. The second approach is the more illuminating.

Return to the ravens. The theses listed above have the uniformity of nature as a consequence: if any is true, then the future will be, with respect to raven color, like the past. Once some non-uniform theses are thrown into the mix, everything changes. Consider, for example, the following thesis, reminiscent of Goodman's grue puzzle: all ravens observed until now are black, the rest green. The Bayesian multipliers for this thesis and the thesis that all ravens are black remain the same as long as all observed ravens are black, which is to say, up until this point in time. Just as probability has been flowing to the latter hypothesis, it will have been flowing to the former. It turns out, then, that the probability flow is not only towards theses that predict blackness for future ravens, but also towards many others. Since the multipliers for these theses have been the same until now, your predictions about the color of ravens will favor blackness only if your initial prior

probabilities—the probabilities you assigned to the different theses before any evidence came in—already favored the thesis that all ravens are black over the grue-like thesis, which is to say, only if you yourself already favored uniformity over diversity.

Many Bayesians have made their peace with Bayesianism's open-minded policy on natural uniformity. Howson (2001) argues, for example, that the Bayesian approach should not be considered so much a positive theory of confirmation—of how evidence bears on hypotheses—as a framework for implementing any theory of confirmation you like.

The Subjectivity of Bayesian Confirmation

Suppose that the Bayesian machine is in good working order: you choose your prior probabilities for the rival hypotheses, and then let the evidence, in conjunction with PCP and the total probability theorem, do the rest. Even then, with your personal input limited to no more than an assessment of the initial plausibility of the rival hypotheses, there is an unsettling element of subjectivity to the process of Bayesian confirmation, which is perhaps best brought out by the following observation: two scientists who agree on the physical probabilities that a hypothesis h assigns to evidence e, and who follow PCP, so assigning the same value to the likelihood $C(e|h)$, may disagree on whether e confirms or disconfirms h.

To see why: e confirms h if the Bayesian multiplier is greater than one, and disconfirms it if the multiplier is less than one. The question, then, is whether $C(e|h)$ is greater than or less than $C(e)$. The scientists agree on $C(e|h)$, but they may have different values for $C(e)$: a scientist who assigns higher prior probabilities to hypotheses that assign higher physical probabilities to e will have a higher value for $C(e)$. It is quite possible for the two scientists' priors for e to fall on either side of $C(e|h)$, in which case one will take e to confirm, the other to disconfirm, h.

A *radical personalist* denies that this is a problem: why should two scientists agree on the significance of the evidence when one was expecting the evidence much more than the other? In the extreme, personalism of this sort approaches the view that Bayesian confirmation theory provides no guidance at all on assessing the significance of evidence, other than by establishing a standard of consistency (Howson 2001); see also the discussion of induction above.

There is some objectivity underlying Bayesianism's subjectivity, however. The two scientists above will, because they agree on the likelihoods, agree on the ordering of the Bayesian multipliers. That is, they will agree on which of any two hypotheses has the higher Bayesian multiplier, even though they may disagree on the size of the multipliers.

An important consequence of this agreement is a result about the convergence of opinion. When hypotheses assign physical probabilities to the evidence, as assumed here, it can be shown that as time goes on, the subjective probability distributions of any two scientists will with very high physical probability converge on the truth, or rather to the class of hypotheses empirically equivalent to the truth. (Even when the likelihoods are purely subjective, or at least only as objective as the probability calculus requires, a convergence result, albeit more limited, can be proved.)

Many Bayesians regard this convergence as ameliorating, in every important way, the subjective aspect of Bayesianism, since any disagreements among Bayesian scientists are ephemeral, while agreement lasts forever. Indeed, that Bayesianism makes some, but not too much, room for scientific dissent may not unreasonably be seen as an advantage, in both a descriptive and a prescriptive light.

Now consider a contrary view: while dissent has its place in science, it has no place in scientific inference. It is fine for scientists to disagree, at least for a time, on the plausibility of various hypotheses, but it is not at all fine that they disagree on the impact of the evidence on the hypotheses—agreement

on the import of the evidence being the sine qua non of science. In Bayesian terms, scientists may disagree on the priors for the rival hypotheses, but they had better not disagree on the Bayesian multipliers. But this is, for a Bayesian, impossible: the priors help to determine the multipliers. The usual conclusion is that there is no acceptable Bayesian theory of confirmation.

A less usual conclusion is that Bayesianism is still viable, but only if some further principle of rationality is used to constrain the prior probabilities in such a way as to determine uniquely correct values for the Bayesian multipliers. This is *objectivist Bayesianism*. Just as PCP is used to determine definite, objective values for the likelihoods, the objectivists suggest, so another rule might be used to determine definite, objective values for the prior probabilities of the hypotheses themselves, that is, for the subjective probabilities $C(h)$.

What principle of rationality could possibly tell you, before you have any empirical evidence whatsoever, exactly how plausible you ought to find some given scientific hypothesis? Objectivists look to the principle of indifference for the answer. That principle, discussed more fully in the entry on Probability and Chance, is in one of its guises intended to specify a unique probability distribution over a set of propositions, such as hypotheses, that reflects complete ignorance as to which of the set is true. Thus the very fact that you have no evidence is itself taken to commit you to a particular assignment of prior probabilities—typically, a probability distribution that is uniform in some sense (Jaynes 1983).

The objectivist envisages all Bayesian reasoners marching in lock-step: they start with precisely the same priors; they apply (thanks to PCP and total probability) precisely the same Bayesian multipliers; thus, they have the same subjective probabilities at all times for everything.

There are various powerful objections to the most general forms of the principle of indifference. Even its most enthusiastic supporters would shy away from claiming that it determines a uniquely correct prior for absolutely

any scientific hypothesis. Thus the lock-step picture of Bayesian inference is offered more as an ideal than as a realistic prospect. To be a modern objectivist is to argue that parts of science, at least, ought to come close to realizing the ideal.

The Problem of Old Evidence

Among the many achievements of Newton's theory of gravitation was its prediction of the tides and their relation to the lunar orbit. Presumably, the success of this prediction confirmed Newton's theory, or in Bayesian terms, the observable facts about the tides e raised the probability of Newton's theory h.

But the Bayesian, it turns out, can make no such claim. Because the facts about the tides were already known when Newton's theory was formulated, the probability for e was equal to one. It follows immediately that both $C(e)$ and $C(e|h)$ are equal to one (the latter for any choice of h). But then the Bayesian multiplier is also one, so Newton's theory does not receive any probability boost from its prediction of the tides. As either a description of actual scientific practice, or a prescription for ideal scientific practice, this is surely wrong.

The problem generalizes to any case of "old evidence": if the evidence e is received before a hypothesis h is formulated, then e is incapable of boosting the probability of h by way of conditionalization (Glymour 1980, chap. 4). As is often remarked, the problem of old evidence might just as well be called the problem of new theories, since there would be no difficulty if there were no new theories, that is, if all theories were on the table before the evidence began to arrive. Whatever you call it, the problem is now considered by most Bayesians to be in urgent need of a solution. A number of approaches have been suggested, none of them entirely satisfactory.

A recap of the problem: if a new theory is discovered midway through an

inquiry, a prior must be assigned to that theory. You would think that, having assigned a prior on non-empirical grounds, you would then proceed to conditionalize on all the evidence received up until that point. But because old evidence has probability one, such conditionalization will have no effect. The Bayesian machinery is silent on the significance of the old evidence for the new theory.

The first and most conservative solution to the problem is to take the old evidence into account in setting your prior for the new theory. In doing this, you are entirely on your own: you cannot use conditionalization or any other aspect of the Bayesian apparatus to weigh the evidence in a principled way. But, because you are free to choose whatever prior you like, you are free to do so in part on the basis of the old evidence.

A second solution requires a radical revision of Bayesian conditionalization, so as to allow conditionalization using not the actual probability of the old evidence, but a (now) counterfactual probability such as your prior for the evidence immediately before you learned it. This provides a natural way to use conditionalization to weigh the old evidence, but the difficulties involved in choosing an appropriate counterfactual prior and in justifying conditionalization on the false prior, rather than the actual prior, have not unreasonably scared most Bayesians away.

The third and perhaps most popular solution suggests that, although conditionalization on old evidence e has no effect on the prior probability of a new theory h, conditionalizing on the fact that h predicts e (for simplicity's sake, assume that it entails e) may have an effect. The idea: until you formulate h, you do not know that it entails e. Once h is formulated and assigned a prior, you may conditionalize on the fact of the entailment; learning that h entails e will have much the same impact on the probability of h, it is supposed, as learning e would have had if it were not already known.

There are two difficulties with this suggestion. The first is that facts about entailment (either of e itself, or of a physical probability for e) are

logical truths, which ought according to the probability calculus to be assigned probability one at all times—making the logical facts as "old" as the evidence itself. Proponents of the present approach to old evidence argue not unreasonably that a sophisticated Bayesianism ought to allow for logical learning, so that it is the requirement that subjective probabilities conform to the probability calculus in every respect that is at fault here, for imposing an unreasonably strict demand on flesh and blood inquirers.

The second (and related) difficulty is that the theory of conditionalization on logical facts is not nearly so nicely developed as the theory of orthodox Bayesian conditionalization. A case can be made that conditionalizing on h's entailment of old evidence will increase the probability of h, but the details are complicated and controversial. See Earman (1992) for further discussion.

Bayesianism Accessorized

Two notable additions to the Bayesian apparatus are ever under consideration.

First is a theory of acceptance, that is, a theory that dictates, given your subjective probabilities, which hypotheses you ought to "accept". Conventional Bayesianism has no need of acceptance: your subjective probabilities are taken to exhaust your epistemic attitudes to the hypotheses, and also to determine, along with your preferences in the usual decision-theoretical way, the practical significance of these attitudes.

Some philosophers argue that there is, nevertheless, work for a notion of acceptance to do, and hold either a simple view on which hypotheses with high subjective probability are to be accepted, or a more sophisticated view on which not only probability but the consequences for science, good and bad, of acceptance must be taken into account (Levi 1967).

Second is a theory of confirmational relevance, that is, a theory that dic-

tates, given your subjective probabilities, to what degree a given piece of evidence confirms a given hypothesis. Conventional Bayesianism has no need of confirmational relevance: your subjective probabilities are taken to exhaust your epistemic attitudes to the hypotheses, and so the dynamics of confirmation are exhausted by the facts about the way in which the subjective probabilities change, which are themselves fully determined, through conditionalization, by the values of the subjective probabilities themselves. Nothing is added to the dynamics of probability change—nothing could be added—by finding a standard by which to judge whether certain evidence has a "large" or "small" impact on the hypotheses; however you talk about probability change, it is the change that it is. (A truly pure-hearted Bayesian need not even define *confirms* and *disconfirms*.)

Many different measures of relevance have, nevertheless, been proposed (Fitelson 1999). The simple *difference measure* equates the relevance of *e* to *h* with the difference between the prior and posterior probabilities of *h* after conditionalization on *e*, or equivalently, with $C(h|e) - C(h)$. The *likelihood measure* equates the relevance of *e* to *h* with $C(e|h)/C(e|\neg h)$. It should be noted that all popular Bayesian measures render relevance relative to background knowledge.

There is no doubt that scientists sometimes talk about accepting theories and about the strength of the evidence—and that they do not talk much about subjective probability. The degree to which you see this as a problem for unadorned Bayesian confirmation theory itself measures, perhaps, your position on the spectrum between prescriptive and descriptive.

References

Earman, J. (1992). *Bayes or Bust?* MIT Press, Cambridge, MA.

Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science* 66:S362–S378.

Glymour, C. (1980). *Theory and Evidence*. Princeton University Press, Princeton, NJ.

Howson, C. (2001). *Hume's Problem: Induction and the Justification of Belief*. Oxford University Press, Oxford.

Jaynes, E. T. (1983). *Papers on Probability, Statistics, and Statistical Physics*. Edited by R. Rosenkrantz. D. Reidel, Dordrecht.

Jeffrey, R. C. (1983). *The Logic of Decision*. Second edition. Chicago University Press, Chicago.

Levi, I. (1967). *Gambling with the Truth*. MIT Press, Cambridge, MA.

Ramsey, F. (1931). Truth and probability. Reprinted in *Philosophical Papers*. Edited by D. H. Mellor. Cambridge University Press, Cambridge.

Strevens, M. (2001). The Bayesian treatment of auxiliary hypotheses. *British Journal for the Philosophy of Science* 52:515–538.