5 | Commentary

5 | Commentary

5.1 | Bayes for Beginners

Although the Bayesian approach to confirmation is not inconsistent with the hypothetico-deductive (H-D) approach (discussed in chapter 4), it differs from it in several fundamental respects. The aim of this first section is to explain and illustrate those differences as a prelude to the analysis of the readings in chapter 5 later in this commentary.

The Relevance Criterion of Confirmation

A distinctive feature of the Bayesian approach is its reliance on the mathematical theory of probability. Unlike the H-D approach, which treats confirmation as a qualitative notion that might be made quantitative later on, Bayesians assume that confirmation is quantitative from the outset. Even such qualitative notions as evidence confirming a hypothesis and evidence confirming one hypothesis more strongly than it does another are analyzed in terms of probabilities with numerical values that lie between 0 and 1. Bayesians contend that this essentially quantitative approach to confirmation in terms of probability theory can solve the puzzles and paradoxes afflicting purely qualitative theories of confirmation, such as Hempel's satisfaction criterion.

The most common way that Bayesians connect confirmation with probability is by adopting the relevance criterion of confirmation, according to which a piece of evidence, E, confirms a hypothesis, H, if and only if E raises the probability of H:

Relevance Criterion of Confirmation	E confirms H if and only if $P(H/E) > P(H)$; E disconfirms H if and only if $P(H/E) < P(H)$.

For convenience, we shall often refer to $P(H)$ as the *prior probability of H*, and $P(H/E)$ as the *posterior probability of H*. $P(H/E)$ is a conditional probability and should be read as the *probability of H given E*.

It should be noted that, although it is defined in terms of quantitative probabilities, the notion of confirmation (often called *incremental confirmation*) defined by the relevance criterion is qualitative. The relevance criterion does not specify how degrees of confirmation should be measured; it merely gives a necessary and sufficient condition for that confirmation. Indeed, there is an ongoing dispute in the literature about whether numerical degrees of confirmation should be a function of the ratio of $P(H/E)$ to $P(H)$ or a function of the difference between them.[1] Regardless of where they stand on this issue, all Bayesians agree that the more E raises the probability of H, the more E confirms H.

The relevance criterion of confirmation differs significantly from the absolute criterion of confirmation, according to which E confirms H if and only if $P(H/E)$ exceeds some suitably high threshold value, say, 0.9. From the point of view of those who endorse the relevance criterion, the absolute criterion confuses confirmation with acceptance. High probability may be an appropriate condition for accepting a hypothesis, but it is not necessary for confirmation. Thus, adherents to the relevance criterion would consider H confirmed by E even though E raised the probability of H only a little, from, say, 0.2 to 0.4, and the posterior probability of H, $P(H/E)$, remained less than 0.5.

Bayes's Theorem and the Axioms of Probability Theory

Bayes's theorem (also called Bayes's rule, law, or equation) lies at the heart of the Bayesian approach to confirmation and gives that approach its name. In this section we shall be concerned solely with Bayes's theorem as a formal result in probability theory. As such, Bayes's theorem, like any mathematical theorem, is entirely uncontroversial. What is distinctively Bayesian about the Bayesian approach to confirmation is not merely its use of Bayes's theorem but its interpretation of the probabilities occurring in the theorem. The Bayesian interpretation of probabilities as subjective degrees of belief will be discussed later, in the section "Probabilities and Degrees of Belief."

Bayes's theorem is a deductive consequence of the three basic axioms of probability theory. Everything else in probability theory can also be deduced from these axioms, supplemented with definitions of notions such as conditional probability. Here are the axioms in their unconditional form.

Axiom 1	Every probability is a real number between 0 and 1: $0 \leq P(A) \leq 1$.
Axiom 2	If A is a necessary truth, then $P(A) = 1$.
Axiom 3	If A and B are mutually exclusive (that is, if it is impossible for both A and B to be true), then $P(A \vee B) = P(A) + P(B)$. This theorem is often referred to as the *special addition rule*.

Strictly speaking, the A, B, C, and so on that probability ranges over are propositions, but we shall, when convenient, talk about the probability of events, theories, classes of theories, and evidence.

Even though the set of axioms is small, several important rules that we shall use later on can be deduced from them.

Negation Rule	$P(\sim A) = 1-P(A)$.
Implication Rule	If A logically entails B, then $P(B) \geq P(A)$.
Equivalence Rule	If A and B are logically equivalent, then $P(A) = P(B)$.
General Addition Rule	$P(A \vee B) = P(A)+P(B) - P(A\&B)$.

The general addition rule is especially useful, since it applies regardless of whether A and B are exclusive. Obviously, when A and B are mutually exclusive, $P(A\&B)$ is 0 and the general addition rule reduces to the special addition rule (axiom 3).

One way to make the general addition rule intuitively obvious is to represent propositions by circles and to let the probability of each proposition equal the area of its circle.[2] When A and B are mutually exclusive, the A-circle and the B-circle do not overlap, and the probability of $(A \vee B)$ is simply the sum of $P(A)$—the area of the A-circle—and $P(B)$—the area of the B-circle (figure 1). When A and B are not exclusive, the circles overlap (figure 2).

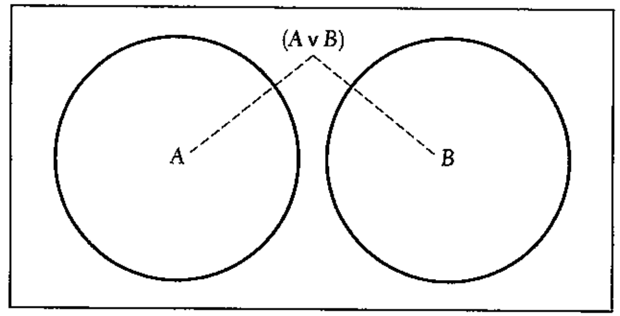

Figure 1

Thus, to calculate $P(A \vee B)$ for figure 2 we add the areas of the two circles as before, but then we have to subtract $P(A\&B)$, which is the area of the overlap, to get the correct answer.[3]

To derive Bayes's theorem from the probability axioms, we need a definition of $P(A/B)$, the conditional probability of A given B.[4] It is:

Definition of Conditional Probability
$$P(A/B) = \frac{P(A\&B)}{P(B)}, \text{ where } P(B) > 0.$$

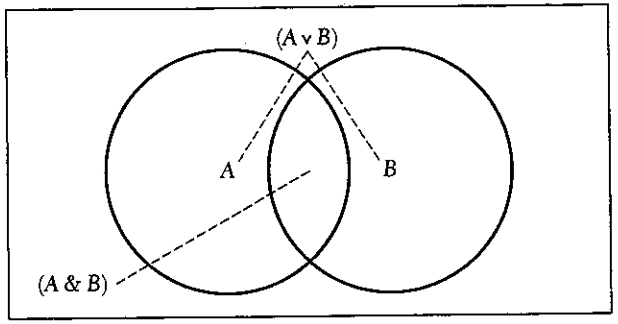

Figure 2

The rationale for adopting this definition of conditional probability can be appreciated by considering figure 2. Suppose that you are told that a dart has been thrown, randomly, at the figure and has landed somewhere inside the B-circle. Given that the dart is inside the B-circle, what is $P(A/B)$, the probability that the dart is also inside the A-circle? The answer is simple: it is the area common to both circles divided by the area of the B-circle. In other words, given that probabilities are proportional to areas, $P(A/B)$ is equal to $P(A\&B)$ divided by $P(B)$.

It is an immediate consequence of our definition of conditional probability that a general multiplication rule holds for the probability of any conjunction (or for the probability of the joint occurrence of any two events).

General Multiplication Rule $\quad P(A\&B) = P(A/B) \times P(B)$.

When $P(A/B) = P(A)$, A and B are said to be statistically independent of one another and the general multiplication rule simplifies to the special multiplication rule.

Special Multiplication Rule $\quad$ When A and B are independent, $P(A\&B) = P(A) \times P(B)$.

It is only a short step from the general multiplication rule to Bayes's theorem. First, we note that, since $(A\&B)$ is logically equivalent to $(B\&A)$, it follows from the equivalence rule that:

$$P(A\&B) = P(B\&A).$$

Substituting, using the general multiplication rule, gives:

$$P(A/B) \times P(B) = P(B/A) \times P(A).$$

Rearranging the terms gives the simplest form of Bayes's theorem:

Bayes's Theorem $\quad P(B/A) = \dfrac{P(A/B) \times P(B)}{P(A)}$, where $P(A) > 0$.

Before trying to apply Bayes's theorem to scientific reasoning and the confirmation of theories by evidence, let us consider a simple example that illustrates its essential features. Imagine that a make of wheelchair, the Samson, is manufactured in just two plants in the United States. One factory is in Boston, the other in Chicago. The Boston plant makes four-fifths of all Samsons; the Chicago plant makes the rest. Of the Samsons manufactured in Boston, one-sixth have a special lightweight aluminum frame, whereas, three-quarters of the Samsons made in Chicago are of this type. You purchase a Samson wheelchair at an auction and discover that it has a lightweight aluminum frame. What is the probability that it was made in Boston?

Let A stand for having a lightweight aluminum frame, let B stand for being made in Boston, and let C stand for being made in Chicago. We want to calculate $P(B/A)$ using Bayes's theorem. The information we are given is that $P(B) = 4/5$, $P(C) = 1/5$, $P(A/B) = 1/6$, and $P(A/C) = 3/4$. Bayes's theorem tells us that

$$P(B/A) = \frac{P(A/B) \times P(B)}{P(A)}.$$

Thus, the numerator equals $1/6 \times 4/5 = 2/15$. But what about the denominator, $P(A)$? Samsons are made in only two places—Boston and Chicago—so the alternatives are $(A\&B)$ and $(A\&C)$, which are mutually exclusive and exhaustive. Thus, by axiom 3,

$$P(A) = P[(A\&B) \vee (A\&C)] = P(A\&B) + P(A\&C),$$

and then, using the general multiplication rule, we get:

$$\begin{aligned} P(A) &= P(A/B) \times P(B) + P(A/C) \times P(C) \\ &= (1/6) \times (4/5) + (3/4) \times (1/5) = 17/60. \end{aligned}$$

Thus:

$$P(B/A) = (2/15) \times (60/17) = 8/17.$$

When using Bayes's theorem, we need not calculate the denominator, P(A), from first principles every time. Instead, we can write down the answer immediately using the total probability rule:

Total Probability Rule $P(A) = P(A/B) \times P(B) + P(A/{\sim}B) \times P({\sim}B)$.

In the wheelchair example, there were just two exclusive alternatives: either the chair came from Boston (B) or from Chicago (C). So, in this example, we can calculate P(A) by substituting C for ~B in the total probability rule:

$$P(A) = P(A/B) \times P(B) + P(A/C) \times P(C).$$

But it is easy to imagine a more complicated example in which aluminum chairs were also made in Detroit (D) and Evanston (E). In this case, the right-hand side of the equation for P(A) will include two extra factors:

$$P(A/B) \times P(B) + P(A/C) \times P(C) + P(A/D) \times P(D) + P(A/E) \times P(E).$$

More generally, when $B_1, B_2, \ldots, B_n$ are n mutually exclusive and exhaustive hypotheses, we can express P(A) as the sum of products as follows.

$$P(A) = P(A/B_1) \times P(B_1) + P(A/B_2) \times P(B_2) + \ldots + P(A/B_n) \times P(B_n).$$

This can be written more succinctly as:

$$P(A) = \sum_{i=1}^{i=n} P(A/B_i) \times P(B_i).$$

BAYES'S THEOREM AND SCIENTIFIC REASONING

The wheelchair example used above to illustrate Bayes's theorem can be regarded as a simple analogue of scientific reasoning. Let the theory, *T*, be that our Samson wheelchair was made in Boston. Initially, before we discovered that the chair is made of aluminum, the probability that it came from Boston was 4/5, since 80 percent of all Samson wheelchairs are made there. In other words, in our example the prior probability of theory *T*, P(T), is 4/5. Once we acquire evidence, *E*, that the chair is made of aluminum, we calculate that the posterior probability of *T* given evidence *E*, P(T/E), is 8/17. So the probability of *T* has dropped, and *E* disconfirms *T*. (In fact, in this example, it is now slightly more probable than not that the chair came from Chicago rather than Boston.)

In light of this example, we can summarize the application of Bayes's theorem to scientific theories as follows:

Bayes's Equation (Version 1) $$P(T/E) = \frac{P(E/T) \times P(T)}{P(E)}.$$

P(T) is the prior probability of *T*, P(E) is the probability of the evidence *E* (what Salmon calls the *expectedness of the evidence*), and P(T/E), the probability of *T* conditional on *E*, is the posterior probability of *T*. The meaning of these terms is pretty straightforward and obvious, but the terminology for P(E/T) can be misleading. P(E/T), the probability of *E* conditional on *T*, is usually referred to as the *likelihood* of *T* (or, as some authors prefer, the likelihood of *T* on *E*). To repeat, the likelihood of *T* is P(E/T), not P(T), and similarly, the likelihood of *T* on *E* is P(E/T), not P(T/E). To avoid possible confusion, we will avoid using the term *likelihood* as a synonym for *probability* and use it solely to refer to P(E/T) and similar expressions.[5]

When given some new evidence *E* for theory *T*, we revise our assessment of the theory's probability by using Bayes's equation to calculate P(T/E) as a function of P(T), P(E/T), and P(E). We then discard our old prior probability, P(T), and replace it with P(T/E). In this way, the posterior probability of *T* becomes our new prior probability:

$$P_{new}(T) = P(T/E).$$

Some authors, such as Salmon, explicitly include background knowledge, *B*, in Bayes's equation, thus making all the probabilities involved conditional. In the wheelchair example, *B* would include the information that the chairs are made either in Boston or Chicago and nowhere else. In real-life scientific reasoning, *B* would include the information about the world and other theories that scientists accept as true or highly probable. Including background knowledge, the equation becomes:

Bayes's Equation (Version 2) $$P(T/E\&B) = \frac{P(E/T\&B) \times P(T/B)}{P(E/B)}.$$

We can also use the total probability rule to expand the denominator as follows.

$$P(T/E\&B) = \frac{P(E/T\&B) \times P(T/B)}{P(E/T\&B) \times P(T/B) + P(E/{\sim}T\&B) \times P({\sim}T/B)}.$$

Since it is cumbersome to write out these equations when every probability is explicitly conditionalized upon the background information, we shall omit the reference to *B* whenever it is convenient. Thus, Bayes's equation can be expressed more simply as:

Bayes's Equation (Version 3)

$$P(T/E) = \frac{P(E/T) \times P(T)}{P(E/T) \times P(T) + P(E/{\sim}T) \times P({\sim}T)}.$$

And in its most general form, when $T_1, T_2, \ldots, T_n$ are mutually exclusive and exhaustive, we have:

Bayes's Equation (Version 4)

$$P(T_k/E) = \frac{P(E/T_k) \times P(T_k)}{\sum_{i=1}^{i=n} P(E/T_i) \times P(T_i)}.$$

Probabilities and Degrees of Belief

In order to apply Bayes's equation (in any of its four versions) to scientific reasoning, we have to decide what the probabilities that appear in it mean and how they should be measured. The pure Bayesian line is that, subject to some important conditions, all probabilities are subjective degrees of belief and that rationality requires us to revise our beliefs using Bayes's equation. The subjective (or so-called personalist) interpretation of probability and the alleged connection between rationality and Bayes's equation are controversial issues that are debated in the readings in this chapter. The purpose of this section is to explain and clarify the main issues involved in that debate.

The right-hand side of Bayes's equation (version 1) contains three sorts of probabilities: $P(T)$, $P(E)$, and $P(E/T)$—the prior probability of theory T, the expectedness of evidence E, and the likelihood of T. In many respects, likelihoods are the least problematic, since, even if $P(E/T)$ is a subjective degree of belief, it seems rational to base that probability on the objective relation between T and E. For example, if T is a deterministic theory that deductively entails E, then everyone agrees that the correct value to be assigned to $P(E/T)$ is one. Similarly, if T is a statistical theory, as in our wheelchair example, then T (together with background information B) will specify the probability that E is true on the condition that T is true.

Assigning values to $P(E)$ and $P(T)$ is more difficult. Later, in the section on Glymour, we discuss the so-called problem of old evidence, that is, the problem of assigning to $P(E)$ some value other than 1 when E is already known to be true. Of course, one might try to use the second or third versions of Bayes's equation to calculate $P(E)$, but as Salmon points out in his article, to do so we would need to calculate $P(E/{\sim}T)$, and since ${\sim}T$ is simply the negation of T, not a specific theory, we cannot infer from ${\sim}T$ the value of $P(E/{\sim}T)$. The fourth version of Bayes's equation avoids the indeterminate character of ${\sim}T$ but only at the price of requiring a complete list of all the possible theories that predict E (or that assign to E some definite probability). In practice, only a small handful of rivals to T (possibly none) will be candidates for serious consideration. Scientists simply do not know what all the logically possible rival theories are. For further discussion of this problem, see the articles by Salmon and Glymour and the sections on their articles later in this commentary.

Even if values for $P(E/T_i)$ were available for every single theory T_i that predicts E (either with certainty or some definite probability), the problem of determining the prior probabilities, $P(T_i)$, of each of these theories would remain. At this point subjectivist Bayesians say that the prior probability of a theory, T, is simply the actual degree of belief that a person has in T. Strictly speaking, on the subjectivist interpretation of probability, there is no such thing as *the* prior probability of T, since degrees of belief are relative to persons and it is perfectly possible that different people believe the same theory to different degrees, even given the same background information. According to subjectivist Bayesians, it is an illusion to think that there is one objectively correct answer to the question, "What is the prior probability of T?"

It seems incredible, on the face of it, that anything useful could be said about scientific inference and confirmation, starting from a basis as subjective as each person's degree of belief. In the remainder of this section, we take a brief look at some of the most important objections to the subjective interpretation of probability and the Bayesian replies to these objections. Many of these issues are explored more fully in Glymour's article and discussed later in this commentary.

Bayesians begin with degrees of belief. What are they, and how can they be measured so as to yield numerical probability values? The pioneers of Bayesian theory (Frank P. Ramsey, Bruno De Finetti, Leonard J. Savage) interpreted degrees of belief in terms of people's behavior. Ramsey proposed, for example, that a person's degree of belief in any proposition be measured by the least odds at which he would be willing to gamble on the proposition being true. In this way, we connect degrees of belief with something we can observe and measure, namely, betting behavior. But even when they are measured in this way, why should we think that the degrees of conviction that a person happens to have in various propositions qualify as probabilities by satisfying the probability axioms? Surely, people violate the axioms in many cases. For example, there might be a proposition Q, such that a person's degree of conviction in Q and his degree of conviction in ${\sim}Q$ do not add up exactly to 1.

The Bayesian response to this objection is the Dutch book argument. Professional gamblers say that a Dutch book has been made against someone if that person accepts a series of bets such that, no matter what the outcome, the person is guaranteed to lose money. No rational person would knowingly gamble in this way. The Dutch book theorem proves that a necessary and sufficient condition for avoiding a book being made against you is that your degrees of belief satisfy the axioms of probability theory. When this condition is satisfied, your degrees of belief are said to be coherent. Thus, when subjective Bayesians interpret probabilities as degrees of belief, a certain amount of idealization is involved. The degrees of belief in question are not necessarily the actual degrees of conviction that a particular person has, but rather the degrees of belief that she would have if she were ideally rational and her degrees of belief were coherent.

Bayesians view the coherence requirement as having the same status and rationale as the requirement of logical consistency. Rationality requires not only consistency with the laws of logic, but also consistency with the probability axioms. Both are necessary conditions for rational belief. What makes the Bayesian position thoroughly subjective is its insistence that coherence (which entails logical consistency) is not only necessary but also sufficient for rationality: no matter how crazy one's degrees of belief may seem to someone else, if they satisfy the probability axioms, then, according to the Bayesians, they cannot be condemned as irrational.

Suppose that Adam and Eve each have a different degree of conviction in the proposition, Q, that it will snow in Phoenix next July. Adam gives it a probability of 0.9, while Eve gives it only a 0.05 chance of being true. As long as Adam also assigns a probability of 0.1 to the proposition that it will *not* snow in Phoenix next July his degrees of belief conform to the probability axioms. Similarly, if Eve's degrees of belief are coherent, then she thinks it 0.95 likely that there will be no July snowfall in Phoenix next year.

As far as subjective Bayesians are concerned, both Adam and Eve are rational with respect to the extremely limited set of propositions consisting of Q and ~Q. But there is more to coherence than merely satisfying the negation rule. One also has to satisfy the special and general addition rules for disjunctions (p. 629), the special and general multiplication rules for conjunctions (p. 630), the implication and equivalence rules (p. 629), and axiom 2 concerning necessary truths (p. 628). Axiom 2 is especially problematic, since it requires that every necessary truth, no matter how complex, be assigned a probability of 1. Thus, while it might seem as though coherence is a very weak condition (because it places no restrictions on the degree of belief that a rational person can assign to any particular contingent proposition), in fact coherence makes very strong demands on the degrees of conviction that can be assigned to the members of any reasonably sized set of propositions (where that set includes many contingent statements, many necessary statements, and all their truth-functional compounds). Moreover, every proposition in the set must be assigned a precise number in the interval from 0 to 1.

Even when we have a coherent set of degrees of belief, there are several tricky issues connected with $P_{new}(T) = P(T/E)$, the Bayesian conditionalization rule that we are supposed to use in order to learn from experience. The first of these problems concerns the prior probability of T, $P(T)$, that appears on the right-hand side of the Bayesian expression for $P(T/E)$. Suppose one were to assign a prior probability of 0 to the theory,

$P(T) > 0$

T. In that case, no amount of evidence could ever confirm it, since its posterior probability, $P(T/E)$, would always remain 0. It is important to realize that the coherence requirement does not solve this problem. Coherence requires that we conform to the probability axioms. The relevant axiom (axiom 2) dictates that all necessary truths have a probability of 1. In conjunction with the other axioms, this entails that all necessary falsehoods must have a probability of 0. But the axioms do not forbid us from also assigning a probability of 0 to any contingent proposition that we might judge to be impossible. The usual Bayesian response is to impose the further demand of *strict coherence*: a set of beliefs is strictly coherent if and only if it is coherent and no contingent proposition is assigned a probability (degree of conviction) of 0 or 1. (This is one response to the Popperian argument discussed in the section "Why All Theories Are Improbable," in the commentary on chapter 1.)

Another problem with the dependence of $P(T/E)$ on $P(T)$ is that the amount by which evidence E confirms (or disconfirms) theory T will depend on the prior probabilities assigned to T by different scientists. How can this subjective influence on the degree of confirmation be reconciled with scientific objectivity? Bayesians reply by appealing to a theorem about the *washing out* (or *swamping*) *of priors*. The theorem shows that as evidence accumulates, the values of $P(T/E)$ calculated by individual scientists with different prior probabilities will tend to converge. In the long run, the initial divergence in the (subjective) prior probabilities becomes irrelevant and (objective) evidence dominates the calculation of confirmation.

Finally (for now), there is the issue of motivating the Bayesian conditionalization rule.[6] Why should we revise our degrees of belief in accordance with Bayes's formula for $P(T/E)$? Remember that rationality, for Bayesians, is supposed to begin and end with the requirement of coherence (or strict coherence) for one's beliefs at any given time. Why, then, is it irrational for someone to violate the Bayesian rule when revising her degrees of belief, so long as her entire set of beliefs remains coherent? If a Bayesian were later to look at that coherent belief set, in ignorance of how the person's degrees of belief had been arrived at, the Bayesian would judge the set rational. The synchronic Dutch book argument (for the coherence of beliefs at a given time) seems to have no relevance to the diachronic conditionalization rule (for how probabilities should change over time). One popular Bayesian response to this challenge is to construct a further Dutch book argument that is explicitly diachronic.[7] The gist of the argument runs as follows. If one makes a series of bets, some of which depend on what one's degrees of belief will be in the future, and if one follows a rule other than the Bayesian conditionalization rule and this alternative rule is known to the person with whom one is betting, then the person with whom one is betting can always construct a Dutch book against one. A necessary and sufficient condition for avoiding a Dutch book under the conditions stated is that one uses nothing but the Bayesian conditionalization rule for changing one's degrees of belief. Just as the synchronic Dutch book argument is used to derive the coherence condition from the presumption of rationality, so, too, the diachronic Dutch book argument is supposed to show how rationality mandates the Bayesian rule for revising one's degrees of belief over time.

5.2 | Salmon on Kuhn and Bayes

In a companion article, written at the same time as the piece in our book, Salmon relates that, when he first began reading Kuhn's *The Structure of Scientific Revolutions* (1962), he was so shocked by Kuhn's repudiation of the distinction between the context of discovery and the context of justification that he set the book aside without finishing it.[8] Later, in 1969, while preparing for a conference on the relation between the history of science and the philosophy of science, Salmon returned to Kuhn's book with renewed interest. Salmon conjectured that the reason for Kuhn's rejection of the traditional discovery-justification distinction lay in Kuhn's commitment to an inadequate conception of scientific justification, namely the H-D account. According to the H-D account, everything connected with the genesis of a scientific theory and its evaluation prior to being tested belongs to the context of discovery and, as such, is irrelevant the theory's epistemic justification; a theory is justified and its acceptance rationally warranted only when the theory has been confirmed, and a theory is confirmed if and only if the predictions deduced from it are observed to be true.

The Inadequacies of Hypothetico-Deductivism

As Salmon notes, a significant limitation of the H-D account of confirmation is that it ignores statistical theories. Statistical theories confine their predictions to assignments of probability to classes of events but do not logically imply that any particular event will occur. By regarding inductive confirmation, in effect, as the inverse of logical deduction, the H-D account excludes from its scope all those theories in which the relation between theory and evidence is not deductive but probabilistic. The Bayesian approach has no such limitation, since it permits the likelihood $P(E/T)$ to assume values less than 1.

In its simplest form, the H-D account seems committed to the view that any theory that logically implies an observational prediction, O, is as well confirmed by that prediction as any other theory that implies O. This flies in the face of common sense and scientific practice. Bayes's equation is attractive because it can do justice to the differential confirmation of rival theories by the same evidence, by appealing to differences in the

initial plausibility of those theories (and hence differences in their prior probabilities). A nice illustration of this is the Bayesian solution to the so-called tacking paradox or the problem of irrelevant conjunction. Suppose that theory T, in conjunction with background information B, entails the true observational prediction E: $(T\&B) \rightarrow E$. Now let I stand for some contingent statement that is logically independent of T and irrelevant to E. It follows trivially that if $(T\&B) \rightarrow E$, then it must also be the case that $(T\&I\&B) \rightarrow E$. Thus, according to the H-D account, E confirms both the original theory T and the augmented theory $(T\&I)$ and, moreover, confirms them by the same amount. The Bayesian analysis agrees with the H-D account that E confirms both theories but disagrees about the degree of confirmation. On the most popular version of the Bayesian analysis, the degree of confirmation of a theory by evidence is a function of the difference between the posterior probability of the theory given that evidence and the prior probability of the theory.[9] On this version, the Bayesian analysis entails that the degree of confirmation conferred by E on $(T\&I)$ must be less than the degree of confirmation that E confers on T alone. Here are the two expressions for the degrees of confirmation of T and $(T\&I)$ on the difference analysis.

$$P(T/E\&B) - P(T/B) = P(T/B) \times \left[\frac{1 - P(E/B)}{P(E/B)}\right].$$

$$P(T\&I/E\&B) - P(T\&I/B) = P(T\&I/B) \times \left[\frac{1 - P(E/B)}{P(E/B)}\right].$$

These expressions are derived from the second version of Bayes's equation, setting $P(E/T\&B)$ and $P(E/T\&I\&B)$ both equal to 1. The factor in the square brackets is the same for both theories, and their respective degrees of confirmation are proportional to their prior probabilities. Since $(T\&I) \rightarrow T$, and I is a contingent statement that is independent of T, it follows from the implication rule that $P(T\&I/B)$ must be less than (or equal to $P(T/B)$. Thus, E confirms $(T\&I)$ by a smaller amount than it confirms T: adding the irrelevant conjunct I to T lowers the confirmation provided by E. In this respect, then, the Bayesian approach to confirmation is a decided improvement over the H-D account.

Although Salmon does not discuss it in his article, the Bayesian approach also promises a resolution of the Duhem problem (that is, the problem of assigning the blame for a failed prediction to a particular member of a group of hypotheses), for not only are some theories confirmed better by the same evidence, but Bayes's equation can also be used to explain how some components of a group of hypotheses and assumptions receive a much larger *dis*confirmation than other components when observations disagree with theoretical predictions.[10] Thus, the Bayesian approach can show us where the blame should be placed when a group of hypotheses and assumptions lead to a false prediction.

Salmon is not alone in thinking that the H-D account is inadequate. An early critic of hypothetico-deductivism, Popper rejected it because he denied the whole notion of inductive confirmation. (See Popper, "The Problem of Induction," in chapter 4 for details.) Other critics, such as Carnap and Reichenbach, accepted that confirmation is an essential part of scientific rationality but insisted that it should be understood in terms of Bayes's theorem. Carnap interpreted the probabilities in Bayes's theorem as a priori logical probabilities; Reichenbach construed them as empirical frequencies. Both were objectivists about probability. More recently, an entire school of statisticians and philosophers of science has arisen—the personalists or Bayesians—that interprets the probabilities in Bayes's theorem subjectively, as degrees of belief.

Salmon thinks that we can use Bayes's theorem to reconcile Kuhn's historical approach to understanding science with the logical empiricism of philosophers such as Carnap, Reichenbach, and Hempel. The key is to incorporate Kuhn's values—criteria for theory assessment such as consistency, simplicity, and fruitfulness—into the Bayesian equation that defines confirmation. Variation in the interpretation of these values and the emphasis placed on them can give rise to differing judgments about the prior probability of a theory. Thus, scientists can reach different but equally rational judgments about how well a theory is confirmed by a particular piece of evidence using the same algorithm (Bayes's equation) because they insert different inputs into that algorithm in the form of different judgments about prior probability. Salmon contrasts this with Kuhn's own suggestion in "Objectivity, Value Judgment, and Theory Choice" (reprinted in chapter 2, above) that scientists reach different conclusions because they use different algorithms. However, careful study of Kuhn's article reveals that when Kuhn talks about using different algorithms he really means inserting different subjective inputs into a Bayesian algorithm, so Kuhn is much closer to Salmon's position than his language might suggest.

Prior Probabilities

As Salmon emphasizes in his article, most of what is philosophically controversial about the Bayesian approach to confirmation depends on the interpretation of prior probabilities in Bayes's equation. Salmon distinguishes three such interpretations: the objective-logical, the objective-empirical, and the subjective. Salmon agrees that the objective-logical interpretation of Carnap and others, according to which probabilities are assigned a priori to all statements on the basis of a formal language and assumptions about the equiprobability of states of affairs, is hopelessly inadequate to the task of analyzing the probability of real-life scientific theories. That leaves the objective-empirical and the subjective interpretation.

In an earlier book, *The Foundations of Scientific Inference*, Salmon adopted Reichenbach's objective-empirical interpretation according to which probabilities are relative frequencies.[11] The basic idea is this: Every hypothesis is either true or false, but when a new hypothesis, H, is first proposed, we do not know which attribute (truth or falsity) it has. In order to make sense of the prior probability (or plausibility) of the single hypothesis, H, we have to place it in a reference class of similar hypotheses. Then, on the basis of past experience, we can see how often hypotheses in this class have turned out to be true. The ratio of the number of true hypotheses to the total number of hypotheses in the reference class is then taken to be the prior probability of H.

In rough outline, this procedure is supposed to be similar to the way frequency theorists handle the problem of the single case. A typical example is the problem of assigning a probability to whether a particular (asymmetrical) coin will show a head on its next toss. In the case of the coin, we estimate this probability by dividing the number of times the coin has come up heads by the number of times the coin has been tossed. Thus, very roughly, the frequency theorist would say that the probability of getting a head on the next toss of the coin is 0.55 if the frequency of heads converges to 0.55 as the number of tosses becomes ever larger. Applying this same approach to the prior probability of hypotheses is extremely difficult. Not least among these difficulties is the problem of specifying the appropriate reference class. What exactly does it mean to talk about hypotheses that are similar to H? Is it a matter of mathematical form, such as the use of inverse-square laws? And if so, why should the success of such laws in one domain of science (such as the study of gravity) make it more likely that they will succeed in another domain (such as the investigation of the strong force binding together particles in the atomic nucleus).[12]

Throughout his career, Salmon has been highly critical of the unfettered subjectivism of the pure Bayesian or personalist interpretation of probabilities as degrees of belief. To Salmon, scientific judgments about confirmation should not depend in any way on the prejudices, emotions, or mood swings of individual researchers. The answer, he thinks, lies in what he (following Abner Shimony) calls *tempered personalism*. Tempered personalism places constraints on prior probabilities that go beyond mere coherence. XX Since experience has taught us that scientists have been moderately successful in the past, no hypothesis advanced by a serious scientist should be given a prior probability that is either 0 or vanishingly small. But, again, since experience tells us even the most promising hypotheses in the past have sometimes turned out to be false, the prior probability of any new hypothesis should be fairly low. The notion of *success* invoked in this discussion is crucial for understanding Salmon's proposal. For Salmon believes that when we assign a prior probability to a new hypothesis, we are trying to estimate the correct, objective probability that the hypothesis will turn out to be successful, and by *successful*, Salmon means *true*. Thus, while it may seem as if Salmon is making significant concessions to Kuhn when he admits consistency, analogy, and professional scientific standing as factors that play a legitimate role in determining the plausibility of new hypotheses, in fact his conception of probability is still, at bottom, objective and frequentist. It is, for example, only because past experience has taught us that hypotheses advanced by cranks very seldom turn out to be true that we should assign them a negligibly small prior probability. As Salmon himself puts it, "prior probabilities . . . can be understood as our best estimates of the frequencies with which certain kinds of hypotheses succeed. . . . The personalist and the frequentist need not be in any serious disagreement over the construal of prior probabilities" (564).

?

One final but important point: Salmon readily admits that it "seems preposterous" (564) that plausibility judgments based on values such as simplicity and symmetry could result in exact numbers for prior probabilities. Like many advocates of Bayes's equation, Salmon appeals to the washing out or swamping of priors to argue that their exact value really does not matter. For as soon as evidence begins to accumulate, the values for the posterior probability of a hypothesis converge. In the long run, the particular values adopted for the prior probability become irrelevant (so long as we avoid the extreme values of 0 and 1). But this convergence argument assumes that different scientists agree on the likelihoods, an assumption that Salmon defends later in his article.

THE EXPECTEDNESS OF EVIDENCE

Methodologists of science commonly hold that a theory receives greater confirmation from the successful prediction of something surprising than from the prediction of something expected. This issue is addressed, in part, in chapter 4 under the guise of the debate over novel predictions. Because the right-hand side of Bayes's equation has $P(E)$, the probability of the evidence E, as its denominator, it follows that, other things being equal, the lower the value of $P(E)$, the greater the value of $P(H/E)$. Thus, the more unexpected the prediction, the greater its confirming power if it should turn out to be true. But what is $P(E)$, the expectedness of the evidence, and how can it be measured?

Salmon uses the total probability rule to express $P(E)$ in terms of prior probabilities and likelihoods, writing all the probabilities involved as conditional on background knowledge B, where B includes initial conditions, boundary conditions, auxiliary hypotheses, and other relevant information.

$$P(E/B) = P(E/T\&B) \times P(T/B) + P(E/{\sim}T\&B) \times P({\sim}T/B).$$

If T is a deterministic theory, then T (in conjunction with auxiliary hypotheses and assumptions) entails E. In such a case, the likelihood $P(E/T\&B)$ equals 1 and $P(E/B)$, the expectedness of the evidence, must be at least as great as $P(T/B)$, the prior probability of T. But assigning an exact number to $P(E)$ is not easy, since it involves knowing the value of the likelihood $P(E/{\sim}T\&B)$, a problem that Salmon addresses in his section on likelihoods.

A second difficulty with $P(E/B)$ that Salmon acknowledges is a version of the problem of old evidence. This is discussed at some length later in this commentary in the section "The Problem of Old Evidence." For the moment we merely note that Salmon thinks that, given his characterization of background information B, the objective value of $P(E/B)$ must always be 1. Since B includes all the details about the experimental setup and the instruments used to observe E, the objective probability that B will occur under those conditions (assuming that the system in question is deterministic) is 1. Thus, Salmon concludes that the expectedness $P(E)$ can only be a subjective probability, reflecting the degree to which a particular scientist finds E psychologically surprising. Given Salmon's hostility towards subjectivism, the conclusion that "expectedness defies interpretation as an objective probability" (566) is highly unwelcome. At the end of his article, Salmon suggests a way to avoid this and a similar problem with the likelihood $P(E/B\&{\sim}T)$, while still permitting objective comparisons among rival theories.

E?

LIKELIHOODS AND THE CATCH-ALL HYPOTHESIS

The main problem with likelihoods concerns the value of $P(E/{\sim}T\&B)$, which appears in the expression for $P(E)$. $P(E/{\sim}T\&B)$ is the probability that E is true given that theory T is false, but since ${\sim}T$ is not a specific theory, the corresponding likelihood is not well defined. Even when we have two competing theories, T_1 and T_2 (such as specific versions of the wave and particle theories of light), $P(E/{\sim}T_1)$ is not equal to $P(E/T_2)$. Although theories T_1 and T_2 are contraries, and thus T_1 entails ${\sim}T_2$ and T_2 entails ${\sim}T_1$, they are not contradictories; thus, ${\sim}T_1$ does not entail T_2, nor is T_1 logically equivalent to ${\sim}T_2$. It is possible that both T_1 and T_2 are false. Thus, if we write out the set of logically exclusive and exhaustive hypotheses, it will include not only T_1 and T_2, but also T_k, the so-called catch-all hypothesis. What is the catch-all hypothesis? Strictly speaking, it is not a single hypothesis at all but a lengthy disjunction of all the possible alternatives to T_1 and T_2, most of which we have never thought about. As Salmon says, trying to guess the ingredients of the catch-all would be like trying to predict the future of science. Even though some of these ingredient hypotheses entail E, this scarcely helps us to answer the question, What is $P(E/T_k)$, the likelihood of T_k? because T_k is the disjunction of *all* the possible alternatives, including those that do *not* entail E. And if we cannot answer this question, then we cannot calculate $P(E)$. Because Salmon regards the problem of calculating the likelihood of the catch-all as completely intractable, he proposes a method for choosing between theories that does not require the calculation of $P(E)$.

SALMON'S BAYESIAN ALGORITHM FOR THEORY PREFERENCE

Salmon agrees with Kuhn that theory choice in science is usually a comparative affair. Typically, the issue is not how well a particular piece of evidence confirms an individual theory, but how well that evidence favors one theory over its rivals. In any given scientific domain only a few theories—usually just two or three—will be competing for acceptance at any given time. Certainly the catch-all hypothesis is seldom a serious option. Thus, despite the intractability of calculating the likelihood of the catch-all and the expectedness of the evidence, the Bayesian approach can still reflect the realities of scientific practice if it can provide a comparative ranking of those hypotheses that are serious rivals. Salmon's proposal is that, in choosing between two theories, T_1 and T_2, on the basis of evidence E, we should compare the posterior probabilities $P(T_1/E\&B)$ and $P(T_2/E\&B)$. An attractive feature of this proposal is that, in forming the ratio of the posterior probabilities, the problematic term $P(E)$ cancels out.

$$\frac{P(T_1/E\&B)}{P(T_2/E\&B)} = \frac{P(E/T_1\&B) \times P(T_1/B)}{P(E/T_2\&B) \times P(T_2/B)}.$$

Assuming that T_1 and T_2 are the only candidates for serious consideration, Salmon's proposal is that, before the discovery of evidence E, scientists should prefer T_1 to T_2 if and only if the prior probability of T_1 is greater than the prior probability of T_2. After the discovery of E, scientists should change their preference from T_1 to T_2 if and only if the posterior probability of T_2 is greater than the posterior probability of T_1. It follows from the Bayesian expression for the ratio of the posterior probabilities that, after the discovery of E, scientists should prefer T_2 to T_1, if and only if

$$\frac{P(E/T_2\&B)}{P(E/T_1\&B)} > \frac{P(T_1/B)}{P(T_2/B)}$$

or, in other words, if and only if the ratio of the likelihoods is greater than the reciprocal of the ratios of the prior probabilities. Salmon refers to this as the *Bayesian algorithm for theory preference*.

Salmon's algorithm is both ingenious and attractive, but it also has its limitations and counterintuitive features. First, it should be clear that in "choosing" T_2 over T_1, we are not deciding to accept T_2 as true or well-

confirmed. We are merely saying that, relative to evidence E, T_2 is better confirmed than T_1. For all we know, T_2 might be extremely improbable and unworthy of acceptance. It is important to remember that, in comparing the posterior probabilities of T_1 and T_2, we are not calculating—nor, if Salmon's pessimism is correct, can we ever calculate—the degree of confirmation of either hypothesis.[13] Thus, the judgment resulting from Salmon's algorithm is relatively weak, since it merely asserts that evidence E supports one theory better than its rival. The degree of that support is left entirely undetermined.[14]

Second (as noted by Wade Savage, the editor of the volume in which Salmon's article first appeared) Salmon's algorithm cannot, in its present form, give us any rational guidance when, as often happens, the body of evidence for T_1 is different from the body of evidence for T_2. For, obviously, $P(E)$ cancels out only when the evidence, E, is the same for both theories. Similarly, Salmon's algorithm does not permit us to judge whether one piece of evidence confirms a theory better than another piece of evidence.

Third, as Salmon himself notes, when both theories are deterministic and, in conjunction with B, entail the evidence E, the ratio of their posterior probabilities given E reduces to the ratio of their prior probabilities. Thus, according to Salmon's algorithm, no amount of evidence can change our initial preference ranking for such theories. For deterministic theories, the likelihoods become irrelevant and the prior probabilities (influenced by Kuhn's criteria for theory choice) dominate completely. Anyone who is critical of the vagueness of Kuhn's criteria and the difficulty of weighing and comparing them is unlikely to be impressed by this as a demonstration of a rational algorithm underlying scientific decisions about theories.

Salmon's response to this third point is contained in sections 8 and 9 of his paper. In section 9, "Kuhn's Criteria," Salmon distinguishes three types of theoretical virtue: informational, economic, and confirmational. Salmon argues that two of Kuhn's criteria—scope and accuracy—fall outside the confirmational category and are thus irrelevant to the prior plausibility of theories. This reduces the task of making the basis of plausibility judgments more precise by narrowing the focus to Kuhn's remaining three criteria, namely, simplicity, consistency, and fruitfulness. ?

In section 8, "Plausible Scenarios," Salmon explains that, when they are first formulated, important scientific theories often have great difficulty in explaining some puzzling phenomenon. He gives as examples the difficulty the absence of detectable stellar parallax posed for the Copernican theory and the problems of giving a coherent account of the optical ether and the phenomenon of selective absorption faced by the wave theory of light. Salmon's point is that the original versions of these theories did not logically entail the phenomena they had difficulty explaining. Indeed, the probability of E (a puzzling phenomenon) given T (the theory in question)