Part VIII

STATISTICAL INFERENCE AND THEORY CHANGE

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193077 Published online by Cambridge University Press

The Appraisal of Theories: Kuhn Meets Bayes

Wesley C. Salmon

University of Pittsburgh

Can statistical inference shed any worthwhile light on theory change? For many years I have believed that the answer is "Yes." Let me try to explain why I think so. On my first reading of Thomas S. Kuhn's *The Structure of Scientific Revolutions* (1962) I was so deeply shocked at his repudiation of the distinction between the context of discovery and the context of justification that I put the book down without finishing it. By 1969, when a conference was held at the Minnesota Center for Philosophy of Science on the relations between the history of science and the philosophy of science, I had returned to *Structure* and formed the view that Kuhn's rejection of this fundamental distinction resulted from his adoption of an inadequate conception of scientific justification. It appeared that he saw scientific confirmation in terms of the traditional hypothetico-deductive (H-D) schema, according to which a scientific hypothesis (or theory) is confirmed by observing the truth of its logical consequences. More precisely, given a hypothesis T, some initial or boundary conditions I, and auxiliary hypotheses A, an observational consequence is deduced. If the consequence turns out to be true that fact constitutes at least a bit of support for T.

The H-D method seemed to me at the time (and still does) grossly inadequate as a schema for the characterization of scientific confirmation (Salmon 1967, chap. VII). Moreover, this opinion had been shared by a number of leading experts on the subject. In his (1949), Hans Reichenbach explicitly rejected the H-D method and advocated the use of Bayes's theorem as the proper schema. In his (1950), Rudolf Carnap adopted a logical interpretation of probability, the structure of which conforms to Bayes's theorem, as an explication of degree of confirmation. Personalists, such as L. J. Savage (1954), were so wedded to Bayes's theorem that they took its name and called themselves *Bayesians*. Karl Popper, who repudiated confirmation altogether, a fortiori rejected the H-D schema. The fact that Popper accepted the use of the same schema as modus tollens when the observational prediction turns out to be false in no way commits him to allowing confirmation when the observational prediction is true.

In my contribution (1970) to the proceedings of the Minnesota conference I adopted Bayes's theorem (see equation (1) below) as the basic schema for scientific confirmation, but without presupposing any particular interpretation of probability. Quite plainly, I argued, if Bayes's theorem is to be used as a schema for confirmation it is

PSA 1990, Volume 2, pp. 325-332

necessary to take account of the prior probabilities that appear therein. The most natural construal of the prior probabilities is as plausibility considerations. To anyone who, like Kuhn, does not think of confirmation in terms of Bayes's theorem, plausibility arguments seem completely heuristic, strongly suggesting that they belong to the context of discovery rather than the context of justification. Nevertheless, Kuhn makes a striking case for the thesis that plausibility considerations play an indispensable role in the choice among scientific theories. Therefore, he seems to conclude, the distinction between discovery and justification is misconceived.

It seems to me that Kuhn is quite right regarding the indispensability of plausibility arguments; however, from the standpoint of Bayes's theorem, these plausibility considerations belong squarely within the context of justification. They appear as explicit terms in the logical schema for the confirmation of scientific hypotheses. Thus, I argued, since we can locate plausibility considerations in the context of justification, there is no need to give up the distinction between the contexts of discovery and justification. In an era in which sharp distinctions had been dropping like flies, I did not succeed in generating much enthusiasm for this one.

The 1983 program of the Eastern Division of the American Philosophical Association contained a Symposium on the Philosophy of Carl G. Hempel. On that occasion I had the singular honor to share the platform with Kuhn and Hempel, the most distinguished representatives, respectively, of the historical approach to philosophy of science and of logical empiricism. On this occasion Kuhn chose to speak about rationality in science, a topic that he and Hempel had been discussing for some time. This symposium struck me as an appropriate opportunity to suggest that a bridge could be built between these two approaches to the philosophy of science — at least insofar as rationality is concerned — by means of Bayes's theorem. As in the Minnesota paper, my attention here was focused on the status of plausibility considerations.

Near the beginning of his paper, Kuhn expressed his appreciation to Hempel. "More than other philosophers of his persuasion, Hempel has examined my views in this area with care and sympathy: he is not one of those who suppose that I proclaim the irrationality of theory choice. But he sees why others have supposed so" (1983, p. 563). Indeed, in (1977) Kuhn had already expressed his dismay at the accusation of irrationality. Although he acknowledged remarks to the effect that theory choice involves persuasion and judgment and is not a matter of proof, and that theory choice goes beyond observation and logic, he insisted that these aspects do not make science irrational — that he did not make theory choice "a matter for mob psychology" (1977, pp. 320-21).

Kuhn's attitude toward rationality takes, as its point of departure, the supposition that mature physical science constitutes *the prime example* of a rational enterprise. If we want to understand the nature of rationality, it is better to examine the practice of that domain of science, in order to try to learn how that practice proceeds, than to lay down a priori conditions of rationality.

> What, I ask to begin with, are the characteristics of a good scientific theory? Among a number of quite usual answers I select five, not because they are exhaustive, but because they are individually important and collectively sufficiently varied to indicate what is at stake. First, a theory should be accurate: within its domain, that is, consequences deducible from a theory should be in demonstrated agreement with the results of existing experiments and observations. Second, a theory should be consistent, not only internally or with itself, but also with other currently accepted theories applicable to related aspects of nature. Third, it should

> have broad scope: in particular, a theory's consequences should extend far beyond the particular observations, laws, or subtheories it was initially designed to explain. Fourth, and closely related, it should be simple, bringing order to phenomena that in its absence would be individually isolated and, as a set, confused. Fifth — a somewhat less standard item, but one of special importance to actual scientific decisions — a theory should be fruitful of new research findings: it should, that is, disclose new phenomena or relationships previously unnoted among those already known. These five characteristics — accuracy, consistency, scope, simplicity, and fruitfulness — are all standard criteria for evaluating the adequacy of a theory. (Kuhn 1977, pp. 321-22).

Kuhn makes it quite clear that the foregoing criteria cannot be applied in any mechanical way. Different scientists may place different relative weights upon them, and, in any case, their application requires judgment. At times, moreover, the criteria may conflict with one another. Simplicity, for example, might conflict with accuracy or scope.

When we look at this list, it seems to me, at least two of the items stand out as having a direct bearing on prior probabilities. Simplicity is often invoked as a plausibility consideration in the physical sciences; a nice example, involving proliferation of fundamental particles, can be found in Harari (1983). In the behavioral sciences, however, simplicity is not always prized. For example, an archaeological explanation in terms of drought alone of the abandonment of a sizable dwelling at Grasshopper, Arizona, near the close of the 14th century would be rejected as too simple — as an oversimplification. Certainly the drought was part of the story, but many other factors are required in a satisfactory theory. Practicing scientists in any given domain must exercise their judgment regarding the degree of simplicity or complexity appropriate in their fields.

In his remarks about consistency, Kuhn states explicitly that he is not concerned solely with the internal consistency of a theory. Clearly, internal consistency is important and desirable, but that is not the whole story. Another major consideration is how well a given theory fits with what we already accept in related domains. On this criterion, Immanuel Velikovski's *Worlds in Collision* (1950) receives an extremely low rating (see Gardner 1957). In contrast, Coulomb's inverse square law of the electrostatic force scores well on this desideratum, in view of Newton's law of gravitation and the then accepted Euclidean structure of physical space.

Kuhn's criterion of consistency is closely related, it seems to me, to arguments by analogy that are invoked to establish plausibility. Let me give three examples, one each from the physical, biological, and behavioral sciences. In physics, Louis de Broglie's hypothesis of wave-particle duality for material particles on the basis of analogy with the same duality for light is a perfect example. In the biological sciences, inference to the carcinogenic nature of saccharin for humans on the basis of the results of experiments with rats constitutes a strong plausibility claim. In archaeology, ethnographic analogy is used to support plausible hypotheses about the use of an artifact in a prehistoric culture on the basis of the observed use of a similar artifact in an extant culture.

There can be little doubt that plausibility considerations are ubiquitous in the sciences; only their status can be open to question. As long as the H-D schema is used to characterize scientific confirmation, plausibility arguments are relegated to heuristics as an aid in the generation of interesting and promising hypotheses. Bayes's theorem, in contrast, shows that they play an indispensable role in the appraisal of theories.

Having discussed two of Kuhn's criteria, I must briefly remark on the remaining three — accuracy, scope, and fruitfulness. All of them can be construed in terms of likelihoods. Consider accuracy. Inasmuch as Kuhn refers explicitly to deducible consequences, the likelihood of the results, given the truth of the theory, is one. If there are results of many observations and experiments, the likelihood of an accurate agreement, if the theory were false, would be small. Similar remarks apply to scope and fruitfulness. If the theory were false it is unlikely that it would extend successfully beyond "the particular observations, laws or subtheories is was initially designed to explain," or that it would "disclose new phenomena or relationships previously unnoted among those already known." Accordingly, it appears that Bayes's theorem provides a logical rationale for the kinds of considerations Kuhn sees as guiding actual scientific practice.

During the discussion in the 1983 APA Symposium, a problem was raised concerning the likelihoods in Bayes's theorem. One useful way of formulating Bayes's theorem is

$$\Pr(T_1|E.B) = [\Pr(T_1|B)\Pr(E|T_1.B)] \,/\, [\Pr(T_1|B)\Pr(E|T_1.B) + \ldots + \Pr(T_k|B)\Pr(E|T_k.B)] \qquad (1)$$

where B is background knowledge, E is a new piece of evidence, and $T_1, \ldots, T_k$ is a mutually exclusive and exhaustive set of theories. In this equation, $\Pr(T_1|E.B)$ is the posterior probability of T_1, $\Pr(T_i|B)$ are prior probabilities, and $\Pr(E|T_i.B)$ are likelihoods. A likelihood is the probability that the evidence actually found would occur, given our background knowledge and the truth of the the theory T_i mentioned in that expression.

Likelihoods are sometimes unproblematic. Consider Galileo's observation of the phases of Venus. On the Copernican theory the probability of this evidence is one; on the Ptolemaic theory the likelihood is zero. Sometimes, however, they are quite problematic. Consider the fact that no annual stellar parallax was observed at the time of Galileo. According to the Ptolemaic theory stellar parallax does not exist; that is why it is not observed. According to the Copernican theory stellar parallax must occur; the fact that we do not observe it can only be explained on the supposition that the fixed stars are unspeakably distant from us — a blatantly ad hoc hypothesis. How probable is it that the Copernican theory is correct and that the stars are so far away? We can easily understand why people rated that likelihood low.

The cosmological problem during the scientific revolution brings out another problem associated with Bayes's theorem (1), namely, the fact that the enumeration of theories in the denominator of the right hand side must be exclusive and exhaustive. Typically, when we have competing theories they are incompatible with one another, but seldom are they exhaustive. In addition to the Copernican and Ptolemaic theories there are other possibilities, e.g., the Tychonic system, in which the earth is stationary, the sun and moon move around the earth, but the other planets move around the sun. On the Tychonic system, the likelihood of absence of observed stellar parallax is one; in addition, the likelihood of the full set of phases of Venus later observed by Galileo is also one.

The Ptolemaic, Copernican, and Tychonic systems are obviously not the only logical possibilities, so if we want the exhaustive enumeration required by equation (1) we must add some other possibilities. The only plausible candidate to complete the set is what Abner Shimony (1970) called "the catchall" — i.e., the hypothesis that says "none of the above." Looking again at equation (1), we can always take T_k to be the catchall. At any given stage of scientific investigation, the catchall is the disjunction of all of the hypotheses we have not yet conceived. What is the likelihood of any given piece of evidence with respect to the catchall? This question strikes me as utterly intractable; to answer it we would have to predict the future course of the history of science. No one is ever in a position to do that with any reliability.

In a recent paper (Salmon 1990) I have offered a solution to the problem of the likelihood on the catchall. Borrowing an idea often emphasized by Kuhn, we may consider what happens if, instead of trying to evaluate one theory in isolation, we make only a comparative appraisal of competing alternatives. The catchall, incidentally, is never one of those in the competition, for it cannot be considered a bona fide scientific theory. Suppose, then, that we have two theories, T_1 and T_2, that we wish to compare. In a way completely parallel to equation (1) we can write

$$\Pr(T_2|E.B) = [\Pr(T_2|B)\Pr(E|T_2.B)] / [\Pr(T_1|B)\Pr(E|T_1.B) + \ldots + \Pr(T_k|B)\Pr(E|T_k.B)] \qquad (2)$$

Comparing the two equations, we note that the denominators on the right hand sides are identical. Assuming that the neither of the numerators is zero — if either were zero we would have no interest in the hypothesis involved in it — we divide (1) by (2) with the result

$$\Pr(T_1|E.B) / \Pr(T_2|E.B) = [\Pr(T_1|B)\Pr(E|T_1.B)] \ / [\Pr(T_2|B)\Pr(E|T_2.B)] \qquad (3)$$

We note that the likelihood on the catchall, as well as its prior probability, have vanished. What remain are only the prior probabilities, posterior probabilities, and likelihoods of the theories we are explicitly comparing. If the ratio of the posterior probabilities is greater than one, we prefer T_1 to T_2; if the ratio equals one, we prefer neither to the other; if the ratio is less than one, we prefer T_2 to T_1.

Even though we have eliminated the most intractable probability that occurs in equations (1) and (2) — the likelihood on the catchall — we are still not in a position to say that the likelihoods are unproblematic. The likelihoods on the theories being compared may still pose difficulties. Remember the Copernican theory and the absence of observed stellar parallax. There is, I think, a rather common strategy for dealing with such cases. By the invocation of suitable auxiliary hypotheses A, one can construct a *plausible scenario* according to which the problematic observation is made into a likely, or even necessary, consequence of the theory T in conjunction with the auxiliary A. For the Copernican system, the plausible scenario places the fixed stars at a huge distance from the earth. The numerator on the right hand side of (3) becomes $\Pr(T_1.A_1|B)\Pr(E|T_1.A_1.B)$. Although the likelihood $\Pr(E|T_1.A_1.B) = 1$, we are left with the prior probability of the *plausible scenario* $\Pr(T_1.A_1|B)$ — namely, with the question of just how plausible the proffered scenario actually is.

Consider another example (see Worrall 1990). In the 1830s David Brewster, a distinguished British optician, still supported the corpuscular theory of light, and refused to adopt the by then widely accepted wave theory. Although he recognized the difficulty for the corpuscular theory in accounting for various diffraction phenomena such as the Poisson bright spot, he believed that the wave theory encountered problems that were equally difficult if not more so. For example, he did *not* find plausible the hypothesis that the universe is completely filled with an aether of suitable mechanical properties to transmit starlight across vast distances of apparently empty space. Moreover, a phenomenon that he discovered — the selective absorption of light passing through a gas — seemed to pose insuperable problems for the wave theory. If you believe that light is a wave phenomenon, then how could you explain why light of a given wavelength passes freely through a gas, while light differing only slightly in wavelength is almost completely absorbed by the same gas? (Brewster avoided talk of wavelengths by referring to indices of refraction.) At the same time, he had no plausible scenario to offer on behalf of the corpuscular theory to deal with such phenomena as the Poisson bright spot. From our late 20th century vantage point, we can see that both the corpuscular and the undulatory theories of the 19th century faced insuperable difficulties. Plausible scenarios were not available.

A contemporary example of the use of plausible scenarios can be seen in the case of dinosaur extinction. The discovery in 1979 of an iridium anomaly — an extraordinarily high concentration of iridium — at the Cretaceous-Tertiary (K-T) boundary near Gubbio, Italy, led to the postulation of a collision of an asteroid or comet with the earth 65 million years ago. This event, it has been claimed, coincided with the extinction of dinosaurs and many other living species. Walter Alvarez and his father, Luis W. Alvarez, advanced a scenario designed to explain how the extinction was caused by the impact (see Alvarez, et al., 1980). This hypothesis has generated an enormous amount of controversy; other scientists maintain, for example, that massive volcanic activity was responsible. They too have fashioned scenarios they find plausible. A historical project studying the development of this controversy is reported by William Glen (1989).

Possible scenarios are especially successful if they enable the theorist to deduce the evidence, thus making the likelihood one, and if they are, indeed, plausible. If this result is achieved for theories T_1 and T_2 by means of auxiliaries A_1 and A_2, equation (3) assumes the following simplified form,

$$Pr(T_1.A_1|E.B) / Pr(T_2.A_2|E.B) = Pr(T_1.A_1|B) / Pr(T_2.A_2|B) \tag{4}$$

in which the ratio of the posterior probabilities equals the ratio of the prior probabilities. Since theory choice does not hinge ultimately on plausibility considerations (prior probabilities) alone, new evidence will be sought to discriminate between the two plausible scenarios. If a *crucial experiment (or observation)* can be found, yielding a likelihood of zero on one of the scenarios and a likelihood of one on the other, eliminative induction can be successfully practiced (see Earman forthcoming, chap. 7).

Two major objections have often been raised against the foregoing account of theory appraisal. First, it has been questioned whether two scientists, differing in their preferences between two theories, always share the same background knowledge B. While it may be true that they sometimes differ, the problem does not arise for the individual scientist who is comparing the merits of two theories for herself or himself; both must be evaluated in light of the individual's total body of background information. Thus, we can reasonably assume that, for any given individual, "B" is univocal in equations (3) and (4).

Second, a far more difficult challenge has been raised regarding the status of the prior probabilities that occur in Bayes's theorem. At the outset of this discussion I mentioned the views of Carnap, Reichenbach, and Savage. As representatives of the logical, frequency, and personalist interpretations of probability, they differ regarding the nature of the priors. For Carnap, they are a priori measures. I do not see how it can reasonably be maintained that we have a priori prior probabilities for the hypotheses seriously considered in science. Reichenbach required that prior probabilities be related to classes of similar theories in terms of the relative frequency of success of hypotheses in such classes. Although Reichenbach never made a convincing case for frequencies in this context, he did, at least, insist that prior probabilities be related to scientific experience. For Savage and other personalists, prior probabilities are degrees of belief subject only to coherence requirements. The extreme subjectivism of this view has been a severe obstacle.

I am inclined to think that a compromise can be made between the frequentist and personalist approaches, along a line suggested by Patrick Suppes (1966, pp. 202-3). Suppes points out that we generally bring to any scientific hypothesis a heterogeneous body of information that cannot be stated explicitly in full. A judgment of prior probability gives at least a rough assessment of the way that knowledge applies to a given hypothesis. It is, to be sure, a subjective judgment, but it also reflects objective experience. It is important that any such prior probability assessment be made in the light of scientific experience, to the exclusion of idiosyncrasies, prejudices, ideologies, and emotions. Moreover, as personalists have often noted, it is unnecessary to have precise numerical values for the priors.

The moral to be drawn from the preceding discussion is twofold. First, the logical empiricists, it seems to me, should take seriously Kuhn's point that in most cases, if not always, scientists are concerned to make comparative evaluations among theories. Seldom, if ever, is a hypothesis judged in isolation from all potential competitors. Second, philosophers with a historical disposition should look closely at the nature of scientific confirmation. They should relinquish the H-D schema and consider the merits of Bayes's theorem as a more adequate account. In this way, I believe, can a significant degree of rapprochement be achieved between these two important schools of philosophy of science.

References

Alvarez, L.W., et al. (1980), "Extraterrestrial Cause for the Cretaceous-Tertiary Extinction", *Science* 208: 1095-1108.

Carnap, R. (1950), *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Earman, J. (forthcoming), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, Mass.: MIT Press/Bradford Books.

Gardner, M. (1957), *Fads and Fallacies in the Name of Science*. New York: Dover.

Glen, W. (1989), "Meteorites, Volcanoes and Dinosaurs: Update on a Historical Project on the Current Debates", *Center for the History of Physics Newsletter* 21: 2-3.

Harari, H. (1983), "The Structure of Quarks and Leptons", *Scientific American* 248, no. 4 (April): 56-68.

Kuhn, T.S. (1962), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

______. (1977), "Objectivity, Value Judgment, and Theory Choice", in *The Essential Tension*. Chicago: University of Chicago Press, pp. 320-39.

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193077 Published online by Cambridge University Press

______. (1983), "Rationality and Theory Choice", *Journal of Philosophy* 80: 563-72.

Popper, K.R. (1959), *The Logic of Scientific Discovery.* New York: Basic Books.

Reichenbach, H. (1949), *The Theory of Probability.* Berkeley & Los Angeles: University of California Press.

Salmon, W.C. (1967), *The Foundations of Scientific Inference.* Pittsburgh: University of Pittsburgh Press.

________. (1970), "Bayes's Theorem and the History of Science", in *Historical and Philosophical Perspectives of Science,* vol. 5, *Minnesota Studies in the Philosophy of Science,* Roger H.Stuewer (ed.). Minneapolis: University of Minnesota Press, pp. 68-86.

________. (1983), "Carl G. Hempel on the Rationality of Science", *Journal of Philosophy* 80: 555-62.

________. (1990), "Rationality and Objectivity in Science, *or* Tom Kuhn Meets Tom Bayes", in *Scientific Theories* vol. 14, *Minnesota Studies in Philosophy of Science,* C. Wade Savage (ed.). Minneapolis: University of Minnesota Press, pp. 175-204.

Savage, L.J. (1954), *The Foundations of Statistics.* New York: John Wiley & Sons.

Shimony, Abner (1970), "Scientific Inference", in *The Nature and Function of Scientific Theories,* Robert G. Colodny (ed.). Pittsburgh: University of Pittsburgh Press, pp. 79-172.

Suppes, Patrick (1966), "A Bayesian Approach to the Paradoxes of Confirmation", in *Aspects of Inductive Logic,* Jaakko Hintikka and Patrick Suppes (eds.). Amsterdam: North-Holland Publishing Co., pp. 198-207.

Velikovski, Immanuel (1950), *Worlds in Collision.* Garden City, NY: Doubleday.

Worrall, John (1990), "Scientific Revolutions and Scientific Rationality: The Case of the Elderly Holdout", in *Scientific Theories* vol. 14, *Minnesota Studies in Philosophy of Science,* C. Wade Savage (ed.). Minneapolis: University of Minnesota Press, pp. 319-54.