A Classic of Bayesian Confirmation Theory

Finnur Dellsén

Penultimate draft – forthcoming in *Metascience*

Book review of Paul Horwich. *Probability and Evidence* (Cambridge Philosophy Classics edition). Cambridge: Cambridge University Press, 2016, 147pp, £14.99 (paperback).

Paul Horwich's *Probability and Evidence*, originally published in 1982, is a classic work in the approach to scientific reasoning known as 'Bayesian Confirmation Theory'. It is therefore fitting that the book has now been reissued as part of the series *Cambridge Philosophy Classics*. The text is identical in content to the 1982 original, apart from the addition of a new (albeit brief) foreword by Colin Howson, another well-known proponent of the Bayesian approach. Horwich's book contains clear, inventive, and stimulating treatments of a whole host of interesting topics in confirmation theory, such as the value of diverse data, the paradox of confirmation (i.e. the raven paradox), and the value of gathering further evidence. Some of Horwich's solutions to these problems are still among the most influential in the current literature, and rightly so.

Probability and Evidence has been less influential in its treatment of the foundations of Bayesian Confirmation Theory. Horwich's discussion of this topic in a lengthy chapter 2 owes much to previous authors, such as Ramsey (1926) and Carnap (1950), and his own contributions do not add much to the Bayesian framework. Furthermore, to a contemporary reader familiar with recent work on foundational issues in Bayesianism, this discussion will seem more dated than the rest of the book. For example, at one point (p. 30) Horwich tells us that he does not subscribe to Bayesian Conditionalization, which dictates that when one learns *E* (and nothing else) one's subjective probability in a hypothesis *H* should equal the conditional subjective probability one had in *H* given *E* right before one learned *E*. This rule is now often considered to be at the heart of Bayesian Confirmation Theory, but Horwich's discussion of it is limited to just two paragraphs (pp. 29-30). Moreover, part of

Horwich's justification for dismissing Bayesian Conditionalization is that "there is no Dutch book rationale for conformity to it." (p. 30) This claim has not aged well, since *diachronic* Dutch book arguments for Bayesian Conditionalization are now among the most widely discussed topics in the literature on Bayesianism.

In a review of the original 1982 version, Woodward (1985: 214) said that "[t]he principal defect of *Probability and Evidence* is its unsystematic character." It's hard not to agree with Woodward's claim here. To be fair, Horwich does not seem overly concerned with presenting a systematic account of scientific reasoning; rather, his focus is primarily on using different bits of Bayesian machinery to solve or dissolve various paradoxes of scientific reasoning. Indeed, Horwich claims to be taking a Wittgensteinian approach in which the goal "is not to formulate and defend a theory that will dictate which of the contradictory propositions is true; but rather to discover in ourselves the sources of our conflicting intuitions." (p. 2) However, as Howson in effect notes in his new preface (p. ix), Horwich does seem to formulate and at least partly defend a grand philosophical theory in the book, viz. Bayesian Confirmation Theory. Indeed, since Horwich's treatments of the aforementioned problems assume that the Bayesian theory is at least roughly correct, it is hard to see how they could have much value for someone who rejects it outright.

However Horwich's claims about his Wittgensteinian approach are to be interpreted, one would at least hope that the positions defended in the book do not conflict with one another. But that isn't always clear. In particular, Horwich's discussion of the foundations of Bayesianism sometimes seems inconsistent with the assumptions he makes in his solutions to the various paradoxical problems. Let me illustrate by returning to Horwich's denial of Bayesian Conditionalization. At one point in his treatment of the (supposed) additional evidential value of prediction over accommodation, Horwich says that "our degree of belief in *H* should increase, with the discovery of D, from $P_E(H)$ to $P_E(H/D)$ – where $P_E(q)$ represents our beliefs before the experiment." (p. 102) Horwich provides no special reason why we should update our degrees of belief in this way in this particular case, so he seems to be presupposing a general rule according to which one's degree of belief in a hypothesis after learning something should equal one's prior conditional degree of belief in the hypothesis given what one learned. This is Bayesian Conditionalization.

As I have indicated, the strongest and most influential parts of *Probability and Evidence* are Horwich's solutions to various puzzles about scientific reasoning. As a case in point, consider Horwich's solution to the paradox of confirmation. *Prima facie*, it's plausible that a generalization of the form 'All As are Bs' is confirmed by a positive instance, i.e. an A that is a B – this is known as *Nicod's criterion*. For example, it seems that the hypothesis that all ravens are black is confirmed by a black raven. Next, note that it is surely also true that if some evidence confirms a proposition, then the same evidence also confirms any logically equivalent proposition – this is known as *the equivalence condition*. But now note that 'All As are Bs' is logically equivalent to 'All non-Bs are non-As'. So by the equivalence condition, anything that confirms the latter confirms the former. By Nicod's criterion, substituting 'A' for 'non-B' and B for 'non-A', we thus end up with the result that a non-B that is also a non-A confirms that all As are Bs. For example, a red shoe would confirm that all ravens are black.

Horwich's approach to the paradox builds on earlier treatments from the Bayesian literature, which generally seek to show that instances of non-black non-ravens (e.g. red shoes) confirm the hypothesis that all ravens are black to a *far lesser extent* than instances of black ravens. However, Horwich's solution is novel in that it introduces a distinction that is arguably quite helpful in solving the problem. As Horwich notes, *how* one gathers some data may make a difference to its evidential significance. (p. 54-55) Horwich thus distinguishes between (RB) discovering that something picked out at random is both a raven and black, and (R*B) discovering, of something known to be a raven, that it is black. There is a corresponding distinction between (~B~R) discovering that something picked out at random is both non-black and a non-raven, and (~B*~R) discovering, of something known to be non-black, that it is a non-raven. Horwich argues that that there is no reason for Bayesians to think that RB confirms the raven hypothesis (that all ravens are black) to a greater extent than ~B~R. However, Horwich argues that the situation changes once we compare the confirmational import of R*B and ~B*~R, in that although ~B*~R does provide some slight confirmation of the raven hypothesis, R*B confirms it to a far greater extent.

I have briefly touched on Horwich's solution to the paradox of confirmation as an example of *Probability and Evidence*'s ingenious and influential treatments of various puzzles about scientific reasoning. These parts of the book justify its status as a Bayesian classic. However, the reissuing under review here has some substantial errors not present in the original. In addition to some minor typographical issues (for example, the logical sign for 'and' ($\wedge$) has sometimes, but only sometimes, been replaced by an upside down V), there are a few more serious errors that are likely to confuse many readers. On p. 50, a type of comparative confirmation – 'E confirms H more than F does' – is defined in terms of the inequality $P(H/E) > P(H)$. However, this cannot be right since *F* does not even occur here; indeed, one finds that the 1982 original, Horwich has written the inequality as $P(H/E) > P(H/F)$ – which makes a great deal more sense. Another issue is that Horwich sometimes refers to other parts of the book using its page numbers, but these don't seem to have been updated from the original. So, for example, there is a reference on p. 30 to a discussion we are told is on pp. 74-77, but that discussion turns out to be on pp. 69-72 in the reissued version of the book.

Despite these errors introduced by the new version, reissuing Horwich's book as part of Cambridge Philosophy Classics is a welcome event. In making *Probability and Evidence* more widely available at a very reasonable price, it will benefit advanced students and scholars with interests in scientific reasoning and confirmation theory. Indeed, as Howson notes in his preface (p. ix), Horwich's simple and elegant writing style "belies the often-challenging nature of the material." Horwich has purposefully written the book with a minimum of formalism, and he assumes no knowledge of mathematics beyond basic arithmetic (although knowing some probability theory will not hurt). For this reason, the book might very well still be of considerable interest to those who are looking for an engaging and readable introduction to the topic of scientific reasoning from a Bayesian perspective.[1]

References

Carnap, Rudolf 1950. *Logical Foundations of Probability*. Chicago: University of Chicago Press.

[1] I am grateful to Chris Dorst for helpful feedback on a draft of this review.

Ramsey, Frank P. 1926. Truth and Probability. In Frank P. Ramsey, *The Foundations of Mathematics and other Logical Essays*, ed. Richard B. Braithwaite, 156-198. London: Kegan, Paul, Trench, Trubner & Co., 1931.

Woodward, James 1985. Review: Paul Horwich, Probability and Evidence. *Erkenntnis* 23: 213-219.