Bayesian probability: A defense

By DOUG SHAVER
March 2010

The Bayesian subjective interpretation of probability attempts to formalize our use of ostensible evidence in justifying beliefs based on that evidence. This paper will assess the interpretation using the conventional criteria of admissibility, ascertainability, and applicability. Using the results of that assessment, I will argue that Bayes' theorem ought to play a central role in any practical epistemology.

Bayes' theorem may be regarded as an expansion of the elementary conditional probability formula $P(A|B) = P(A \wedge B)/P(B)$, where A is an event whose likelihood of occurring is affected by the occurrence of B. For a subjective interpretation, the likelihood or probability of an occurrence is assumed to measure our degree of belief, or credence, that it has occurred or will occur.[1] It seems reasonable to assume that we are concerned with making our credences rational, or in other words that we should wish to make them correspond, as best we can make them correspond, to whatever actual probabilities obtain in whatever portion of the real world constitutes our universe of discourse.

For epistemological purposes, we are interested in what Bayes' theorem might tell us about the reasonableness of inferring a particular hypothesis, h, from some fact that we think is evidence, e, for that hypothesis. In that case we are interested in the probability of that hypothesis given the evidence, or $P(h|e)$. In its simplest form, Bayes' theorem yields

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

[1]D. H. Mellor, *Probability: A Philosophical Introduction* (New York: Routledge, 2005), 7.

But if it's justification we're after, we need to derive the quantities on the right side somehow, and to that end Carrier proposes the following expanded version with *b* for background information:

$$P(h \mid e \wedge b) = \frac{P(h \mid b)P(e \mid h \wedge b)}{P(h \mid b)P(e \mid h \wedge b) + P(\sim h \mid b)P(e \mid \sim h \wedge b)}\,.$$ [2]

This formulation forces us to include in our analysis some estimate of the likelihood that our hypothesis is wrong notwithstanding our evidence and background information. As Carrier notes, the right side is of the form $A/(A + B)$, where A is the outcome of interest and $(A + B)$ represents all possible outcomes, since h and $\sim h$ are exhaustive and mutually exclusive.[3]

Since Bayes' theorem is entailed by formulas that are not themselves controversial, its admissibility seems uncontroversial as well. And, given admissibility and ascertainability, applicability seems obviously to follow, provided that our ascertained subjective probabilities correspond at least approximately to objective real-world probabilities. If we can achieve that correspondence, then an assessment of Bayes' theorem will rest on its ascertainability. A thorough argument for this achievability cannot be attempted in this essay, but I will offer some suggestions for why we may think it is possible.

A major challenge is making sense of the probability of an unknown outcome that has already occurred, such as the unexamined tossed coin. The probability that it landed heads is in fact either zero or one. Epistemically, though, our situation is indistinguishable from what it was before the coin was tossed, at which time we felt securely justified in supposing that the probability of its landing heads was 0.5. The pertinent fact of our epistemic situation in either case, before or after the toss, is our ignorance. Before the toss, we do not know what will happen, and after the toss we do

---

[2]Richard C. Carrier, "Bayes' Theorem for Beginners: Formal Logic and Its Relevance to Historical Method and Tutorial," (2008), 3, http://www.richardcarrier .info/CarrierDec08.pdf.

[3]Carrier, 27.

not know (before we look) what did happen, but in each case we have no reason to expect one outcome rather than the other: No more reason to affirm “It will land (or landed) heads” than to affirm “It will land (or landed) tails.” This is so even under determinism, where in principle the outcome is as fixed pre-toss as post-toss. There is thus a useful sense in which probabilities are just a measure of our ignorance.

The measurement is clearly not direct, but there is a relationship. To say there is a 1/6 probability of getting an ace with one throw of a die but a 1/2 probability of getting heads with one throw of a coin is not exactly to say we are three times as ignorant in one case as the other, but it does have something to do with there being three times as many possible outcomes when throwing a die as when throwing a coin and with our having no reason in either case to more confidently expect any one of the possible outcomes than any other.

The role of ignorance in probability assessments sheds some light on the Monty Hall Paradox: The host’s opening of one unchosen door gives us information that we did not have when we made our original selection. If we chose door A, the discovery that door C does not contain the prize—combined with our knowledge that the host (presumably) knew, before he opened the door, that it did not—gives us a reason to think it is more likely behind door B than door A. Objectively, it either is or is not behind door B, but having lost some of our ignorance, we now have some reason to choose B in preference to A. Or do we, really? Any interpretation of probability is in some way related to the notion that if we played the game an indefinite number of times, we would win the prize twice as often if we switch every time than if we don’t. Assuming we get to play only once, though, it is not entirely clear what difference this ought to make, especially since, if we’re like most people, our intuition contradicts the formal mathematics, insisting that doors A and B both have probability 1/2 of concealing the prize. Let us say that 1/2 is our intuitive credence and 2/3 is a rational credence. Why should we prefer the latter?

We should prefer it because, never minding how often we play “Let’s Make a Deal,” having rational credences is a good way of making decisions in general, assuming we make those decisions based on what we expect their outcomes to be. If, faced with a choice between any A and B, we habitually choose whichever has the higher actual probability of a favorable outcome, then, all else being equal, we can reasonably expect to get those favorable outcomes more often than not. Granted that we are often in no position to know the actual probabilities or whether all else is equal, a habitual preference for basing our decisions on available data, however meager relative to what we wish we had, seems to be the only rational option.

Back to the coin. Suppose the coin had already been tossed nine times and came up tails every time. I might succumb to the gambler's fallacy and be inclined to wager a large sum of money that it came up heads this time. I would have a credence that $P(H) > 0.5$, but this would not be a rational credence, particularly if I were familiar with the relevant statistical principles. Given only two possible outcomes, and no logical reason to be more confident of one than the other, the only rational credence I can have is $P(H) = 0.5$. On any particular occasion, I might get away with relying on the gambler's fallacy and win a large sum of money, but if I do it habitually, it is a near certainty that I will not like the results.

Now, we stipulated a fair coin. Suppose it was not, but that I had no good reason to think it was unfair. I believed, for whatever reason was sufficient for me in that situation, that it was an ordinary coin. If it was actually a two-headed coin, then $P(H)$, both prior to and after the toss, was 1.0. Given my ignorance, though, it was still rational for me to believe $P(H) = 0.5$. No reasonable theory of justification can require us to believe or disbelieve anything on the basis of information not at our disposal. Probability in this case is, again, somehow measuring my ignorance, but it does that, as we saw, even in the case of a fair coin. Probability is thus a kind of function of a relationship between what we know and what we don't know, and a rational credence depends on an accurate assessment of that relationship.

Similar reasoning applies when we consider any set of n outcomes that are equiprobable under the classical interpretation. For throwing a die, $n = 6$, and to say that all six are equiprobable is just to say that until we observe the actual outcome, then for any two of the six, we should have no more confidence in one than the other.

Complications arise when we try to extend this notion of credence to events, observed or unobserved, that are not obviously members of some set analogous to outcomes of coin tosses or die throws. Our beliefs about historical events are an example. “Buffalo Bill” Cody claimed in his autobiography that he was a rider for the Pony Express, and most Americans have taken his word for it. Some historians, noting a lack of corroborating evidence where they thought there ought to be some, doubt his claim. That is to say, if R = “Cody was a Pony Express Rider,” they think $P(R) < P(\sim R)$ to some significant degree of difference. But any coin toss has an obvious reference class from which we get a relevant frequency distribution. To what class does Cody’s claim to have been a Pony Express rider belong? Writers of autobiographies? Pony Express riders in general? Owners of Wild West shows? A definitive response, assuming one is possible, would need insights not just from probability and induction theories but also from historiography and the philosophy of history. We cannot discuss them here, but let’s proceed on the assumption that a resolution of the disputes is at least possible.

In objective historical fact, either $P(R) = 1$ or $P(R) = 0$. There are, or at one time were, certain facts that would settle the matter if we had them. Certain employees of Russell, Majors and Waddell, if their testimony had been obtained, could have told us one way or the other. Evidence of that sort is apparently irretrievable, though, and so of those facts, we are ignorant, and so we cannot justify a rational credence of either $P(R) = 1$ or $P(R) = 0$, since the first would correspond to a feeling of certainty that he did and the second, certainty that he did not. We have some reasons to think he did, and we have some reasons to think he did not, and we assign $P(R)$ a value closer to one or zero depending on which reasons we find more persuasive. What Carrier seems to be suggesting is that a sufficiently rigorous application of Bayes’ theorem can help us discern the logical consistency with which we perform our evaluation of the relevant facts at our disposal. I think his suggestion is a good one.

References

Carrier, Richard C. “Bayes’ Theorem for Beginners: Formal Logic and Its Relevance to Historical Method and Tutorial.” (2008). http://www.richardcarrier.info/CarrierDec08.pdf.

Mellor, D. H. *Probability: A Philosophical Introduction*. New York: Routledge, 2005.

(The original version of this essay was written as an assignment for a class in inductive logic taught by Dr. Susan Finsen at Cal State San Bernardino.)

Back to site home.

(This essay was last updated on June 1, 2012.)