Is Epistemic Probability Pascalian?

James B. Freeman
Hunter College of The City University of New York

ABSTRACT: What does it mean to say that if the premises of an argument are true, the conclusion is probable or to indicate the relation between premises and conclusion in a non-deductive argument as epistemic probability? Authors of introductory logic texts which present basic probability theory in connection with inductive logic assume that epistemic probability satisfies the laws of the Pascalian probability calculus. However, there are strong arguments against taking the logical, Bayesian personalistic, or limiting frequency interpretations of probability as properly explicating epistemic probability. After reviewing one argument against the logical interpretation, we shall explore whether the propensity interpretation, when supplemented by the non-Pascalian concept of an argument's weight, gives an adequate account of epistemic probability for at least one type of non-deductive argument.

1. The Problem

When teaching introductory logic courses, saying that while deductive arguments establish their conclusions with necessity, inductive arguments establish their conclusions only with probability is commonplace. Alternatively, if all the premises of a valid deductive argument are true, the conclusion will be true, but if all the premises of a cogent inductive argument are true, we may allow only that the conclusion is probable. This characterization is distinctly problematic. There is a clear, pre-analyzed notion of truth. We know what to expect if told a statement is true. But to be told that a statement is probable lands us in a philosophical thicket. Indeed, some theoreticians of probability would allow that ascribing probability or a degree of probability to a statement, as opposed to a relation between two statements, is meaningless. Although not all logicians may go this far, clearly the relational notion of conditional probability for them is primary. Just as valid deductive arguments do not show that their conclusions are necessary truths, but that the conclusions follow necessarily from the premises, so cogent non-deductive arguments show that their conclusions follow "with probability." But what does this mean? There are many alternative interpretations of probability. Authors of introductory logic texts which contain a chapter on probability make a further assumption here: Whatever the relation of following with probability might mean, the concept of probability involved satisfies the Pascalian axioms. This means that to understand the relation between premises and conclusion in a non-deductive argument, which we may designate as epistemic probability, we need to find a proper interpretation of these axioms. There are four principal interpretations of Pascalian probability–logical, Bayesian personalistic, limiting frequency, and propensity. We hold that one cannot explicate epistemic probability through the first three. To present the arguments why is beyond the scope of this presentation. But we shall present one argument against the logical interpretation, which might seem the first candidate to entertain in evaluating arguments. We shall then argue that the propensity interpretation may be one factor–but not the only factor–in a proper interpretation of epistemic probability adequate for one type of non-demonstrative argument.

2. Logical Probability

According to the logical interpretation, probability, like deductive entailment, is a logical relation between two statements, one of which may be the conjunction of a finite set of premises. The logical probability of the conclusion given the premises is a measure of the overlap between the situations where the premises are all true and the conclusion is true. Deductive entailment is a limiting case of logical probability, where the situations where the premises are true are completely included in the situations where the conclusion is true. But understanding epistemic probability as a *logical* and thus *a priori* relation is highly problematic. Salmon highlights why. The truth or falsity of statements of logical probability "are results of definition and pure logic alone; they have no synthetic or factual content" (1967, p. 75). Hence "This theory provides no reason for supposing any connection whatever between what is probable and what happens often.

It seems to provide no basis for expecting the probable in preference to the improbable" (1967, p. 79). But critics have indicated many instances where epistemic probability rightly depends on empirical considerations. Cohen (1989) for example points out that on the logical interpretation the information that a given subject who has had a toxic reaction upon taking a drug D and also has a history of heart problems has just as much effect on the probability that everyone taking drug D shows a toxic reaction as does the information that this subject had the reaction and drinks an average of 60 ounces of water a day! Although we cannot present the details of Cohen's argument here, we take this result as telling. The logical interpretation sees probability as a relation between beliefs, propositions, statements. Should our question be determining how strongly the premises of an argument support its conclusion and in particular how strongly that support must be for the premises to constitute grounds adequate for the conclusion, attempting to identify some relation between the premises and the conclusion as whole propositions seems an obvious first strategy. But it is not the only strategy, at least for some arguments. To introduce an alternative, however, we need to review some considerations from argumentation theory.

3. Arguments, Warrants, and The Propensity Interpretation

As Toulmin (1958) has taught us, corresponding to an argument is a warrant. We agree with Hitchcock in (1985) that warrants should be understood as inference rules rather than statements. As long as the argument is not a non-sequitur, premise and conclusion will share at least one content expression. In formulating the warrant, we may replace all occurrences of at least one of these content expressions with variables, which need not be individual variables, although for our purposes here, we shall confine ourselves to warrants involving just individual variables. Hence a paradigm example of a warrant we are considering here is

From x is in New York City

To infer x is in New York State

Shifting attention from the statements which constitute an argument to the warrant the argument instances is crucial to seeing how another interpretation of Pascalian probability may be a candidate for explicating epistemic probability. Now, as Cohen points out, frequency interpretations understand probability as concerning "events, classes of events, properties, or other features of reality" (1989, p. 59). But clearly, to speak of an event such as getting a head on the toss of a coin is to speak of an event satisfying a certain property. Likewise, given a class of events, we may attempt to identify the defining property of that class. The same holds for sets in general. Thus we may think of frequency interpretations as dealing with properties, including relations. We may then intelligibly speak of the frequency of the conclusion of a warrant being satisfied, given that the premises were satisfied.

For the sake of simplicity, let us deal just with monadic properties. Where

From $\varphi_1 x$, $\varphi_2 x$, ..., $\varphi_n x$

To infer ψx

expresses the warrant of an argument, should $\Pr(\psi x / \varphi_1 x \ \& \ \varphi_2 x \ \& \ ... \ \& \ \varphi_n x) \geq m$, where m is regarded as appropriately high, e.g. $m = .95$, would we be entitled to say that premises and conclusion of an argument instancing this warrant are adequately connected to transfer the acceptability of the premises to the conclusion? A moment's reflection should indicate that this proposal fails in general. Since $m \neq 1$, such a warrant is subject to rebuttals or defeaters. Let 'χx' express such a defeater and 'a' name some individual. Then if one had a justified belief that $\varphi_1 a$

& $\varphi_2 a$ & ... & $\varphi_n a$ & χa, one should not, on the basis that $\varphi_1 a$ & $\varphi_2 a$ & ... & $\varphi_n a$ holds, find that ψa is acceptable. The force of the warrant has been defeated. Notice that the failure of the warrant need not be due to its being supported by biased evidence. In the sample backing the warrant, the proportion of items which are χ may adequately mirror the proportion of χ's in the general population. Rather, it is a question of whether a critical challenger should find ψa acceptable, given that she finds χa acceptable together with $\varphi_1 a$, & $\varphi_2 a$ & ... & $\varphi_n a$. The answer is clearly negative.

Suppose, by contrast, that one had no evidence that χa nor evidence that the instantial evidence backing the warrant was biased. Under these circumstances, do the premises constitute an adequate *prima facie* case for the conclusion? This proposal is still problematic, because there are (at least) two frequency interpretations to distinguish, the relative or limiting frequency interpretation and the propensity interpretation. Although time will not permit our presenting the justifying arguments, we are setting aside the limiting frequency interpretation. On the propensity interpretation, that $\Pr(\psi x/\varphi x) = .x$ means that φ's have a tendency, propensity, disposition to be ψ's x per cent of the time. The evidence that $\Pr(\psi x/\varphi x) = .x$ consists in observed frequencies of ψ's among φ's in one or more trials. The propensity theory takes these frequencies as indicating something about the structure (or nature) of elements in the reference class.

How may we understand epistemic probability on the propensity interpretation? Where **P** represents the conjunction of premises of an argument, that $\Pr(C/\mathbf{P}) = .x$ means that where

From $\varphi x_1, x_2, \ldots, x_n$

To infer $\psi\ x_1, x_2, ..., x_n$

is the warrant of the argument from **P** to *C*, in x per cent of cases where φ holds of some n-tuple $<a_1, a_2, ..., a_n>$ of elements of the domain, ψ holds of that n-tuple because of potentially operative factors in $<a_1, a_2, ..., a_n>$'s satisfying φ. How adequate is this explication of epistemic probability? Perhaps the central issue concerns how judgments of epistemic probability may be supported or justified. This again is tantamount to asking how warrants may be backed. For the type of warrants to which we are restricting our discussion here, the backing consists of evidence from observations of actual frequencies in samples. When then would such an observation justify a judgment of probability?

This evaluation proceeds in two steps. First, we observe in a particular sample the relative frequency among n-tuples of elements of the domain which are φ those which are ψ. Hereafter we shall speak simply of φ's and ψ's. Suppose we observe a frequency of .y. Consider the interval (.y-ε, .y+ε). Assume for the sake of illustration that ε = .01. The second step is to verify the reliability of our test. This involves addressing the following question: Given the size of the sample, what is the probability that the actual proportion .x of φ's in the reference class which are also ψ is within this interval? This is a question of statistics. There are several ways of addressing this question, such as the method of confidence intervals, but that issue is beyond the scope of this paper.

In explicating epistemic probability through a propensity interpretation, we understand a warrant, "From φ*x*, To infer ψ*x*" as in effect asserting that φ's have a propensity to be ψ's.

Suppose we have a justified belief that φ's have a great propensity to be ψ's. Can we reliably reason according to this warrant? Does it properly transfer acceptability? As we indicated, for the arguments considered here, the warrants are backed "from below" by samples. Suppose then that we take a sample of 10,000 φ's and discover that 9,800 of them are ψ also, i.e. the frequency of ψ's in our sample is .98. Now given a sample of that size, we estimate that the frequency of φ's which are ψ in the overall population or reference class is 98 ± 1%. Since our sample size is 10,000, reasoning in accordance with the method of confidence intervals, and where f indicates the actual frequency of φ's which are ψ's, not just in our sample, we have $\Pr[(.98 - 1) \leq f \leq (.98 + 1)] \geq .95$. So on the basis of our sample data, we estimate that the proportion of φ's which are ψ falls in the interval (.97, .99), where our backing procedure is right 95% of the time. If we can be satisfied with a 95% confidence level in our backing procedure and that procedure shows that the propensity of φ's to be ψ's falls in the interval (.97, .99), surely we have a good *prima facie* case that our inference rule transfers acceptability. But can we conclude that we have a good case *simpliciter* for the reliability of the inference rule?

Unfortunately, the problem of defeaters returns right here. It could be that the frequency of ψ's among φ's in general is .98 ± 1, but the frequency of ψ's among φ's which are also χ's is .02 ± 1, where our backing of this result again has a 95% confidence level. Hence, even if we find the premise that φa acceptable, we may not infer ψa, if we also accept χa, unless of course that defeater can itself be defeated. We shall indicate how in connection with a further point we need to note.

For a potential defeater to have force, it is sufficient but not necessary that it be acceptable. It is only necessary that the burden of proof be on the proponent of the argument to show that the *negation* of the defeater *is* acceptable or that its defeating force is otherwise undercut. If it is a genuine question whether χa or $\sim\chi a$ holds–should $\Pr(\chi x/\varphi x) = \Pr(\sim\chi x/\varphi x) = .5$–and the person to whom the proponent is addressing the argument recognizes this fact, then she, the challenger, has every right and indeed the dialectical duty, to ask the proponent to please show $\sim\chi a$ or to show that some further counterdefeating condition operates here, i.e. to show that $\Pr(\psi x/\varphi x \ \&\ \chi x \ \&\ \Delta x)$ is appropriately high for some property Δ. In selecting our sample of 10,000 φ's, we have indicated no provision ensuring that the sample is representative. But, as Cohen points out, if we take such potentially defeating factors into account when selecting the backing sample for a warrant, we shall no longer be basing our claim of the reliability of the backing procedure solely on the size of the sample. We shall be taking into account not just enumerative but variative induction. Assume that the generalization corresponding to the warrant has been established as a law of nature. For the warrant to be reliable in a sufficiently wide range of cases, then, it is necessary not just that the frequency of φ's which are ψ's be appropriately high and the backing certify the propensity within a given confidence interval, the backing must also satisfy variative criteria.

Some will reply that these considerations say nothing new. We all acknowledge that the sample must be representative. Some philosophers speak of a requirement of total evidence, while others, e.g. Hempel (1965), talk of a requirement of maximal specificity for the reference class. What shall we say to such requirements as a preliminary condition for establishing the reliability of an inference rule? Cohen has one significant practical objection. Such requirements cannot be satisfied in practice; they are ideal. Suppose we know that Sam is a truck driver. We want to know whether Sam will live to 65. But there is a lot more that we may or could know about Sam relevant to the issue of his longevity.

> Even to know just the *available* evidence (physical, meteorological, geological, astrophysical, epidemiological, economic, socio-political, etc.) bearing on a person's survival to sixty-five one would have to work away indefinitely, since what is not learned today might, with sufficient effort, be learned tomorrow. (Cohen, 1989, p. 101)

Rather, in assessing the reliability of an inference rule, we must take a different tack. Instead of seeking some ideal maximally specific reference class, we must consider the evidential spread of a sample and be able to compare samples for evidential spread. We want our sample evidence to have as wide a spread as practically possible. What does this involve?

We may offer this motivation here. Suppose Jones' physician tells him that he will not have a heart attack within the next five years, since his cholesterol level is 150. Is the premise an adequate ground for the conclusion? The warrant of the inference is

From x has a cholesterol level of 150

To infer x will not have a heart attack over the next five years

If statistical records are available, we may identify a sample of persons whose medical records as of a certain date more than five years ago indicated a cholesterol level of 150 and determine the percentage of those who did not have a heart attack over the next five years. Suppose the relative frequency of those in our sample with no "coronary event" over the five years subsequent to the finding of a serum cholesterol level of 150 is .96. Suppose our enquiry has a 95 per cent confidence level, i.e. where .x is the frequency of the non-occurrence of heart attacks over a five year period following a blood test within the reference class of persons found to have a cholesterol level of 150 in the test, $\Pr[.95 \leq .x \leq .97] = .95$. If 95 per cent is an acceptable reliability level for our inference rule, our test shows that our warrant is reliable.

However, there are many other factors bearing on whether one has a heart attack, for example the ratio of high density to low density cholesterol. Suppose we did a series of surveys of persons all with serum cholesterol levels of 150 but with varying high density to low density cholesterol ratios. Suppose all these tests had a 95 per cent confidence level and that on the basis of all of them, we could predict that the rate of avoiding heart attacks within five years of persons with a cholesterol level of 150 and having one of these values of the high density/low density ratio was 95 ± 1 per cent. We again have confirmation of the reliability of our inference rule, but our ground seems to be more adequate. We have ruled out a potential defeater. Should we carry out further tests, targeting other potential defeaters but taking into account those already considered, and all showed that differing values of these variables did not affect the rate of heart attack avoidance over a five year period, we would have backing evidence of even greater weight. These considerations show that we can compare bodies of backing evidence not only for size but for weight, and in this case where the results are positive, the greater the weight, the stronger the argument. What are the implications of this recognition for a Pascalian explication of epistemic probability?

Our considerations at this point suggest that an adequate conception of epistemic probability needs to supplement a propensity interpretation of Pascalian probability with a consideration of what Cohen, following Keynes, calls the weight of the evidence. Should we be able to specify a weight regarded as indicating a weight sufficient to transfer acceptability, in conjunction with a sufficiently high probability, we would have a specific answer to the question of when the premises of certain arguments adequately ground their conclusions. Such arguments have warrants of the general form 'From φx/To infer ψx' which are backed by evidence of the frequency of ψ's among φ's. By requiring the premises to have appropriate weight in addition to the warrant of the argument having a sufficiently high frequency probability, we have supplemented explicating epistemic probability solely through an interpretation of Pascalian probability. Can this supplement concerning the additional weight be understood in Pascalian terms?

We may argue that it does not on several grounds. First, a basic theorem of the Pascalian calculus tells us that $\Pr(\psi x/\varphi x) = \mathrm{n}$ if and only if $\Pr(\sim\psi x/\varphi x) = 1 - \mathrm{n}$. But both of the inference rules 'From φx/To infer ψx' and 'From φx/To infer $\sim\psi x$' may be backed by the same body of evidence, instances of φ's which are either ψ or not-ψ. So the weight of evidence for both is the same. Hence the premise 'φs' supports the conclusions 'ψs' and '~'ψs' with the same weight, although the probability (unless $\Pr(\psi x/\varphi x) = \Pr(\sim\psi x/\varphi x) = .5$) is different. Weight does not follow the Pascalian rule for negation. Neither does weight follow the Pascalian multiplication rule: $\Pr(\psi x \,\&\, \chi x/\varphi x) = \Pr(\psi x/\varphi x) \times \Pr(\chi x/\varphi x \,\&\, \psi x)$. We cannot present the argument here.

Let us take stock of where our argument sketch has brought us to this point. We understand epistemic probability as the relation between an argument's premises and conclusion by virtue of which the premises may give some support to the conclusion and, if the epistemic probability is sufficiently high, will transfer their acceptability to the conclusion. We have been exploring whether epistemic probability can be explicated in terms of an interpretation of Pascalian probability. We have here argued that the logical interpretation is not appropriate. However, we have found that the propensity interpretation dovetails nicely with our understanding of how strongly the premises of an argument support its conclusion by virtue of the degree of reliability of the argument's warrant. Here we are understanding that the warrant is backed by statistics concerning the relative frequency with which the conclusion is satisfied when the premises are satisfied. But we have found that argument strength is not simply a matter of statistical strength here. So, following Cohen's contrast of enumerative with variative induction, we have argued that both enumerative and variative inductive considerations are necessary to establish that a warrant (of the sort we have been considering) is properly backed to be sufficiently reliable for acceptability transfer. However, insofar as enumerative induction is involved, we are at this point taking the enumerative evidence as indicating the conditional probability of the warrant's conclusion on the basis of the premises, where probability is understood as propensity. Hence we allow, at this point, that Pascalian probability is not out of place in explicating epistemic probability for all types of arguments, in particular for non-deductive arguments be backed "from below." The extent to which Pascailian probability is a proper factor in explicating epistemic probability for other types of arguments remains to be investigated.

References

Cohen, L. Jonathan. 1989. *An Introduction to the Philosophy of Induction and Probability*.

Oxford: Clarendon Press.
Hempel, Carl G. 1965. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press.
Hitchcock, David. 1985. "Enthymematic Arguments." *Informal Logic* 7: 83-97.
Salmon, Wesley C. 1967. *The Foundations of Scientific Inference*. Pittsburgh, PA: University of Pittsburgh Press.
Toulmin, Stephen E. 1958. *The Uses of Argument*. Cambridge: Cambridge University Press.