BOOK REVIEW

THE PROBABLE AND THE PROVABLE. By L. Jonathan Cohen. Oxford: Clarendon Press, 1977. Pp. xvi, 363.

*Reviewed by Carl G. Wagner**

In matters involving uncertainty, probability theory, and its applied companion, statistical decision theory, have often proved to be valuable aids to rational decisionmaking. Thus, it is natural to expect that the conceptual tools furnished by these theories might also be applicable to the analysis or practice of legal decisionmaking.[1] Indeed, the language of the law, replete with grammatical variants of the very term "probability," and with phrases such as "more likely than not" and "beyond a reasonable doubt," seems to invite explication in formal probabilistic terms. Yet a number of arguments have been advanced in opposition to such an undertaking. Some individuals, having encountered probability only in the narrow application of forecasting, have concluded, for example, that only future uncertainties may be analyzed probabilistically, and that issues dealt with by the law, which characteristically involve uncertainty about the past, must therefore lie beyond the reach of such analysis.[2] Others, acquainted only with the interpretation of probabilities as relative frequencies attached to sequences of repeated "experiments," have noted that triers of fact face situations too idiosyncratic to allow the application of statistical generalizations,[3]

* A.B. 1965, Princeton University; Ph.D. 1969, Duke University; Associate Professor of Mathematics, University of Tennessee. This review was prepared while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences. I am grateful for financial support provided by the National Science Foundation (Grant No. BNS 76-22943 A02), the Andrew W. Mellon Foundation, and the University of Tennessee.

1. *See, e.g.*, Finkelstein & Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489 (1970); Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065 (1968); Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021 (1977).

2. For an incisive correction of this misconception, see Tribe, *Trial By Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

3. In a criminal trial, for example, it is clearly erroneous to view the probability of a defendant's guilt as the fraction of times that an individual would be guilty in a large number of identical cases. Common experience indicates that no two cases are exactly alike, and so no such fraction can be calculated.

and have concluded on these grounds that probability theory must be inapplicable to legal decisionmaking.

To hold for the above reason that probability is inapplicable to legal decisionmaking is to fail to appreciate that the axioms of probability theory do not specify a fixed interpretation of probability numbers, but rather a set of consistency conditions for the assignment of such numbers. To equip a set of declarative sentences from some field of inquiry such as science or law with a probability "measure" is simply to assign to each sentence in the set a number between zero and one, subject to the following restrictions: (1) logically equivalent sentences must be assigned the same number; (2) self-contradictory sentences must be assigned the number zero, and tautologies the number one; and (3) if two sentences cannot be true simultaneously, then their disjunction (the compound sentence formed by connecting them with the word "or") must be assigned that number which is the sum of the numbers assigned to its constituent sentences.

One way of generating such probability numbers, commonly used in scientific inquiry, is to assign to an assertion the number expressing the fraction of times the assertion is found to be true in a sequence of experimental tests. In this case probability numbers express the relative frequencies with which assertions are found to be true.

As noted above, triers of fact will rarely have information sufficient for assigning frequency-based probability numbers to contested assertions. Their data, after hearing the evidence, consist of unquantified subjective degrees of belief in the assertions at issue. To assign numerical measures to such degrees of belief, consistent with the axioms of probability theory, might seem an impossible task. As a result of the groundbreaking work of Savage[4] and de Finetti,[5] however, we are fortunate to possess a method by which an individual may so quantify his subjective degrees of belief.[6] Probability numbers assigned by this method measure what is, for the purposes of the law, the crucial

4. *See* L. SAVAGE, FOUNDATIONS OF STATISTICS (1972).

5. *See* de Finetti, *Probability: Interpretations*, in 12 INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL SCIENCES 496 (D. Sills ed. 1968).

6. Roughly speaking, an individual quantifies his degree of belief in the truth of some assertion by a process of self-interrogation involving the preferential ranking of certain bets. To do this he makes use of comparisons with frequency-based probabilities connected with some simple enterprise such as choosing a card at random from a well-shuffled deck. He then asks himself a series of questions ("Would I prefer betting on X's guilt to betting that the card will not be a spade? Yes? Then I must assign X's guilt a probability greater than .75. Would I prefer betting on X's guilt, even to betting that the card will not be the ace of spades?" and so on) by which he refines his subjective estimate of the probability of the contested assertion. Thus an individual's subjective probabilities reflect his degrees of belief as manifested in his betting preferences.

characteristic of a contested proposition, namely, the degree of belief accorded it by a trier of fact. Moreover, these numbers are assigned in such a way that the general theorems of probability theory are available as inferential tools for legal decisionmaking.

Nevertheless, Laurence Tribe has contended that the explicit quantification of forensic probabilities would be undesirable on grounds of policy, arguing that such a practice would undermine the finality with which legal proceedings are supposed to settle contested issues.[7] For example, it follows from the axioms of probability theory that attaching a numerical weight p to one's belief in some material proposition entails attaching the weight 1-p to belief in its negation. The fact that there are some values of p that are less than 1 but still compatible with belief "beyond a reasonable doubt" detracts from the solemn assurance implicit in that standard by officially conceding that it does not mean belief beyond a quantifiable doubt. Furthermore, the authority of a verdict rendered on the basis of a "preponderance of evidence," usually identified with the condition $p > 1/2$, fares even worse in the light of probabilistic candor. Tribe also fears that if triers of fact were encouraged to think in explicit mathematical terms, they would give undue weight to statistical evidence (conveniently quantified for them by others) rather than attempt the difficult task of integrating such evidence with other relevant considerations to arrive at a fully informed subjective probability.[8]

Such arguments, based on a commitment to protecting the ritual aspects of the legal process, depend on certain questionable assumptions about human intellect and psychology.

Is the ability of human beings to accept a legal standard that declares final that which is not entirely free from doubt as fragile as Tribe fears? Are juries so impressionable with regard to "hard" pieces of numerical evidence as to warrant excluding such evidence from their consideration? At present we lack the careful empirical studies required to answer these questions decisively. Even if Tribe's pessimism should be confirmed by such studies, however, this would not rule out the possibility of training jurors in the reasoning processes required to deal with probabilistic evidence in a competent manner. We might, therefore, still profitably study the probabilistic reasoning of an ideal Bayesian subjectivist[9] juror, in the expectation of clarifying and evalu-

7. Tribe, *supra* note 2, at 1372-75.

8. *Id.* 1363.

9. A "Bayesian subjectivist" revises his subjective probability assignments in the light of new evidence according to a formula (Bayes' Rule) discovered by the Reverend Thomas Bayes in

ating actual practice from a theoretical, normative perspective.

In *The Probable and the Provable*, however, the philosopher L. Jonathan Cohen[10] aims to demolish the idea that classical probability (which he calls "Pascalian probability") is even a theoretically appropriate model for the law's apparently probabilistic language. Cohen pursues this task in a chapter entitled "Six Difficulties for a Pascalian Account of Judicial Probability." The thrust of his argument is that the language of forensic, or legal, probability differs so significantly in its formal structure from that of Pascalian probability that no meaningful parallels can be drawn. In particular, Cohen claims that a Pascalian version of forensic probability is at odds with the legal rules that establish the standards of proof necessary for the acceptance of conjunctive propositions, such as the degree of proof necessary to establish the conjunction of proximate cause and injury in a tort action, and control inferential argument, such as the rules setting the required standard of proof in arguing from one proposition to an intermediary proposition, and through this inference, to a conclusion. Since these two points lie at the core of Cohen's presentation, it is important to examine them in some detail.

1. *The Difficulty About Conjunction.* In seeking to state forensic probability in Pascalian terms, the first task is to specify a Pascalian measure for the degree of belief in some material proposition M that is justified by evidence E. The traditional choice for such a measure has been the so-called "conditional probability[11] of M, given E," and it is this choice that Cohen criticizes.

Let us first consider the Pascalian analysis of the notion of a pre-

---

1763. The simple intuition underlying this rule is that if the probability of a piece of evidence increases with the truth of a contested proposition, the acceptance of the evidence as true should increase the estimated probability of the contested proposition. See Kaplan, *supra* note 1, at 1083-91, for a lucid discussion of the algebraic details of Bayes' Rule.

10. Mr. Cohen is a Fellow at The Queen's College, Oxford University.

11. As a simple illustration of conditional probability, consider a single toss of a fair, six-sided die. Lacking any other information, we would assign an unconditional probability of $3/6 = 1/2$ to the assertion (M) that an odd number turns up. If, however, we know that a number greater than or equal to 4 has turned up (E), this will lower our estimate of the probability of M to $1/3$, since only one of the numbers 4, 5, and 6 is odd. The number $1/3$ is the conditional probability of M, given E.

An equivalent formulation of conditional probability, which is often convenient for theoretical purposes, is as follows: The conditional probability of M, given E, is equal to the unconditional probability of M and E, divided by the unconditional probability of E. In the above example, we could thus have calculated the unconditional probability that an odd number greater than or equal to 4 turns up ($1/6$) and divided this by the probability that a number greater than or equal to 4 turns up ($3/6$), arriving at $(1/6)/(3/6) = 1/3$ as the conditional probability.

ponderance of evidence. Suppose that we are given material proposition M and evidence E, and that we abbreviate the assertion that M is not the case by "not-M." Further, let us denote the conditional probability of M, given E, by Pr[M,E], and the conditional probability of not-M, given E, by Pr[not-M,E]. It follows from the axioms of probability and the definition of conditional probablity that the sum of these two numbers is always equal to one.[12] Therefore under the Pascalian theory, there is a preponderance of evidence for M if and only if the conditional probability of M, given E, is greater than one-half. Otherwise not-M is just as probable as or more probable than M.[13] Now, Cohen argues, suppose that M_1 and M_2 are component points of plaintiff's contention in a civil case. (Suppose, for example, that it is a tort action and M_1 proposes actual injury and M_2 proposes proximate causation.) The legal rule is that plaintiff prevails if he proves each point by preponderance, that is, if he proves that $Pr[M_1,E] > 1/2$ and $Pr[M_2,E] > 1/2$. But suppose, for example that $Pr[M_1,E] = .8$ and $Pr[M_2,E] = .6$. The law in this case certainly regards plaintiff as having proved his entire contention, M_1 and M_2, injury and causation, by a preponderance. If Pascalian probability were readily applicable to legal decisionmaking in such a case, we would expect the conditional probability of the conjunctive proposition M_1 and M_2, given E, to be "probable" as well, that is, greater than one-half. In fact, however, under traditional Pascalian theory, the probability of the conjunctive propositon M_1 and M_2, given the independence of M_1 from M_2, is equal to the product of their individual probabilities: $.8 \times .6 = .48$.[14] Thus, what Cohen shows to be legally "probable" he also shows to be simul-

12. In symbols, we may express this as $Pr[M,E] + Pr[not\text{-}M,E] = 1$, or, $Pr[not\text{-}M,E] = 1 - Pr[M,E]$. This is the so-called "complementarity principle" for conditional probabilities.

13. This is because "preponderance" amounts to the assertion that the conditional probability of M, given E, is greater than the conditional probability of not-M, given E ($Pr[M,E] > Pr[not\text{-}M,E]$). Since the sum of these two numbers is equal to one, the latter inequality holds exactly when the conditional probability of M, given E, is greater than one-half ($Pr[M,E] > ½$).

14. The universally applicable rule for computing the probability of a conjunction M_1 and M_2, given E, is as follows:

$$Pr[M_1 \text{ and } M_2, E] = Pr[M_1, E] \times Pr[M_2, E \text{ and } M_1].$$

If, in the context of E, the truth of M_1 has no bearing on the truth of M_2, that is, if M_2 is *independent* of M_1, then

$$Pr[M_2, E \text{ and } M_1] = Pr[M_2, E].$$

Hence, in the case of independence, we get the simplified special formula $Pr[M_1 \text{ and } M_2, E] = Pr[M_1, E] \times Pr[M_2, E]$.

For the formally inclined reader, the mathematical derivation of the general product rule is as follows: By the remarks in note 11 *supra*, we have

$$Pr[M_1 \text{ and } M_2, E] = \frac{Pr[M_1 \text{ and } M_2 \text{ and } E]}{Pr[E]}$$

taneously improbable under traditional Pascalian analysis.

This tension between legal practice and the Pascalian model can be approached in more than one way. An individual might concede on the basis of this example that the Pascalian model is not always perfectly descriptive of legal practice, and simply view this as a defect in the inferential rules of the law. This need not imply a wholesale rejection of the logic underlying legal practice, particularly if one is convinced that examples like the foregoing occur in practice with sufficiently low frequency. Several considerations suggest the reasonableness of such a view.

First, such material propositions as injury and proximate causation, signified here by M_1 and M_2, are rarely, if ever, independent. Without the independence assumption, we must employ the more general Pascalian multiplication rule $Pr[M_1 \text{ and } M_2, E] = Pr[M_1, E] \times Pr[M_2, M_1 \text{ and } E]$.[15] If, as one would often expect, $Pr[M_2, M_1 \text{ and } E]$ is sufficiently larger than $Pr[M_2, E]$ to make the foregoing product greater than one-half, the conjunction of M_1 and M_2 would be "probable" in both the legal and the Pascalian senses.

Second, defendants often dispute only one of the points in a conjunctive complaint, for example, conceding that a plaintiff indeed suffered injury, but denying the existence of proximate cause. This, in Pascalian terms, results in the automatic assignment of the value 1 to the probability of the uncontested point. In such cases there is no difficulty about conjunction, given a probability greater than one-half for the contested proposition.

Cohen explicitly considers the possibility of such mitigating factors, but remains unplacated. As long as, on a Pascalian analysis, there remains the indisputable possibility of having to countenance an anomaly in the standards of conjunctive proof, no matter how remote, Cohen asserts that the Pascalian model must be rejeeted.

2. *The Difficulty About Inference Upon Inference.* While in the case of conjunction, the Pascalian model suggests that the existing rules of legal proof are excessively loose, Cohen shows that this model occa-

---

$$= \frac{Pr[M_1 \text{ and } E]}{Pr[M_1 \text{ and } E]} \times \frac{Pr[M_2 \text{ and } (M_1 \text{ and } E)]}{Pr[E]}$$

$$= \frac{Pr[M_1 \text{ and } E]}{Pr[E]} \times \frac{Pr[M_2 \text{ and } (M_1 \text{ and } E)]}{Pr[M_1 \text{ and } E]}$$

$$= Pr[M_1, E] \times Pr[M_2, M_1 \text{ and } E].$$

15. See note 14 *supra*.

sionally makes certain accepted legal limitations on inference upon inference appear too strict. The simplest inferential arguments proceed from evidence E to some intermediate proposition I, and then from E and I to the target material proposition M. The legal rule in civil cases requires that all inferences in the chain prior to the last must be proved beyond a reasonable doubt, which translates into a probability very close to one, while the last may be proved on preponderance, that is, with a probability greater than one-half.[16] Let us say that the material proposition (M) in this situation is the negligent failure to stop a car resulting in the injury of a pedestrian, and that our evidence (E) is comprised of: (1) identification of the driver, (2) testimony that the car was in good working order, (3) testimony that it was a clear day, (4) testimony that traffic was light, and (5) the driver's claim that he failed to see the pedestrian. Furthermore, let us say that I, the intermediate proposition, is that there were no distractions for the driver. The inferential argument thus runs from evidence that a specific person, driving a well-maintained car on a traffic-free road on a clear day failed to see a pedestrian, to the intermediate proposition that, given the evidence, it is most probable that the driver had no significant distractions, to the material proposition that, given the evidence and the probability that the driver had no distractions, the driver was negligent in some fashion for failing to stop. By the laws of Pascalian probability, it follows that the ultimate probability of our material proposition M, given our evidence E, is at least as large as the product of the probability of our intermediate proposition I, given our evidence, and the probability of M, given E and I.[17] Suppose now that the probability of I, that the driver had no distractions, given E, were quantified by a juror at the level .75, and that the probability of our material proposition of negligence, signified by M, given E and I, were quantified at the level .8. If we apply the aforementioned inequality for the product of these num-

16. See J. WIGMORE, ON EVIDENCE § 41, at 439 (3d ed. 1940) (quoting New York Life Ins. Co. v. McNeely, 52 Ariz. 181, 193-96, 79 P.2d 948, 953-55 (1938)).

17. Symbolically, we have $\Pr[M, E] \geqslant \Pr[I, E] \times \Pr[M, E \text{ and } I]$. The proof is as follows: Since the assertion M and E is true whenever the assertion M and E and I is true, it is clear that $\Pr[M \text{ and } E] \geqslant \Pr[M \text{ and } E \text{ and } I]$.

Hence,

$$\frac{\Pr[M \text{ and } E]}{\Pr[E]} \geqslant \frac{\Pr[M \text{ and } E \text{ and } I]}{\Pr[E]} \times \frac{\Pr[E \text{ and } I]}{\Pr[E \text{ and } I]}$$

that is,

$$\frac{\Pr[M \text{ and } E]}{\Pr[E]} \geqslant \frac{\Pr[E \text{ and } I]}{\Pr[E]} \times \frac{\Pr[M \text{ and } E \text{ and } I]}{\Pr[E \text{ and } I]}$$

According to the remarks in note 11 *supra*, this is equivalent to

$$\Pr[M, E] \geqslant \Pr[I, E] \times \Pr[M, E \text{ and } I].$$

bers, we find that the probability of M, given E, is at least as large as .6 and thus appears to satisfy, in Pascalian terms, the standard of a preponderance of evidence (.6 > 1/2). Nevertheless, this proof would be unacceptable to the courts because the quantified probability of I, the intermediary proposition, is .75 and as such is certainly not equal to the significantly higher quantitative threshold for proof beyond a reasonable doubt. Thus, legal standards again fall short of meshing with a Pascalian system of probability.

Cohen observes, moreover, that the Pascalian system does not capture the asymmetry of the rule inference upon inference. If, for example, Pr[I,E] = .999 and Pr[M, E and I] = .6, the courts allow the inference from E to M. But if Pr[I,E] = .6 and Pr[M,E and I] = .999, the inference from E to M is disallowed. Yet on the Pascalian account, we have in both cases that Pr[M,E] > .5994. Thus, once again, the Pascalian model fails to be descriptive of legal practice.

This failure, however, is surely not as disturbing as Cohen makes it out to be. While, from the Pascalian point of view, the rule regulating inference upon inference may seem more conservative than necessary, this could be regarded as offsetting those instances in which the criteria for conjunctive proof appear to be too liberal. Also, one can understand the asymmetry of the existing rule governing inference upon inference as a necessary pragmatic modification of the rule the Pascalian model would prescribe. Earlier inferences in a chain are predicated on more remote, less specific conditions than later ones. From the standpoint of caution, these earlier inferences should require proof at a very high level of probability.

One may, in short, concede that the Pascalian model does not fit legal practice on purely *descriptive* grounds, while continuing to view it as a *normative* model, by which the departures of practice can be measured and perhaps rationalized on pragmatic grounds. From this perspective, the legal rules governing proof of conjunctions and arguments based on inference upon inference appear as compromises with a Pascalian ideal, dictated by the fact that jurors are generally incapable of applying the sophisticated multiplication rules of Pascalian probability.[18]

The recognition that the realities of legal practice may fall short of the Pascalian ideal surely need not be the occasion for despair about the general fairness of such practice.

Nonetheless, this way of dealing with the aforementioned difficul-

18. I am indebted to John Kaplan for convincing me of this point.

ties is unacceptable to Cohen. For him, legal practice is the norm and the goal is to construct a formal model that reflects this practice in every detail. He proposes as an alternative to Pascalian probability a method of weighing propositions according to the "inductive support" that can be marshalled on their behalf. Roughly speaking, inductive support accrues to a proposition to the extent that it survives a sequence of attempts to falsify it. The "inductive probability" of material proposition M on evidence E is measured by the inductive support available for the universal inference of M from E. Significantly, inductive probabilities need not be real numbers; virtually any set of sequentially ordered objects will suffice for inductive scorekeeping, for all that is required is a way of keeping track of the level to which a proposition has resisted falsification. Inductive probabilities cannot be combined by anything resembling the usual arithmetical operations such as multiplication or addition, and hence there is no *algebra* of inductive probabilities to speak of.[19]

Given this paucity of mathematical structure, one might expect the foundations of inductive probability to be quite elementary. They are, to the contrary, forbiddingly complex, yet frustratingly vague at crucial points.[20] The important concept of relevance is furnished with no formal inductive analysis.[21] Furthermore, when Cohen begins his axiomatic rendering of inductive probability, he lists nearly fifty preliminary definitions and axioms, and this list is incomplete. He then refers readers to his earlier book, *The Implications of Induction*,[22] for portions of the development, including proofs of a number of theorems omitted from the present monograph.

Readers who, despite these roadblocks, wish to study Cohen's theory of inductive probability should begin with Chapter 19, "An Inductivist Resolution of Six Difficulties for a Pascalian Account of Judicial Proof." Consider first his resolution of the conjunction problem. Cohen shows in an earlier chapter that the conjunction of two or more propositions *about the same subject matter* has inductive probability

19. This precludes anything like Kaplan's Pascalian analysis of the interplay between standards of proof, rules of evidence, and societal and individual utilities, as there is no way to multiply inductive probabilities with utilities. *See* Kaplan, *supra* note 1.

20. In order to avoid certain inconsistencies in support-grading, for example, tests of a proposition must be ordered in terms of decreasing "falsificatory potential." *See* L. COHEN, THE PROBABLE AND THE PROVABLE 138-39 (1977). Cohen's informal discussion of this key notion does not provide a convincing argument that such an ordering always exists.

21. By contrast, the Pascalian account of the relevance of E to M amounts simply to the assertion that the conditional probability of M, given E, is greater than the unconditional probability of M.

22. L. COHEN, THE IMPLICATIONS OF INDUCTION (1970).

equal to the smallest of their individual probabilities.[23] A probability function with this property appears to mesh nicely with the legal standards for proof of conjuctions. The difficulty is that the notion of "same subject matter" is a very restrictive one and Cohen admits that the component points in a conjunctive complaint may not satisfy this condition. In such cases, he says, "the probabilities of the component elements are incommensurable, in which case no probability value can be assigned to their conjunction and separate assignments to each must suffice."[24] To be sure, this protects the inductive model from clashing with legal practice, but it does so by opting to say nothing about legal practice in the most interesting cases. Such a strategy is barely distinguishable from that of adopting a "partial Pascalian" model whereby one computes the probability of a conjunction by the multiplication rule only if no anomaly results.

Cohen's resolution of the problem of inference upon inference is even less satisfying. His inductive model never furnishes an inductive probability for M on evidence E, since the probabilities of the component links are always "incommensurable."[25] In short, the inductivist account provides no formal analysis of inference upon inference, thus trivially avoiding the possibility of disparity between model and reality. Pressed to say something to explain the legal rules governing inference upon inference, Cohen relies on the same observations about remoteness and the desirability of caution cited earlier[26] to rationalize the legal use of a criterion more stringent than that suggested by a Pascalian analysis.[27]

One cannot fail to be disappointed that so much effort should, in the end, produce such circumscribed results. On the other hand, Cohen has delineated, more clearly than any other critic of the Pascalian model of forensic probability, the points of tension between that theory and actual practice. Moreover, he has investigated in detail an alternative model that we had at least no a priori grounds for rejecting.

It has been just over ten years since Kaplan published the first substantial formal analysis of forensic probability and the processing of probabilistic evidence.[28] Since then, these issues have been vigorously discussed in the journals, and we are now witnessing the appearance of

23. L. COHEN, *supra* note 20, at 221.
24. *Id.* 266.
25. *Id.* 268.
26. See text accompanying note 18 *supra*.
27. L. COHEN, *supra* note 20, at 269.
28. *See* Kaplan, *supra* note 1.

the first book-length works devoted to this subject.[29] Despite its limitations, Cohen's ambitious attempt to construct a radically different model of forensic probability is a positive addition to this emerging genre.

29. *See, e.g.*, R. EGGLESTON, EVIDENCE, PROOF, AND PROBABILITY (1978); MODELING THE CRIMINAL JUSTICE SYSTEM (S. Nagel ed. 1977). *See also* G. SHAFER, A MATHEMATICAL THEORY OF EVIDENCE (1976).