Some Thoughts on Wesley Salmon's Contributions to the Philosophy of Probability

Paul Humphreys†

Wesley Salmon provided three classic criteria of adequacy for satisfactory interpretations of probability. A fourth criterion is suggested here. A distinction is drawn between frequency-driven probability models and theory-driven probability models and it is argued that single case accounts of chance are superior to frequency accounts at least for the latter. Finally it is suggested that theories of chance should be required only to be contingently true, a position which is a natural extension of Salmon's ontic account of probabilistic causality and his own later views on propensities.

1. Introduction. Over a period of roughly twenty years, I had many conversations with Wes Salmon about relative frequencies, definitions of randomness, and propensities. They were always illuminating and I learned a great deal from his wise counsel. But one thing upon which we always differed was the merits of propensity accounts of probability. I came to the view, albeit slowly, that some version of a single case propensity account was not only tenable but was indispensable for a full account of chance. Wes, despite a number of complimentary remarks in his writings about propensities, never, I think, completely lost his suspicion that they violated the empiricist's code of conduct. I felt, and still feel, that this was a great pity, because one of Salmon's most important contributions to philosophy, and he had many important contributions, was his insistence during the later part of his life that ontic approaches to issues in causation and explanation were both intellectually profitable and likely to be true. Single case propensities—or as I prefer *chances*—fit more naturally into this realist orientation than do relative frequencies. Thus, what I shall do in this paper is to sketch an account of chances that is consistent

---

†To contact the author, please write to: University of Virginia, Corcoran Dept of Philosophy, 512 Cabell Hall, P.O. Box 400780, Charlottesville, VA 22904-4780; e-mail: pwh2a@virginia.edu.

Philosophy of Science, 71 (December 2004) pp. 942–949. 0031-8248/2004/7105-0026$10.00

with Wes Salmon's ontic position on causation and with his later views on propensities. In no way am I claiming that this is, or would have been, his own view, simply that it is one which fits fairly naturally into the framework that Salmon had constructed by around 1990.

Relative frequentists talk, for the most part, in terms of reference classes, sequences, or data sets. Propensity advocates prefer to speak of chance setups, mechanisms, and dispositions. In order to avoid potentially misleading and philosophically loaded terminology, I shall use the term *'chances'* rather than 'propensities' and *'generating conditions'* rather than 'mechanisms'. The arguments in favor of propensities are, for the most part, centered around its advantages in dealing with the problem of the single case. In particular, it is the ability to attribute values of chance to a physical system when little or no frequency data has been obtained from that system that is the single greatest advantage of a commitment to chances grounded in generating conditions. There is a sense in which that is correct, but it requires adjusting our demands about what theories of chance can provide.

2. Salmon on Propensities. In 1979, before the ontic approach to causation had reached its later prominence in Salmon's thinking, he made the following assessment of the propensity approach. Having discussed the need to consider statistically relevant factors in assigning frequency values to the single case, he continues:

> When we consider the single case version of the propensity interpretation of probability, it appears to me that precisely the same problem arises. One cannot say simply that the probability of an outcome is a property of some unspecified chance setup; it is necessary, instead, to characterize the chance setup in terms of its *relevant* characteristics. Single case propensity theorists readily admit that observed frequencies provide the evidence by means of which to ascertain the propensity of a given chance setup to produce a result. In order to decide which chance setups to observe in order to collect the statistical data upon which to base a judgement about the value of a propensity, the propensity theorist must make precisely the same type of judgement as the frequentist concerning those features of the situation that are relevant to the occurrence of a given outcome, and those that are irrelevant. Thus, for example, in attempting to ascertain the propensity of *this* die to land with side six up on *this* toss, we must take account of the cubical shape of the die, the homogenous distribution of mass within it, the fact that it is shaken sufficiently before throwing and is tossed with sufficient force. Contrariwise, we ignore the color of the die, the question of whether the throw is made during the day

> or the night, and whether the thrower is left or right handed. For the single case propensity theorist, as for the frequentist attempting to deal with the single case, the question of which factors are relevant and which irrelevant is equally vital. It thus appears that the single case propensity interpretation, no less than the virtual sequence propensity interpretation, provides no substantial improvement over Reichenbach's limiting frequency interpretation, but at most an interesting reformulation which *may* include some shift of focus. (Salmon 1979, 14–15)

Salmon is of course correct in saying that the chance setup must be specified for the single case. But the kind of relevance that affects chances is different from that which affects reference classes; it is causal relevance that is important for chances, not statistical relevance. Moreover, the causally relevant factors cited by Salmon operate not between the events upon which relative frequencies are defined, such as "side six up," but upon the underlying structures that generate the frequencies, such as the die itself. The difference is profoundly significant. Causally relevant factors affect the relative frequencies indirectly by affecting the generating conditions before each datum appears. Statistically relevant factors affect the relative frequencies directly by division of the reference class after the data have been generated.[1] This point would be trite except for the fact that the switch enables us to avoid the trivialization that always haunts the partitioning of reference classes. For relative frequencies it is sufficient to consider the case of finite or infinite binary sequences of the form 001011101100. . . . Such sequences are extensional objects that can, qua extensional entity, be partitioned in many different ways and, except in the extreme cases of infinite sequences within which only a finite number of the elements of the sequence are 1s or 0s, there will always exist some subsequence with a different limiting frequency from that possessed by the original sequence. This trivialization was pointed out by Alonzo Church early in the development of definitions of random sequences (Church 1940) and many subsequent attempts were made to avoid the trivialization, including some ingenious attempts by Salmon himself (see e.g., Salmon 1984, Chapter 3).[2] There is an extensive literature on this issue and this is not the place to add to it. Instead, I want to argue that this causal aspect of chances makes the theory of probability an inap-

---

[1] This point must be understood conceptually. In prospective epidemiological studies, which relevant factors are to be used to divide the reference classes can be decided before the data are generated. But the division cannot be made on the basis of those factors until the class exists, i.e., when the data themselves exist, after the fact.

[2] In contrast, Reichenbach's account of limiting relative frequencies does not require randomness.

propriate constraint on any theory of chances, whether conditional or absolute.

3. Criteria of Adequacy. Salmon laid out three criteria of adequacy for interpretations of the probability calculus in his classic 1966 lectures *The Foundations of Scientific Inference* (published in book form with the same title in 1967). They are:

1. Admissibility: "the meanings assigned to the primitive terms in the interpretation transform the formal axioms . . . into true statements."
2. Ascertainability: "This criterion requires that there be some method by which, in principle at least, we can ascertain values of probabilities."
3. Applicability: "We are seeking a concept of probability that will have practical predictive significance."[3]

Salmon's preference for relative frequencies over propensities rested on his belief that propensities failed to satisfy the admissibility criterion for interpretations of probability and that the ascertainability criterion could ultimately only be satisfied for propensities indirectly through appeal to relative frequencies. This position is most clearly expressed in his article "Dynamic Rationality" (1988), where he has this to say:

> The problem is that the term "propensity" has a causal aspect that is not part of the meaning of "probability". I have no objection to the concept of propensity as such. I believe that there are probabilistic causes in the world, and they are appropriately called "propensities". There is the propensity of an atom to decay, a propensity of a tossed die to come to rest with the side 6 uppermost, a propensity of a child to misbehave, a propensity of a plant sprayed with an herbicide to die, etc. Such propensities produce relative frequencies. We can find out about many propensities by observing frequencies. (Salmon 1988, 14).

And:

> It is my view that—in quantum mechanics and everywhere else—physical probabilities are somehow to be identified with frequencies. One reason for this is that relative frequencies constitute an admissible interpretation of the probability calculus if the axioms require only finite additivity. . . . The conclusion I would draw is that there

---

[3] The three criteria can be found in Salmon 1967, 63–64.

> are two kinds of probabilities, personal probabilities and relative frequencies. (15)

Salmon is correct that propensities—chances—have a causal aspect. And it is exactly that aspect which makes the admissibility criterion an unreasonable constraint to place on an account of chance. For realists about chance, or at least realists about the processes that give rise to chance phenomena, theory should be required to conform to the phenomena, not the other way around. In addition, under Salmon's approach to causation, an adequate theory of causation will be only contingently and not necessarily true. To begin exploring these claims, let me add a fourth criterion of adequacy to Salmon's original three.

4. Explanatory Value: the attribution of probabilities should provide an explanation, when an explanation exists, for why the entire probability distribution has the form that it does.

4. Theory-Driven versus Frequency-Driven Models. On this account, the relative-frequency approach does poorly, for the distribution of values in a reference class or sequence is a brute fact for frequentists. Sometimes this is appropriate. In other cases it is not. We can see the difference in the distinction between what, adapting some terminology from Salmon, we can call frequency-driven probability models and theory-driven probability models. One of the standard examples of the Poisson distribution, cited often in texts, is the distribution of flying bomb hits on south London during World War II. And indeed the number of areas with k hits fits a Poisson distribution remarkably well (Clarke 1946). But no plausible explanation exists for this fact. Our theories of aerodynamics, human intentions, rocket-propelled ordinance, and so on provide no explanation of why the conditions for a Poisson distribution to hold are satisfied in this historical case.[4] So this is a clear example of a frequency-driven model—the observed distribution of frequencies fits the Poisson and that is it.

Compare that to the use of theory-driven models to account for fluctuations in electron-photon cascades in materials hit by electrons. The original, rather crude, Poisson model proposed in 1937 by Bhabha and Heitler was explicitly based upon physical theory and, independently of data, was in the same year criticized by Furry for entailing physically

---

[4] These conditions are, simplifying slightly, that the underlying stochastic process has stationary independent increments, in sufficiently small intervals at most one event can occur, and in any nonzero interval the probability of an event occurring is strictly between 0 and 1.

unrealistic consequences (Bharucha-Reid 1960, chap. 5). These physical considerations led first to a simple birth process model and subsequently to Polya process models. The crucial difference between the frequency-driven approaches and the theory-driven approaches lies in the fact that the latter explicitly describe the underlying chance processes that are assumed to give rise to the observed distribution of frequencies, whereas the former do not.

Salmon recognized this distinction, or one like it, in the 1988 paper from which I earlier quoted. In the course of that paper (Salmon 1988, 23), Salmon introduced an example which can usefully illustrate our approach. In 1980, Salmon tells us, the Pennsylvania state lottery was ‘fixed’ so as to favor certain outcomes. The apparatus used in the lottery consisted of three machines, each containing ten ping-pong balls numbered 0 through 9. The ping-pong balls were mixed by an air jet to randomize the outcomes and each machine was independent of the others. One number was drawn from each machine and the resulting three-digit number was then the winning number for that day. Each three-digit number between 000 and 999 had, on the basis of an elementary probability model, a probability of 10^{-3} of being drawn. However, on at least one occasion white paint had been injected into all of the balls except those numbered 4 or 6. It is evident that under these circumstances, the probability of numbers containing only 4s and 6s was greatly increased.

The important point about this example is that before any drawing had been made from the altered apparatus, the effects of the physical changes on the distribution of chances—and hence frequencies—of the outcomes can be predicted, at least in the sense that the general form of the distribution can be predicted. It follows that in the case of theory-driven models, chances (or propensities) based on explicit considerations of the generating conditions can satisfy the applicability criterion for interpretations of probability. Perhaps more importantly, it is by appeal to the structure of the generating conditions for the frequencies that the distributional form of the outcomes can be explained for theory-driven cases. Much emphasis has been placed by philosophers on arriving at the numerical values of probabilities—what Salmon called the ascertainability criterion. Yet from the perspective of probability theory and scientific investigation it can also be a significant achievement to predict the type of distribution associated with a phenomenon. This prediction can, in certain cases of theory-driven models, be made by analyzing the structure of the generating conditions for the chances, leaving the specific parameter values to be determined empirically, usually by frequency measurements.

5. The Contingency of Chances. This emphasis on the generating conditions fits well with the mechanistic account of probabilistic causality developed by Salmon in the early 1980s. Although mechanisms play their primary role in the propagation and production of causation, a material object considered as a structure is taken to be a special case of a process (Salmon 1984, 140). Moreover, the ability to transmit its own structure is the hallmark of a genuine process: "The basic causal mechanism, in my opinion, is a causal process that carries with it probability distributions for various types of interactions" (203).

However, it is here that I suspect we would have parted company. Salmon had rightly noted that chances are embedded within a network of causes and are a part of the world. Salmon had insisted on many occasions that a theory of causation need only be contingently true; it is sufficient to describe the way that causation operates in our world. To assume that a theory of causation that adequately describes our world would continue to be true had the world been different, especially in the laws that it obeys, was to require too much. I suggest that a similar attitude is appropriate in dealing with chance. Relative frequencies can be described by a necessarily true theory because relative frequencies are arithmetical objects. For quite different reasons, the theory of normed, nonnegative, countably additive measures on sigma-fields, which has come to be identified with the theory of probability in many quarters, is a mathematical theory, and hence necessarily true.

In contrast, how particular causes affect particular chance distributions is a contingent matter and since chances are embedded in causal networks, how chances behave is also contingent. What if we were to abstract from those causal networks and consider a theory of pure chance? In that case there is no evident reason, once we divest ourselves of our frequentist intuitions, why the values of chance should be additive nor why they should be normalized. So I suggest that the admissibility criterion is ill advised in the case of chances and that it be replaced by the explanatory-value criterion. That position fits well, I believe, with Salmon's later views on propensities, causes, and his ontic realism.

6. Concluding Remarks. Wes Salmon was an inspiring philosopher and a wonderful human being. I wish he were here to respond to this paper because his comments were always illuminating, helpful, and insightful. But in an obvious sense he always will be a participant in these discussions. For like the work of his mentor and philosophical hero Hans Reichenbach, Salmon's philosophy constitutes a permanent addition to our subject and a rich source of ideas upon which we can all build.

REFERENCES

Bharucha-Reid, Albert T. (1960), *Elements of the Theory of Markov Processes and Their Applications.* New York: McGraw-Hill Book Company.

Church, Alonzo (1940), "On the Concept of a Random Sequence", *American Mathematical Society Bulletin,*2nd series, 46: 130–135.

Clarke, R. D. (1946), "An Application of the Poisson Distribution", *Journal of the Institute of Actuaries* 72: 48.

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

——— (1979), "The Philosophy of Hans Reichenbach", in Wesley C. Salmon (ed.), *Hans Reichenbach: Logical Empiricist.* Dordrecht: D. Reidel Publishing Company, 1–84.

——— (1984), *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

——— (1988), "Dynamic Rationality: Propensity, Probability, and Credence", in James H. Fetzer (ed.), *Probability and Causality: Essays in Honor of Wesley C. Salmon*. Dordrecht: D. Reidel Publishing Company, 3–40.