The Probability Monopoly

Bart Kosko

Shuzan (926–992 A.D.) once held up his bamboo stick to an assembly of his disciples and declared: "Call this a stick and you assert; call it not a stick and you negate. Now, do not assert or negate, and what would you call it? Speak! Speak!" One of the disciples came out of the ranks, took the stick away from the master, and breaking it in two, exclaimed, "What is this?"

Daisetz Teitaro Suzuki

An Introduction to Zen Buddhism

PROBABILITY does not exist. The authors Laviolette and Seaman [5] grant this at the start and turn to deFinetti [2] for support. I agree that probability does not exist in the physical sense. It does not take up space or time in the space–time continuum we call our world. The same holds true for the existence of numbers.

The question is whether probability exists in the formal system of math. It does in the sense that probability measures are finite positive measures [1] that model "independent" sets as measure products. The authors [5] think probability is more than this but they fail to show it.

I claimed [4] that probability is not a theoretical primitive. We can often eliminate it in favor of a "fuzzy" or multivalued containment operator. The probability $P(A)$ of set A on space X is the degree $S(X \subset A)$ that the part A contains the whole X, which the identity $P(A) = P(A|X)$ suggests. Probability is the whole in the part. That makes no sense if you assume, as scientists have assumed by default since the days of Aristotle, that containment is bivalent, or black and white. It follows at once if containment is fuzzy even though the sets A and X need not be fuzzy. The proof that relative frequency equals $S(X \subset A)$ takes only two lines [4]. But who said containment is bivalent? That is an extreme boundary case, and in this case just the whim of math culture. I doubt that in a hundred or so years multivalued operators will still shock or offend anyone.

Probability is a very special case of fuzziness. It always faces two limits. First, it works with bivalent sets A. So $A \cap A^c = \emptyset$ and $A \cup A^c = X$. So $P(A \cap A^c) = 0$ and $P(A \cup A^c) = 1$ for all sets A. That forces us to draw hard lines between things and non-things. We cannot do that in the real world. Zoom in close enough and that breaks down. As Quine [7] says:

"Diminish a table, conceptually, molecule by molecule: when is a table not a table? No stipulations will avail us here, however arbitrary. . . . If the term 'table' is to be reconciled with bivalence, we must posit an exact demarcation, exact to the last molecule, even though we cannot specify it. We must hold that there are physical objects, coincident except for one molecule, such that one is a table and the other is not."

No one wants to draw a line where earth's atmosphere ends and space begins, or where life starts in a fetus, or where an electron cloud stops. These are matters of degree. Here and in general curves are more accurate than lines. Thing A overlaps non-thing A^c. Fuzzy sets are just those sets A such that $A \cap A^c \neq \emptyset$ and $A \cup A^c \neq X$. To apply probability to a problem we must assume away enough structure to draw a bivalent line. Zadeh [10] found the only way to save probability here. We must work with the probability of a fuzzy set as when the doctor says there is 90% chance that you have a mild cold. That lets fuzzy sets into sigma-algebras in a big way. But you still cannot take probabilities of too many fuzzy or nonfuzzy sets, which is the next point.

Second, probability measures need small infinities. A probability measure maps the sets in a sigma-algebra to the unit interval [0, 1]. Let c be the power of the continuum. Then the sigma-algebra can have a cardinality at most c. The Borel sigma-algebra $B(R^n)$ on R^n has cardinality c [8]. That is why we use it when we work with probabilities on real intervals or interval products and not the more intuitive real power set 2^{R^n} that has cardinality 2^c. The real power set is too big for a probability space. The gap between c and 2^c is vast. But every space has fuzzy subsets no matter how great its cardinality. These fuzzy subsets and their nonfuzzy neighbors contain one another to some degree. So the whole-in-part relation degree $(X \subset A)$ holds to some degree even though $P(A)$ is not defined.

I used the inexact ellipse to make this point. There is no trouble mapping all geometric figures to fuzzy shades of gray in [0, 1]. We can even use exact error measures to give us the operational definitions of a fuzzy ellipse that the authors [5] seem to feel belong to probabilists. There are at least 2^c sets of fuzzy ellipses. Probabilists do not have this ease of definition and this set of choices. The set of possible geometric figures is too big to admit a probability space. Even the set of (continuous) curves, which has cardinality c, makes trouble. The authors [5] would like to say that each inexact oval has some probability that it is an ellipse. They say so but do not show so. The trouble is each figure has the status of a real point. So each figure has probability zero. Of course an inexact oval also makes us doubt whether there is a "randomness" that hovers about the oval. I do not think there is.

Manuscript received July 18, 1992; revised September 12, 1992.

The author is with the Department of Electrical Engineering, Signal and Image Processing Institute, University of Southern California, Los Angeles, CA 90089-2564.

IEEE Log Number 9214997.

The authors [5] hold the Bayesian view that probability is a "state of knowledge" or brain state. They claim that probability models this brain state and models it better than fuzzy math can model it. The claim is wrong or moot if probability is a subset of fuzzy set theory. It also smacks of a retreat and means that a branch of math would not be or would not apply if there were no brains in the world—as there were not until a billion or so years ago. The authors say, "Uncertainty, like beauty, exists in the eye of the beholder," Sure it does. But the stable relative frequencies we observe when we flip coins, or hybridize peas, or watch atoms vibrate suggest that the uncertainty, its formal structure, lies in more matter than just our brains. That is what led our brains to work out probability theory in the first place.

Still I think the Bayesians are on to something when they find probability in our minds. To capture it they cite the axioms of Savage or Kolmogorov or deFinetti. But what they are after does not depend on axioms. It depends on genes and the minds they have grown in the past billion years or so of neural evolution.

We have a *probability instinct*. We are forward-looking creatures. Natural selection favors organisms that can rank future alternatives with some accuracy. Will it rain? Do the mud prints mean a deer is near? Should we attack? Will the rock hit the target? Mammalian evolution alone has had a quarter-billion years to shape our uncertainty "reasoning" or hunches. This has made probability or "randomness" a type of Jungian archetype in our psychology. The selecting environment has built it into the structure of our thoughts along with time and space and causal connection.

Eons later Western theorists saw some of these intuitions reflected in the bivalent and artificially precise games of chance of their day. This was a cultural fluke—as when our ancestors first ended up with five fingers per hand instead of four or six, or when we first started driving on the right side of roads in America and the left side in Britain. On this rare branch of the cultural tree, probabilists imposed in the tradition of Aristotle, the bivalent structure of games of chance onto the world and into the math of uncertainty. Then there were no uncertainty competitors for a few hundred years. In that time modern science and bivalent math took shape. By default this gave the probability view a monopoly on uncertainty.

It could have turned out otherwise. The authors [5] chide my "multiculturalism" when I suggest that our math and world view might be different today if modern math had taken root in the A-AND-not-A views of Eastern culture instead of the bivalent A-OR-not-A view of ancient Greece. To dismiss this as "unfortunate deconstructionism" is just to name call and to ignore historical fact. Both Lao-Tze and the Buddha championed the A-AND-not-A view of simultaneous opposites. The Taoist yin-yang symbols makes this clear and today adorns the flag of South Korea and Mangolia. The Buddha built his whole world view on first breaking out of the black–white shell of words that still binds much of Western culture and the modern science it spawned. This lies at the heart of *satori* enlightenment [9] in Zen Buddhism in Japan. Today Japan and South Korea lead the world in fuzzy commercial products. That may well involve a cultural influence. In any case I cannot imagine any major Eastern thinker who would claim that $P(A \cap A^c) = 0$ holds for *all* events A. That is the height of logical and cultural extremism. The probability monopoly rests on it.

Fuzzy theory challenges the probability monopoly. Probabilists have attacked it with gusto to keep their monopoly status, to have, as Jaynes [3] and Lindley [6] want, the only uncertainty theory in the unit interval [0, 1]. But the fuzzy math is sound. Its world view of shades of gray has a deep intuitive ring. And the new fuzzy products have come into their own in the marketplace. The probability monopoly is over.

References

[1] K. L. Chung, *A Course in Probability Theory.* New York: Academic, 1974.

[2] B. DeFinetti, *Theory of Probability: A Critical Introductory Treatment,* vol. 1. New York: Wiley, 1974.

[3] E. T. Jaynes, "Where do we stand on maximum entropy?" in *The Maximum Entropy Formalism,* Levine and Tribus, Eds. Cambridge: M.I.T. Press, 1979.

[4] B. Kosko, "Fuzziness vs. probability," *Int. J. General Syst.*, vol. 17, no. 2, pp. 211–240, 1990.

[5] M. Laviolette and J. W. Seaman, Jr., "The efficacy of fuzzy representations of uncertainty," *Int. J. Approx. Reason.*, Special Issue on Probability vs. Fuzzy Sets, 1992.

[6] D. Lindley, "The probability approach to the treatment of uncertainty in artificial intelligence and expert systems," *Stat. Sci.*, vol. 2, pp. 17–24, 1987.

[7] W. V. O. Quine, "What price bivalence?" *J. Philosophy,* vol. 78, no. 2, pp. 90–95, Feb. 1981.

[8] W. Rudin, *Real and Complex Analysis.* New York: McGraw-Hill, 1974.

[9] D. T. Suzuki, *An Introduction to Zen Buddhism.* New York: Grove, 1964.

[10] L. A. Zadeh, "Probability measures of fuzzy events," *J. Math. Analysis and Appl.*, vol. 10, pp. 421–427, 1968.