activity concerning the definition of conditioning and its relation to the familiar Bayes rule. When one interprets conditional probabilities as updates of probabilities in the light of new evidence, then it is suggested that we may have more flexibility in the choice of updated conditional probabilities than is allowed in classical probability (Diaconis and Zabell, 1982; Shafer, 1982). Conditional probability in a lower envelope setting is thoroughly treated in Walley (1987, Chapter 7).

Finally, subjective probability is complementary to objective, frequentist-based probability, but the two approaches taken together neither exhaust the domain of random phenomena nor the possible interpretations for the axioms surveyed (Fine, 1983). Neither theory accounts for intrinsic limits to the precision with which we can model random phenomena when we need to account for hesitancy in the case of individual beliefs and unstable relative frequencies in the frequentist case. Nor do they exhaust the possibilities for interpreting probability and reasoning about random phenomena.

ADDITIONAL REFERENCES

DIACONIS, P. and ZABELL, S. (1982). Updating subjective probability. *J. Amer. Statist. Assoc.* **77** 822–830.

FINE, T. L. (1983). Foundations of probability. In *Encyclopedia of Statistical Sciences* (S. Kotz and N. Johnson, eds.) **3** 175–184. Wiley, New York.

GRIZE, Y. L. and FINE, T. L. (1986). Continuous lower probability-based models for stationary processes with bounded and divergent time averages. To appear in *Ann. Probab.*

KAPLAN, M. (1971). Personal communication reported in Fine (1973, page 31).

KUMAR, A. and FINE, T. L. (1985). Stationary lower probabilities and unstable averages. *Z. Wahrsch. verw. Gebiete* **69** 1–17.

KYBURG, H. E., JR. (1974). *The Logical Foundations of Statistical Inference*. D. Reidel, Dordrecht, Netherlands.

LEVI, I. (1980). *The Enterprise of Knowledge*. MIT Press, Cambridge.

PAPAMARCOU, A. and FINE, T. L. (1986). A note on undominated lower probabilities. *Ann. Probab.* **14** 710–723.

SHAFER, G. (1982). Bayes's two arguments for the rule of conditioning. *Ann. Statist.* **10** 1075–1089.

WALLEY, P. (1987). *Rationality and Vagueness*. Chapman and Hall, London. In preparation.

Comment

Teddy Seidenfeld

This essay provides an informative overview of axiomatic theories whose common theme is a development of quantitative personal probability from the qualitative or comparative binary relation '$A > B$', understood as "A is subjectively more probable than B." The pathbreaking works of Ramsey, de Finetti, and Savage contribute to this project by giving the comparative probability relation an operational, decision theoretic basis. Roughly put, they propose that A is subjectively more probable than B provided the lottery L_A, having a desirable prize awarded if event A occurs and status quo otherwise, is preferred to ($\succ$) the lottery L_B which has the desirable prize awarded if B occurs.

Definition

$$(1) \qquad A > B \quad \text{iff} \quad L_A \succ L_B.$$

In Savage's hands, quantitative personal probability is reduced to the qualitative relation $\geq$ which, in turn, is reduced to (weak) preference among lotteries $\succsim$. For Ramsey and de Finetti, quantitative probability may be "elicited" directly from choices among gambles and agrees with the comparative relation $>$ (defined above). But the common thread is that rational belief is constrained by coherent preference, and binary choices reveal preferences.

In Section 3 of his paper, Professor Fishburn turns his attention to theories of personal probability different from the strict Bayesian position of Ramsey, de Finetti, and Savage. Specifically, he rightly considers a liberalization which relaxes the assumption that $>$ is a weak order. To understand why this is a reasonable change from the norms of strict Bayesianism recall, e.g., Smith's (1961) idea for "medial odds," to permit a spread in the odds as Levi (1980, Section 7.3) so aptly puts it.

Consider a wager on event A with a combined stake s at odds $p:1-p$ ($0 \leq p \leq 1$). You bet on A by putting up ps (which is lost in case A fails to occur) with the prospect of winning $(1-p)s$ in case A occurs. (These wagers are a special case of Smith's bets "on A against B," obtained by letting B be the sure event.) Also, there is the associated wager against A, equivalent to a bet on $\neg A$ at odds of $(1-p):p$, where you place $(1-p)s$ on $\neg A$, lost in case A occurs, with the prospect

Teddy Seidenfeld is Associate Professor of Philosophy, Department of Philosophy, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

of winning ps provided $\neg A$ occurs. Suppose the stake is in units "linear in utility." (This is an assumption in the "Dutch Book" arguments of de Finetti (1975, Sections 3.1.4 and 3.2), Shimony (1955), and Smith. Specifically, what is needed is that the combination of favorable/indifferent/unfavorable bets is favorable/indifferent/unfavorable.)

Now imagine you face a choice among three options:

O_1 bet on A at odds $p:(1-p)$,

O_2 bet on $\neg A$ at odds $(1-p):p$,

and

O_3 abstain from betting altogether.

Then, according to the strict Bayesian theory, your preferences satisfy one of three profiles:

$$O_1 \succ O_3 \succ O_2,$$

or

$$O_2 \succ O_3 \succ O_1,$$

or

$$O_1 \approx O_2 \approx O_3.$$

You ought to be content with the life of the gambler. Regardless of the odds posted, there is no Bayes profile which corresponds to a strict preference for abstaining. Your "fair" odds, the point at which you are indifferent among the three options, identifies your quantitative personal probability. It agrees with your qualitative probability, a weak order, defined by (1).

Smith's proposal is that an agent be allowed to refuse betting either on or against an event without thereby supposing he is indifferent among the three options. Smith raises the possibility (outside the strict Bayesian norms) that, *for more than one odds ratio*, the agent chooses to abstain rather than to bet. The idea here, which has been aired by others before and since, is a representation of belief by a (convex) set **P** of probability functions. (See, for example, Levi (1974) and Giron and Rios (1980) for related discussion.)

I understand Smith to suggest that the agent bets on or (exclusively) against A at the posted odds whenever this maximizes expected utility for *each* probability in the set **P**. Otherwise, i.e., when neither bet is judged advantageous with respect to **P**, the agent chooses to abstain. (Smith opts for abstaining also if the posted odds are "fair.") This proposal makes Bayesian theory the special case when **P** is a unit set.

Levi (1980, Section 7.3) gives a detailed account of this interpretation. When the posted odds are "medial," each of the three options maximizes utility for some p in **P**. Then the three options are noncomparable in terms of preference, $\succsim$. Levi advocates a lexicographic decision procedure where a consideration of "security," e.g., maximin with respect to payoffs, determines which of the preferentially noncomparable options is (are) admissible. Thus, refusing to bet is uniquely admissible whenever the posted odds are "medial" and "security" is assessed by maximizing the minimum payoff (or by maximizing the minimum expectation with respect to **P**). This gives a smooth reconstruction for Smith's account of how the gambler behaves when he holds a spread in his odds.

Here we find a sound, decision theoretic basis for removing the assumption that (normatively) the preference relation $\succsim$ is complete. Not surprising, the qualitative probability relation $>$ based on this modified preference relation also admits noncomparability. However, the resulting qualitative relation, $A > B$, is not an *interval order*. Nor is it equivalent to determing whether the upper probability $\mathbf{P}^*(B)$ is less than the lower probability $\mathbf{P}_*(A)$.

Here is a simple illustration of that point.

Example 1. Let the agent's medial odds on B correspond to the set $\mathbf{P} = (p\colon .2 \le p(B) \le .9)$; hence, $\mathbf{P}_*(B) = .2$ and $\mathbf{P}^*(B) = .9$. Let C be a probabilistically independent event with a determinate probability .5, i.e., $p(C \mid B) = .5$ for each p belonging to **P**. For instance, let B be the event "precipitation in Pittsburgh on April Fool's Day, 1986," and let C be the event "the next flip of this 'fair' coin lands heads-up." Define the event $A = B \cup C$. Then the agent (strictly) prefers the lottery L_A to the lottery L_B; hence, $A > B$ according to the definition (1). But $\mathbf{P}_*(A) = .6$ and $\mathbf{P}^*(B) = .9$, contrary to the representation (related to an interval order):

$$A > B \quad \text{iff} \quad p(A) > p(B) + \sigma(B),$$

where $p = \mathbf{P}_*$ and $\sigma(B) = [\mathbf{P}^*(B) - \mathbf{P}_*(B)]$.

There is more in the Ramsey–de Finetti–Savage program that is lost under this interesting liberalization of strict Bayesian theory. Not only does preference admit noncomparability of options, it fails to be basic binary as well. That is, the preference relation on a set of options is not defined in terms of the preference relation on its paired subsets.

Example 2. Let D be an event of maximally indeterminate probability, $\mathbf{P}_*(D) = 0$ and $\mathbf{P}^*(D) = 1$. Consider a choice among these three options: to bet on or (exclusively) against D at even odds, or to pay an (insurance) fee of .1 utile to refuse to bet. Although in any of the three paired comparisons both options are noncomparable by $\succsim$, i.e., each maximizes expected utility for some p in **P**, in the choice among the three options the insurance is $\succsim$*-inadmissible* (*dispreferred*) since it fails to maximize expected utility for each p in **P**.

This phenomenon constitutes a violation of Sen's (1977, page 64) Property γ. By his Proposition 8 (1977, page 64), the choice function given by $\succsim$ is not *normal* or, equivalently, $\gtrsim$ fails to be basic binary. (If we include Levi's lexicographic "security" considerations in determining admissible choices, then the choice function violates Sen's (1977, page 64) Property α, as Example 2 illustrates. With "security," the insurance is uniquely admissible in a choice between it and one of the two bets. But the insurance is inadmissible as a choice among the three options.)

In *Foundations* (Section 7.2) Savage uses a group decision rule that fails to satisfy his **P1**: the postulate that preference is a weak order, where an option is admissible if it maximizes utility for some p in $\mathbf{P}$. (The set $\mathbf{P}$ corresponds to the convex combination of personal probabilities held by the individuals in the group.) Again, in Section 13.5, he defends group decision rules that violate **P1**. Is it not wise to propose the same norms in group and in individual choices? I think so. In that case, we can adopt Savage's own reasons to argue for the liberalization of strict Bayesian norms proposed by Smith, Levi, et al. But what is left of Savage's project to complete the reduction of quantitative personal probability to choice (without extraneous notions of probability)? Can the set $\mathbf{P}$ be recovered from choice behavior without the tacit assumption of a utility function as used in the "Dutch Book" arguments, i.e., without requiring that the conjunction of favorable gambles be favorable? I believe that remains an open question.

ADDITIONAL REFERENCES

GIRON, F. J. and RIOS, S. (1980). Quasi-Bayesian behavior: A more realistic approach to decision making? In *Bayesian Statistics II* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, eds.) 17–38. University Press, Valencia.

LEVI, I. (1974). On indeterminate probabilities. *J. Phil.* **71** 391–418.

LEVI, I. (1980). *The Enterprise of Knowledge*. MIT Press, Cambridge.

SEN, A. K. (1977). Social choice theory: a re-examination. *Econometrica* **45** 53–89.

SHIMONY, A. (1955). Coherence and the axioms of confirmation. *J. Symbolic Logic* **20** 8–20.

Comment

Mervyn Stone

Dr. Fishburn's review is both authoritative and timely. It is good to see a paper that dares, in this new journal, to propagate by style and content the best *Annals* tradition—clear exposition and comparison of important mathematical structures, unclouded by the polemical discussion that inevitably arises when mathematical concepts are ultimately related to the problems of induction and decision.

It will be interesting to see if the present discussants let him get away with it. They may not—for the simple reason that there is a sizeable school of "infinitarians" who will be disposed to sift through Dr. Fishburn's fine deposits for items to advance their cause (see Scozzafava (1984) for examples of the art). For the purposes of this discussion, an "infinitarian" is one who will not countenance the restriction to countable additivity, and is prepared to defend any implications of this stand, including those that are regarded by some as manifest counterexamples to the view that finite additivity rules OK.

Dr. Fishburn raises a polemical little finger, as it were, when he states that the assumption of monotone continuity is "quite appealing." It is somewhat paradoxical that the effect of the monotone continuity axiom, whose statement involves countable infinities, is to allow one then to forget about "infinity" as a point in the sample space, and get on quietly with the job of using infinity, in the sample space as a whole, as a framework for useful approximation of necessarily finite, practical induction and decision. In contrast, the axioms for merely finite additivity do not explicitly involve infinity, but have unresolved problems of infinity that ought, I think, to disturb the practical inferencer or decision-maker who adopts a finitely additive P of the type in question.

One widely considered example has

$$S = \{(x, \theta): x = 1, 2, 3, \ldots, \quad \theta = 0, 1\}$$

with

$$P(\Theta = 0) = P(\Theta = 1) = 1/2,$$
$$P(X = x \mid \Theta = \theta) = 2^{-(x+\theta)}.$$

Note the missing probability 1/4 in the countable union of $(x, 1)$, $x = 1, 2, 3, \ldots$. Formally, P is "nonconglomerable in the x margin." The setup implies

$$(1) \quad \left.\begin{array}{l} P(X > 12) > 1/4 \\ P(\Theta = 1 \mid X > 12)/P(\Theta = 0 \mid X > 12) > 1000 \end{array}\right\}.$$

Mervyn Stone is Professor of Statistics, Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, United Kingdom.