Comment

Simon French

This is an excellent review of the literature: all-embracing, scholarly, thought-provoking. Genest and Zidek are to be congratulated on organizing and summarizing such a wide body of the literature so very effectively. I, for one, am very grateful to them. Given the paper's undoubted virtues, it seems almost churlish to criticize; but praise alone makes for dull, unconstructive discussion, and I do take issue with the sentiments behind their opening paragraphs.

Genest and Zidek suggest that a common motive for studying the aggregation of opinion problem is to find some concept of consensus, which might replace the concept of objectivity which is supposedly central to the scientific method. While this is certainly a commonly expressed motive, it is not one that bears inspection by those of us who call ourselves subjectivists.

De Finetti begins his treatise with the now famous sentence "probability does not exist." Probabilities do not model some physical properties or propensities that exist in some objective sense within a system. If one tries to interpret them as such, as indeed the frequentists do, one is led into a maze of paradox and inconsistency from which ad hockery provides the only escape. To a subjectivist, probabilities model the beliefs of an observer of a system. But, in fact, they do more than this. The probability calculus provides procedures whereby the observer may guide the evolution of his beliefs in such a way that he is led to self-consistency. Probabilities are truly subjective. They belong to an individual observer; they are part of his thought processes.

Simon French is a Senior Lecturer in the Department of Decision Theory, University of Manchester, Manchester, M13 9PL, United Kingdom.

Accepting this, one is immediately led to the question: what to one person is a second person's probability? The answer a subjectivist must give is: "not a probability." For once it has left the second person's mind it is no longer part of his thought processes and, therefore, not a probability. Thus an average or aggregate "probability" taken over a group of observers is certainly not a probability. It belongs to no one. Averaging over individuals does not make probabilities more "objective" in any sense. It destroys probabilities.

Genest and Zidek are right to point to the move toward identifying intersubjectivity currently being made within the Bayesian community. But intersubjectivity and averaging are very different. Intersubjectivity studies are certainly concerned with groups of observers of the same system, but they are not concerned with identifying an average opinion. Rather, these studies are concerned with identifying circumstances under which the thought processes of the group members will run in parallel. When will their individual probabilities happen to take the same numerical values or, more generally, when will their individual probability distributions happen to share a common functional form? Intersubjectivity is about consensus in the *strict* sense of that word, that of unanimous agreement.

In the light of these points it will be understood why in my own work on the aggregation problem I have always addressed the question of how one individual should update his beliefs on hearing the opinions of others. It will also be understood why I do not believe that any of the work on aggregation will lead to a replacement for the concept of objectivity.

Comment

Robert L. Winkler

I would like to congratulate Genest and Zidek on an excellent job of reviewing a wide variety of work involving the combining of probability distributions. The increasing interest in using subjective probabilities in practice has led to greater attention being given to the problem of combining such probabilities, making this review and bibliography particularly welcome and timely. In this discussion I will indicate my own biases about fruitful approaches for combining probabilities and probability distributions.

Robert L. Winkler is IBM Research Professor, Fuqua School of Business, Duke University, Durham, NC 27706.

As indicated by Genest and Zidek, one stream of work has focused on the form of the combining rule, specifically on what formulas can be derived from various sets of assumptions. Another direction has been to view the experts' probabilities as information and to aggregate this information, explicitly or implicitly modeling the experts through the likelihood functions that are used in the aggregation process. Although it is certainly interesting and informative to see the formulas that arise under certain sets of assumptions, I find it preferable to approach the combination of probability distributions from a modeling perspective. The emphasis in the literature, however, has been more on the former than on the latter, perhaps because the former is "cleaner" and more amenable to obtaining general results.

Contrary to Genest and Zidek's characterization of the assumptions that have been studied as "rather weak and qualitative," I would claim that most are too strong and restrictive. For example, as noted by them, the condition of external Bayesianity implies that the aggregation rule, or pooling mechanism, is the same before and after new data are observed. A situation where this might be reasonable is one in which the experts are viewed as exchangeable *a priori* and they all see the same new data. In most cases, however, I would expect the weights assigned to the experts (in a linear opinion pool, a logarithmic opinion pool, or yet some other aggregation formula) to change as new data are seen.

The independence-preservation property is also more restrictive than it might appear at first glance. Harrison (1977) presents an intriguing analysis to show that uncertainty about one's own calibration can lead to seemingly unrelated events being perceived as dependent for a single expert. If this problem can arise for a single expert, can we expect to preserve independence across experts? In the same vein as Harrison's work, suppose that experts have somewhat similar "error processes" via shared cognitive biases. Uncertainty about the parameters of these error processes could cause violations of the independence-preservation property, just as uncertainty about nuisance parameters can complicate statistical analysis in general.

Another example of an innocuous-sounding assumption that is much more problemátic than it might seem is the assumption that if all of the experts agree on the probability of an event, then the aggregated probability should equal the experts' common probability. If all of the experts consulted say that the probability of E is 0.6 (i.e., $P_1 = P_2 = \cdots = P_n = 0.6$), should the combined probability also be 0.6? On the surface, this does not seem unreasonble, but it depends on the situation at hand. Suppose that I am about to draw a ball at random from an urn, E corresponds to the ball being red, and the experts' probabilities are based on samples from the urn. If the proportion of red balls in the urn can take on any value, then a combined probability of 0.6 makes sense. But if the proportion of red balls in the urn is known to be either 0.2 or 0.8, then each successive P_i makes it seem more likely that the proportion is 0.8. As a result, the combined probability would be somewhere between 0.6 and 0.8, violating the assumption that it should equal the experts' common probability. Dependence among the experts (e.g., from overlapping samples), uncertainty about calibration (e.g., possible miscounting when seeing the samples), and my own prior probability for E can also affect the combined probability, but the purpose of the example is to show that the assumption can be violated without these complications.

An axiomatic approach to combining probabilities presented by Morris (1983) has generated considerable interest and numerous written reactions. A set of such reactions from Clemen (1986), French (1986), Lindley (1986), and Schervish (1986) has just been finalized for publication in *Management Science* along with a reaction from Morris. Much of the discussion focuses on the assumptions investigated by Morris and relates to the issue of starting from assumptions vis-à-vis modeling. Ironically, the concluding comments in Morris (1983) seem to be very much in the spirit of modeling.

My own feeling is that different combining rules are suitable for different situations, and any search for a single, all purpose, "objective" combining procedure is futile. Moreover, the question of who decides upon the general form and the specific parameters of a combining rule is important and cannot be swept under the rug. However, this issue is just as crucial when working from basic assumptions as it is when taking the modeling approach. Since there is no single combining procedure for all seasons, a subjective element cannot be avoided in the acceptance of basic assumptions or in the modeling of experts. An analogy may be drawn with classical statistics, which involves elements of subjectivity in practice despite the desire of some statisticians to view it as "objective." Some situations will engender more agreement than others on the combining rule that seems most appropriate, but an "objective" rule is an unattainable goal.

Another issue that seems to be a red herring is that of the form in which the judgments of uncertainty are expressed. Probabilities, odds, and log odds are equivalent, as are cumulative functions and mass or density functions. Why, then, should combining rules dealing with odds, log odds, or cumulative distributions be treated separately? In terms of the elicitation of probabilities, the relevant question is which measures are easiest for an expert to think about and work with effectively. In terms of combining rules, the relevant question is one of modeling convenience.

In general, the modeling approach allows a great deal of flexibility. In addition to not being tied to a specific form of combining rule, it allows judgmental inputs other than probabilities. Some experts might give entire probability distributions, others might give a few probabilities (e.g., an interval of values with a probability, a few cumulative probabilities), others might give one or more summary measures (e.g., a mean), and yet others might provide only qualitative information (e.g., a statement that an event is quite likely). Through the modeling approach, any of these types of information can be included.

The form of a combination rule, not just the specific parameters for a given form (e.g., the weights in a linear opinion pool), should be based on modeling considerations. For example, suppose that each expert provides a density for a variable x. Expert j's density might be viewed as being based on certain underlying conditions C_j embodying expert j's model of the world as it relates to x. If we view the sets of conditions C_1, $C_2, \cdots, C_n$ as alternate possibilities and assign probabilities $P(C_1), P(C_2), \cdots, P(C_n)$ summing to one, the result is a linear opinion pool with weights $w_j = P(C_j)$ for $j = 1, \cdots, n$. This amounts to saying that one of the experts is "right" and we are uncertain about which one it is. On the other hand, it might seem more reasonable not to think of experts as being right or wrong, but to take a different tack. If the experts are viewed as independent, perfectly calibrated information sources, with the joint likelihood equal to the product of individual likelihoods, then a multiplicative rule would result.

In some cases we may wind up with complicated modeling that involves both linear and multiplicative combination rules. Suppose that our modeling leads us to a multiplicative combination rule but that we are uncertain about some relevant parameters (e.g., the degree of dependence among information sources). If we model this uncertainty with a discrete distribution over the parameters, then the final combined probability will be a mixture of multiplicative formulas (i.e., a linear opinion pool for combining a variety of possible logarithmic opinion pools).

From a practical standpoint, the modeling approach may be difficult to apply in actual situations. Therefore, simple combination rules that treat the experts symmetrically (e.g., a simple average or geometric mean) deserve careful attention. Such rules may be relatively robust in addition to being each to use. In recent work involving the combination not of probabilities but of point forecasts (e.g., Makridakis and Winkler, 1983; Winkler and Makridakis, 1983; Clemen and Winkler, 1986), simple averages of forecasts tended to perform at least as well as, and usually better than, fancier combination rules. This raises the question of whether modeling in the combination of probabilities can be effective enough in practice to justify the time and effort that it requires. More research is needed on the modeling of experts (and other sources of probabilities, e.g., econometric models in economic forecasting), including work on complicating factors such as calibration and dependence. The consideration of procedures involving group interaction or feedback is also worthy of study. The review and bibliography of Genest and Zidek sets the stage for future research by making work from a wide variety of outlets readily available, indicating the current "state of the art," and suggesting some promising directions for future work.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grants PRA8413106 and ATM8507495.

ADDITIONAL REFERENCES

CLEMEN, R. T. and WINKLER, R. L. (1968). Combining economic forecasts. *J. Bus. Econ. Statist.* **4** in press.

FRENCH, S. (1986). Calibration and the expert problem. *Manage. Sci.* **32** in press.

HARRISON, J. M. (1977). Independence and calibration in decision analysis. *Manage. Sci.* **24** 320–328.

LINDLEY, D. V. (1986). Another look at an axiomatic approach to expert resolution. *Manage. Sci.* **32** in press.

MAKRIDAKIS, S. and WINKLER, R. L. (1983). Averages of forecasts: some empirical results. *Manage. Sci.* **29** 987–996.

SCHERVISH, M. J. (1986). Comments on some axioms for combining expert judgments. *Manage. Sci.* **32** in press.

WINKLER, R. L. and MAKRIDAKIS, S. (1983). The combination of forecasts. *J. R. Statist. Soc. Ser. A* **146** 150–157.