Structural Safety 52 (2015) 144–149

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

Aggregation of not independent experts' opinions under ambiguity

Marcello Basili [a,*], Luca Pratelli [b]

[a] DEPS, University of Siena, Italy
[b] Naval Academy of Livorno, Italy

ARTICLE INFO

Article history:
Received 22 November 2013
Received in revised form 4 March 2014
Accepted 1 April 2014
Available online 1 May 2014

Keywords:
Ambiguity
Aggregation
Entropy
Multiple priors
Quantiles

ABSTRACT

We consider an aggregation scheme of opinions expressed through different probability distributions or multiple priors decision model. The decision-maker adopts entropy maximization as a measure of risk diversification and a rational form of prudence for valuing uncertain outcomes. We show a new aggregation rule based on the composite value function that is able to represent asymmetric attitude on extreme events (optimism with respect to windfall gains and pessimism with respect to catastrophic events) and a rational prudence on ordinary events. We define when the new rule preserves stochastic dominance.

1. Introduction

In recent years, literature aggregating experts' opinions on ambiguous events has been increasing in climatology, environmental sciences, medical disciplines, computer sciences, engineering, law, economics, etc. [1–7]. Combining experts' opinions is a formal process for eliciting a common judgement in the form of subjective probability distribution, called consensus distribution, about the value of some decision-relevant quantities or event occurrences. Experts' judgements elicitation is a multidisciplinary process apt to fill data gaps or partial scientific knowledge. Mathematical and behavioural approaches for combining experts' opinions have been proposed and used. Mathematical Bayesian aggregation models manage individual probability distributions to obtain a single combined probability distribution operating with different degrees of complexity: equal-weight, best-expert, copula, etc. [8–13]. Judgements in the form of subjective probability distributions can be obtained from subject-matter experts by interacting with them (sharing assessment) in elicitation protocols such as Delphi, Q-Methodology, Nominal Group Technique and Kaplans approach. Even though elicitation protocols suffer from many problems such as polarization, strategic manipulation, overconfidence, self-censorship, pressure to conform, and more extreme probability estimates in order to generate some kind of consensus distribution [14,15], a number of algorithmic approaches based on Bayesian theory have been offered [16,17].

If the decision-maker (DM) does not have access to a reliability measure for each expert (e.g. likely loss or other measures of effectiveness), a measure of correlation between experts' judgements (e.g. experts' calibration and weighting) and faces ambiguous events, a situation may arise in which there is not enough information to form a unique reliable probability distribution. Instead, a bounded set of reasonable probability distributions considered reasonably possible may be formed. Therefore, the combination of experts' opinions using a Bayesian statistical approach can lead to inconsistent and incoherent consensus distribution.

Here, we introduce an approach to form a consensus distribution that adopts the quantile function in a setting with multiple priors. The quantile is a generalization of the concept of median and quantile-based decision criteria shows an increased influence in statistical literature. The set of all probability distributions of each expert on possible events, i.e. a closed and convex set of probability distributions, reflects incomplete knowledge and inadequate empirical data. The multiple priors approach is a method for probability-based ambiguity characterization, and the set of experts' probability distributions can be considered to reflect the DM's assessment of the reliability of available information, that is, the perception of ambiguity among all possible consequences. The DM distinguishes between a set of ordinary or familiar outcomes, which are considered more reliable and closer to the DM's life experiences, and two tails that include results attached to more uncertain and extreme (unfamiliar) events. As experimental and behavioural literature shows the existence of a positive correlation between the DM's attitude to extreme

* Corresponding author. Address: DEPS, University of Siena, Piazza San Francesco 7, 53100 Siena, Italy. Tel.: +39 0577232787; fax: +39 0577232739.
E-mail address: basili@unisi.it (M. Basili).

http://dx.doi.org/10.1016/j.strusafe.2014.04.001
0167-4730/© 2014 Published by Elsevier Ltd.

outcomes and the competence effect [18–20], the DM is supposed to be pessimistic with respect to purely catastrophic losses, neutral on ordinary outcomes and optimistic with respect to windfall gains. For ordinary outcomes, the DM adopts diversification as opposed to concentration, an intuitive and consistent strategy for reducing likely loss, to derive an appropriate probability distribution among the set of multiple priors. As entropy reflects the diversification degree of a portfolio [21] and the maximum entropy density is the *least informative* without introducing extra ad hoc information, the DM makes use of the maximum entropy principle to elicit the probability distribution, which is prudent when valuing ordinary outcomes. When working with the two tails, the DM adopts the maximin and the maximax approaches, respectively, and obtains a fat-tailed consensus distribution. The DM elicits the consensus distribution by combining the particular probability distribution π_j that maximizes the entropy in the set of familiar events and is the probability distribution *closest to uniformity* with maximin and maximax distributions on the two tails. The suggested aggregation rule can preserve stochastic dominance, in some circumstances.

The paper is organized as follows: Section 2 provides a discussion of related literature. Section 3 introduces the new aggregation rule. Section 4 shows a very simple application of the new aggregation rule with three experts. Section 5 shows how the suggested aggregation rule preserves stochastic dominance. Section 6 includes some concluding remarks.

2. Related literature

Failure of the frequentist theory due to incomplete, sparse or unavailable data made the Bayesian theory the normative theory for solving the problem of aggregating experts' opinions. Some problems exist with the full implementation of consensus rules in an axiomatic Bayesian approach to the experts' priors and updating by a likelihood function. In fact, "although the Bayesian aggregation methods are theoretically appealing, difficult issues remain concerning how to characterize the degree of dependence among the experts and how to determine the quality of the expert judgements" [7] such as: the arbitrariness of the pooling weights, the use of invariant combination rules, the dependence on the experts' information, dependence on experts' probability distributions (e.g. stochastic dependence) and calibration of experts' opinion [22,23]. In empirical studies, Clemen and Winkler [24] and Figlewski and Urich [25] found that correlations among expert forecasts can be above 0.80. In fact, "in some situations leading to a weighted average as a combining rule, high dependence may cause very extreme weights (e.g., highly negative weights) and combined probabilities that are much higher or lower than any of the individual experts' probabilities....Another awkward situation is one in which there are large disagreements among experts without readily identifiable explanations. For example, suppose that several experts assess density functions for a random variable and the densities hardly overlap. Here the modeling can be critical in the sense that the results are very sensitive to the nature of the model. Under some assumptions, the combined density function could be extremely spread out, much more so than any of the individual densities; under other assumptions, the combined density function could be tighter and could have very little overlap with any of the individual densities. Much attention has been paid to situations with agreement (e.g., the example where everyone agrees on a probability of 0.55)" [26]. In scenarios characterized by ambiguity and stochastically dependent experts' opinions, the Bayesian axiomatic approach to consensus distribution does not appear satisfying, even in the sophisticated versions of copula models.

Ambiguity is different from mathematical risk and emerges when individuals face vague and incomplete information. Ambiguity influences the perception of possible actions and induces human beings to elicit probabilities and apply decision rules that violate axioms of the rationality paradigm, based on the Bayesian approach, of the expected utility theory. Decision theory under uncertainty rests on the Savage subjective expected utility theory [27], even if a representation of uncertainty exists that is based on objective probability distributions in outcomes or lotteries [28]. The Savage approach centres around two fundamental assumptions: a complete and available list of possible future states of the world and subjective beliefs over the state space (uncertain prospect) that are represented by a well-defined (additive) probability function. In uncertain settings, individuals are supposed to be able to undertake expected cost/benefit analysis in information gathering and therefore reach an informational optimum satisfying the requirement of consistency (Dutch Book or Arbitrage in Gambling). Information processing consists of updating the prior probability distribution, during which a signal is received on the realization of the state (Bayes rule). These assumptions follow from the implicit hypothesis that individuals are rational in a strong sense, that they have a complete knowledge of all possible states of the world and can manage to deduce all logical propositions contained in the axioms of the theory. Both assumptions have been questioned and abandoned in: bounded rationality models [29,30] and non-expected utility models [31–35] in which agents have distorted probabilities, contractions or expansions of prior linear probabilities that are capable of accommodating individuals' perception of probabilities through weighting functions.

In the past two decades, different approaches have been proposed to calibrate the aggregation of experts' opinions through the DM's ambiguity attitude. These methods can be included in three main classes distinguished on the base of ambiguity representation: methods based on Dempster's rule of combination or theory of evidence [36–38], combination rules based on possibility distributions and fuzzy measures [39–41] and methods of aggregation based on multiple priors or capacity [42–44].

Among such different approaches developed for aggregating experts' opinions, the maximum entropy method allows consideration of correlation, reliability and competence of experts in a scenario characterized by ambiguity. The maximum entropy approach is a method for aggregating a set of opinions into a single probability distribution with greater multiplicity that is more capable of realization in nature or more likely to occur. The maximum entropy principle was introduced by Jaynes [45,46] in physics as a generalization of the classical Principle of Insufficient Reason of Laplace. It is a general method to choose a probability distribution under uncertainty and elicits the most unbiased uniform distribution of all possible methods. Maximum entropy probability, which is a measure of conflict of evidence, is a measure of the diversification degree and a rational form of prudence. Some axiomatized aggregation formulas are based on maximum entropy interference, such as [47,48], which described a combination of two or more expert opinions in a single calibrated distribution, in which the calibration or competence measure reflects the quality of an expert's prediction and correlation of the dependence between and among experts' predictions. Competence measure is expressed by a real valued function (a monotone increasing function from 0 to 1) of the absolute distance or the quadratic absolute distance between the observed result and the predicted result. More complex appears to include pairwise interaction among experts because of sharing information sources, education backgrounds, theoretical dispositions, common training and experience. Levy and Delic [47], and Myung et al. [48] expressed statistical pairwise dependence (covariance) introducing correlation among experts' opinions in the constraints. With the exception of trivial cases of full and null dependence, the constrained problem of entropy maximization does not admit a closed form solution but requires application of numerical methods to find the solution such as [49], which