Erkenn (2011) 74:277–288
DOI 10.1007/s10670-010-9256-9

ORIGINAL PAPER

Peer Disagreement and Independence Preservation

Carl G. Wagner

Received: 27 April 2009 / Accepted: 22 September 2010 / Published online: 20 November 2010

Abstract It has often been recommended that the differing probability distributions of a group of experts should be reconciled in such a way as to preserve each instance of independence common to all of their distributions. When probability pooling is subject to a universal domain condition, along with state-wise aggregation, there are severe limitations on implementing this recommendation. In particular, when the individuals are epistemic peers whose probability assessments are to be accorded equal weight, universal preservation of independence is, with a few exceptions, impossible. Under more reasonable restrictions on pooling, however, there is a natural method of preserving the independence of any fixed finite family of countable partitions, and hence of any fixed finite family of discrete random variables.

1 Introduction

There has been a recent resurgence of interest among philosophers in the *epistemology of disagreement*, an inquiry that seeks to determine how two individuals ("you" and "I") should revise their beliefs in the face of disagreement when each regards the other as an *epistemic peer*.[1] Feldman (2007), Christensen (2007), and Elga (2007), have all argued that in such cases one should give equal weight to the opinion of a peer and to one's own opinion. As Kelly (2005) has

[1] There are a number of ways to explicate the notion of epistemic peer. We might each judge that we are equals on "intelligence, perspicacity, honesty, thoroughness, and other relevant epistemic virtues" (Gutting 1982, p. 83). Alternatively, we might each think that "conditional on our disagreeing, we are each equally likely to be mistaken." (Elga 2007). The precise explication of this notion is unimportant for our purposes here.

C. G. Wagner (✉)
Department of Mathematics, The University of Tennessee, Knoxville, TN 37996, USA
e-mail: wagner@math.utk.edu

observed, the natural framework for exploring this *equal weight view* is one in which beliefs are encoded as judgmental probabilities, and this is the framework adopted here.[2]

Suppose that you and I have each assessed a probability distribution over a countable set S of possible states of the world, and that my distribution P_1 differs from your distribution P_2. A natural way to implement the equal weight view here would be for each of us to revise our priors to their arithmetic mean $P = ½P_1 + ½P_2$.[3] But many individuals have found arithmetic averaging to be problematic as a method of pooling probabilities, citing, *inter alia*, the fact that such averaging may fail to preserve instances of independence common to P_1 and P_2. However, Genest and Wagner (1987) have shown that, under historically typical constraints on pooling, if $|S| \geq 5$, only dictatorial pooling ensures the universal preservation of independence. Applied to the epistemic peer problem, this theorem entails that, under the aforementioned constraints, insisting on universal independence preservation necessitates that I must either stand pat in all cases of disagreement with you, or in all cases adopt without modification your probability assessments. More generally, in the case of n individuals, each of whom regards each of the others as an epistemic peer, the only way to guarantee universal preservation of independence is for everyone to adopt the distribution of a single member of the group.

Do these results constitute a "conundrum" for the equal weight view (indeed, for any sort of weighted averaging as a method for reconciling differing probability assessments) as, for example, Shogenji (2007) has suggested? Our aim here is to examine this question in detail. First, we review results from the probability pooling literature that culminate in the Genest-Wagner dictatorship theorem. This review reveals that the problem lies with three unreasonable demands on pooling: *universal domain, state-wise aggregation,* and *universal preservation of ind*ependence. Under more reasonable restrictions on pooling, however, we show that there is a natural method of preserving the independence of any fixed finite family of countable partitions, and hence of any fixed finite family of discrete random variables. We conclude with some remarks about the apparent difficulty of devising pooling methods that preserve a fixed set of cases of *conditional* independence.

2 Probability Pooling and Independence Preservation: A Survey

In what follows, S denotes a countable set of possible states of the world, assumed to be mutually exclusive and exhaustive. A function $P: S \rightarrow [0,1]$ is a *probability distribution* on S if and only if $\sum_{s \in S} P(s) = 1$. Each probability distribution P

[2] If, for example, my doxastic options regarding propositions are limited to full belief, disbelief, and suspension of judgment, it is not even clear how to reconcile the full belief (or disbelief) of one peer with suspension of judgment on the part of another. As Thomas Kelly (2005) nicely puts it, how can the views of two epistemic peers, one atheist and the other agnostic, be reconciled in accord with the equal weight view?

[3] That is, for each state of the world $s \in S$, $P_3(s) = ½ P_1(s) + ½ P_2(s)$.

gives rise to a *probability measure* (which, abusing notation, we also denote by P) defined for each set $E \subseteq S$ by $P(E) := \sum_{s \in E} P(s)$.

If n is a positive integer, we call a sequence $(P_1,\ldots,P_n)$ of probability distributions on S a *profile*. A *pooling operator* T furnishes a method of reconciling the possible differences among the distributions $P_1,\ldots,P_n$ in the form of a single distribution $Q = T(P_1,\ldots,P_n)$ which, depending on the context, may furnish

1. a rough summary of the current probability distributions $P_1,\ldots,P_n$ of n individuals;
2. a compromise adopted by these individuals in order to complete an exercise in group decision making;
3. the probability distribution of a decision maker external to a group of n experts (who may or may not have assessed his or her own prior over S before consulting the group) upon being apprised of the probability distributions $P_1,\ldots,P_n$ of these experts; and, of particular interest here,
4. a "rational" consensus to which all individuals have revised (or ought rationally to revise) their initial probability distributions $P_1,\ldots,P_n$ after consideration of the expertise of each of their colleagues; in particular, the revision of the distributions of n individuals, each of whom regards each of the others as an epistemic peer.

There is an extensive literature on probability pooling (see the article of Genest and Zidek (1986) for a summary and appraisal of work done through the mid-1980s), which parallels in many respects the older and even more extensive literature on social welfare functions in the sense of Arrow (1951). Following the example of social welfare theory, pooling theories posit certain constraints on pooling, and then attempt to identify the pooling operators that satisfy those constraints. Typical constraints have included, for example:

2.1 Universal Domain (UD)

The domain of the pooling operator T consists of the set of *all logically possible profiles*, i.e., if Δ denotes the set of all probability distributions on S, and Δ^n its n-fold Cartesian product, then $T: \Delta^n \to \Delta$.

2.2 State-wise Aggregation[4] (SA)

For each $s \in S$, there exists a function $f_s: [0,1]^n \to [0,1]$ such that for all $(P_1,\ldots,P_n) \in \Delta^n$,

$$T(P_1,\ldots,P_n)(s) = f_s(P_1(s),\ldots,P_n(s)). \tag{2.1}$$

2.3 Zero Preservation (ZP)

For each $s \in S$ and all $(P_1,\ldots,P_n) \in \Delta^n$, if $P_1(s) = \cdots = P_n(s) = 0$ then $T(P_1,\ldots,P_n)(s) = 0$.

[4] State-wise aggregation is often termed *irrelevance of alternatives* or (confusingly) *independence*.

2.4 Universal Independence Preservation (UIP)

For all $(P_1, \ldots, P_n) \in \Delta^n$ and for all subsets E and F of S, if $P_i(E \cap F) = P_i(E)P_i(F)$ for $i = 1, \ldots, n$, then $T(P_1,\ldots,P_n)(E \cap F) = T(P_1,\ldots,P_n)(E)\, T(P_1, \ldots, P_n)(F)$.[5]

While SA alone allows for the functions f_s to vary with s, when $|S| \geq 3$, adding the restriction ZP completely removes this flexibility, entailing not only that there is a single function $f\colon [0,1]^n \rightarrow [0,1]$ such that, for all $s \in S$, $f_s = f$, but that f is a *weighted arithmetic mean.*

Theorem 2.1 (*Wagner* 1982) *If* $|S| \geq 3$, *a pooling operator T satisfies UD, SA, and ZP if and only if there exists a sequence* $(w_1,\ldots,w_n)$ *of nonnegative real numbers summing to 1 such that, for all* $s \in S$ *and all* $(P_1,\ldots,P_n) \in \Delta^n$, $T(P_1,\ldots,P_n)(s) = w_1P_1(s) + \cdots + w_nP_n(s)$.

Remark 2.2 If $|S| = 2$, every pooling operator satisfying UD trivially satisfies SA, and so there is a rich variety of such operators satisfying UD, SA, and ZP. See Lehrer and Wagner (1981, p. 110) for a characterization of the subclass of such operators for which $f_1 = f_2$. It is clear that weighted arithmetic pooling may fail to satisfy condition UIP. Indeed, only the most extreme forms of such pooling satisfy UIP.

Theorem 2.2 (*Lehrer and Wagner* 1983) *If* $|S| \geq 3$, *a pooling operator T satisfies UD, SA, ZP, and UIP if and only if it is dictatorial, i.e., if and only if there exists a* $d \in \{1,\ldots,n\}$ *such that for all* $(P_1,\ldots,P_n) \in \Delta^n, T(P_1,\ldots,P_n) = P_d$.

As shown in Wagner (1984), dropping condition ZP is of no help since for $|S| \geq 3$, pooling operators satisfying UD, IA and UIP must be dictatorial or imposed.

Remark 2.3 If $|S| = 2$, every pooling operator satisfying UD and ZP satisfies UIP, since if E and F are subsets of S independent with respect to P, then either p(E) or P(F) is equal to 0 or 1.

Theorem 2.2 is perhaps unsurprising, for SA is an extremely strong condition, requiring that the "consensual" probability assigned to each state s depends only (through the function f_s) on the probabilities assigned by individuals to that state, immediately precluding the sort of holistic approach that would seem necessary to preserve the independence of events which typically comprise multiple states.

In addition, SA requires that $\Sigma_s f_s(P_1(s),\ldots,P_n(s)) = 1$, without any normalization. Would allowing normalization accommodate UIP in non-dictatorial fashion? In what follows we restrict attention to probability distributions that assign a positive probability to each $s \in S$ in order to avoid consideration of minor variations on the principal result.[6] Accordingly, we subject our pooling operator T to the following axiomatic constraints:

[5] Advocates of universal preservation of event independence include Laddaga (1977), Laddaga and Loewer (1985), Schmitt (1985), and (implicitly) Barlow et al. (1985)

[6] See Wagner (1984), where dropping ZP while maintaining SA allows for externally imposed, as well as dictatorial pooling, depending on the fine details of how independence preservation is articulated.

2.5 Universal Positive Domain (UPD)

The domain of T consists of all logically possible profiles of positive probability distributions on S, i.e., if Δ^+ denotes the set of all positive probability distributions on S, and Δ^{+n} its n-fold Cartesian product, then T: $\Delta^{+n} \rightarrow \Delta^+$.

2.6 Normalized State-wise Aggregation (NSA)

For each $s \in S$, there exists a function g_s: $(0,1)^n \rightarrow (0,1)$ such that for all $(P_1, \ldots, P_n) \in \Delta^{+n}$,

$$\sum_{s \in S} g_s(P_1(s), \ldots, P_n(s)) < \infty,$$

and

$$T(P_1, \ldots, P_n)(s) = g_s(P_1(s), \ldots, P_n(s)) \,/\, \sum_{s \in S} g_s(P_1(s), \ldots, P_n(s)). \tag{2.2}$$

2.7 Universal Positive Independence Preservation (UPIP)

For all $(P_1, \ldots, P_n) \in \Delta^{+n}$ and for all subsets E and F of S, if $P_i(E \cap F) = P_i(E)P_i(F)$ for $i = 1, \ldots, n$, then $T(P_1, \ldots, P_n)(E \cap F) = T(P_1, \ldots, P_n)(E)\, T(P_1, \ldots, P_n)(F)$.

When $|S| = 3$, *any* pooling operator satisfying UPD preserves independence in a trivial way since events E and F cannot be independent with respect to $P \in \Delta^+$ unless one of E or F is S or the empty set. When $|S| = 4$, there is a rich variety of pooling operators satisfying UPD, NSA, and UPIP:

Theorem 2.3. (Abou-Zaid 1984; Sundberg and Wagner 1987) *Suppose that* $|S| = 4$ *and T satisfies UPD and NSA, where at least one of the functions* g_s *in formula* (2.2) *is Lebesgue measurable. Then T satisfies UPIP if and only if there exist arbitrary real constants* $a_1, \ldots, a_n$ *and* $b_1, \ldots, b_n$ *such that*

$$T(P_1, \ldots, P_n)(s) \propto \prod_{i=1}^{n} [P_i(s)]^{b_i} \exp\{a_i P_i(s)[1 - P_i(s)]\} \tag{2.3}$$

for all $(P_1, \ldots, P_n) \in \Delta^{+n}$ *and all* $s \in S$.

The pooling formulae arising from (2.3) include dictatorships ($a_i \equiv 0$, $b_i = \delta_{i,d}$, the Kronecker delta for fixed $d \in \{1, \ldots, n\}$); normalized weighted geometric means ($a_i \equiv 0$); and the method which imposes the uniform distribution for all profiles $P_1, \ldots, P_n$ ($a_i \equiv 0$, $b_i \equiv 0$).[7]

Alas, when $|S| \geq 5$, we again confront a dictatorship result:

Theorem 2.4. (Genest and Wagner 1987) *If* $|S| \geq 5$, *a pooling operator satisfies UPD, NSA and UPIP if and only if it is dictatorial.*

[7] The symbol $\propto$ is read "is proportional to," and indicates that the quantities on the right-hand side of formula (2.3) are normalized so as to sum to 1 over all $s \in S$.

As seen in formula (2.2), NSA allows the probabilities assigned by every individual to every state to enter into the determination of the consensual probability of each state. When |S| = 4, this added flexibility accommodates UPIP in a variety of interesting ways. As shown by Theorem 2.4, however, NSA is insufficiently flexible to accommodate UPIP when $|S| \geq 5$. This is perhaps not surprising, since, for a given state s, NSA relegates the probabilities assigned by individuals to states other than s to a very limited role, where they function only to ensure that consensual probabilities sum to 1 over all possible states.[8]

At first glance Theorem 2.4 may appear to devastating to the equal weight approach to resolving disagreement among epistemic peers, and, more generally, to resolving disagreement among any group of experts. As we argue in the next section, however, this theorem is the result of subjecting pooling to unreasonable universality conditions. Under more reasonable restrictions on pooling, however, there is a natural method of preserving the independence of any fixed finite family of countable partitions, and hence of any fixed finite family of discrete random variables.

3 Independence Preservation Under Reasonable Conditions

As noted earlier, probability pooling theory has been developed in analogy with axiomatic social choice theory, as formulated by Arrow (1951). In the process, certain universal domain conditions that are posited in the latter theory (motivated in part by a commitment to the ethical requirement that normative principles be universalizable) have been adopted more or less uncritically in the former theory. But universal domain conditions are considerably less compelling in the case of probability pooling. Of course, generality of application is a desirable feature of any decision method, but not at the expense of being deprived of a sensitive treatment of important classes of special cases. Moreover, conditions like UD and UPD more or less force one to employ pooling methods that satisfy SA or NSA. For how, by means of a (necessarily) finite set of instructions, is one to specify $T(P_1,\ldots,P_n)$ for *every logically possible profile* $(P_1,\ldots,P_n)$ without proceeding state-by-state? While one usually thinks of pooling axioms as normative conditions supported by rationality arguments, state-wise aggregation lacks this sort of normative provenance. Rather, SA and NSA are unavoidable responses to universal domain requirements with no independent rationale, and a straight jacket we should be happy to be free of, if only we could.

By contrast, advocates of universal independence preservation do regard this restriction as normative, but the rationale that they offer in support of imposing it typically amounts to citing special cases of agreed-upon independence, such as physical independence, where the desirability of preserving independence under pooling is uncontroversial. But citing such cases is inadequate as a justification for demanding preservation of *every single instance* of independence common to the distributions of a set of individuals. And no such justification could possibly be offered, for there are many instances of purely fortuitous independence. Consider

8 What is perhaps surprising is the fact that the modest extension of SA represented by NSA accommodates UIP in so many interesting ways when |S| = 4.

the case of a single toss of a die, where E = "die comes up even" and F = "die comes up a multiple of 3." You regard the die as fair, but I think that the die is weighted in such a way that the probabilities of 1, 5, and 6 are each 1/6, the probabilities of 2 and 4 are each 1/12, and the probability of 3 is 1/3. On each of our distributions, E and F are independent. But this independence is an incidental feature of our distributions, one which we may not even have noticed until it was called to our attention.[9] In this case, and in a multitude of similar cases that can be imagined, there is certainly no reason to care about preservation of independence. It should be emphasized, by the way, that our aim here is not to give a full account of the sorts of independence that warrant preservation (assuming such an account is even possible), but rather to describe a natural method of reconciling the possibly differing distributions of a group of individuals who agree on the independence of a certain family of partitions, *as well as on the desirability of preserving the independence under pooling*, a task to which we now turn our attention.[10]

We begin with a simple example. Suppose, that you and I regard each other as epistemic peers, and we agree that the outcomes of a sequence of two tosses of a coin are independent, but disagree about the probability of the coin landing heads, with your assessment of that probability being ¼ and mine being ½. What does it mean to reconcile our resulting distributions over the set S = {hh, ht, th, tt} under, say, an equal weighting scheme? The fact that the arithmetic mean of our distributions fails to preserve the independence, common to both our distributions, of, for example, the events E = "heads on the first toss" and F = "heads on the second toss," is simply a red herring. The sensible way to proceed here would be for each of us to adopt the value 3/8 = ½(¼ + ½) as the probability of heads, and then to exploit the independence of outcomes on different tosses to assess our common distribution over S.[11,12]

3.1 Independent Partitions

The above method is easily generalized, and indeed to the case of probability measures defined on a σ-algebra $\mathbf{A}$ of subsets (the "events") of an arbitrary set Ω. Suppose that P is such a probability measure and that E and F are subsets of S

[9] This example comes from Genest and Wagner (1987). See also Lehrer and Wagner (1983, p. 343).

[10] A presumably uncontroversial necessary condition for preserving independence is that such independence should play a role in the actual or potential construction of individuals' respective distributions.

[11] There are of course a number of other possibilities for averaging the probabilities ¼ and ½, consistent with the spirit of equal weighting. Indeed, given any strictly monotonic function α, we might revise our original probabilities that the coin lands heads to the (normalized) quasi-arithmetic mean $\alpha^{-1}(\frac{1}{2}[\alpha(\frac{1}{4}) + \alpha(\frac{1}{2})])/\sigma$, and our probabilities that the coin lands tails to $\alpha^{-1}(\frac{1}{2}[\alpha(\frac{3}{4}) + \alpha(\frac{1}{2})])/\sigma$, where $\sigma := \alpha^{-1}(\frac{1}{2}[\alpha(\frac{1}{4}) + \alpha(\frac{1}{2})]) + \alpha^{-1}(\frac{1}{2}[\alpha(\frac{3}{4}) + \alpha(\frac{1}{2})])$.

[12] This is also the sensible way to proceed in order to preserve other features common to our two distributions. Suppose, for example, that you and I agree that the random variable X has a Poisson distribution, but you think that $E(X) = \mu_1$ and I think that $E(X) = \mu_2$. We should clearly each revise our original distribution to a Poisson distribution with $E(X) = \frac{1}{2}(\mu_1 + \mu_2)$, or some other quasi-arithmetic mean of μ_1 and μ_2 (see note 11, *supra*). Here, by contrast, mindless state-by-state averaging of our original probabilities that X = k (k = 0,1,…) would produce a non-Poisson density function.

that are independent with respect to P (henceforth, "P-independent"). It is often assigned as an exercise in elementary probability texts to show that the independence of E and F entails (indeed, is equivalent to) the independence of E and F^c, the independence of E^c and F, and the independence of E^c and F^c. In other words, what is really at issue in judging the events E and F to be P-independent is the P-*independence of the partitions* $\mathbf{E} = \{E, E^c\}$ and $\mathbf{F} = \{F, F^c\}$, where, in general, the countable, **A**-measurable partitions $\mathbf{E^{(1)}} = \{E_{1,1}, E_{1,2}, \ldots\}$, $\mathbf{E^{(2)}} = \{E_{2,1}, E_{2,2}, \ldots\}, \ldots$, and $\mathbf{E^{(k)}} = \{E_{k,1}, E_{k,2}, \ldots\}$ of S are said to be P-independent if, for every sequence $\left(E_{1,j_1}, E_{2,j_2}, \ldots, E_{k,j_k}\right) \in \mathbf{E^{(1)}} \times \cdots \times \mathbf{E^{(k)}}$, $P\left(E_{1,j_1} \cap E_{2,j_2} \cap \cdots \cap E_{k,j_k}\right) = P\left(E_{1,j_1}\right)P\left(E_{2,j_2}\right) \cdots P\left(E_{k,j_k}\right)$. Note that what is usually termed the *total* P-*independence* of events $E_1, \ldots, E_k$ is equivalent to the P-independence of the partitions $\mathbf{E^{(1)}} = \{E_1, E_1^c\}, \ldots, \mathbf{E^{(k)}} = \{E_k, E_k^c\}$.

Suppose now that $P_1, \ldots, P_n$ are probability measures on **A** assessed by individuals of possibly differing expertise, reflected in a sequence $(w_1, \ldots, w_n)$ of nonnegative real weights summing to 1 (for epistemic peers, $w_1 = \cdots = w_n = 1/n$). Suppose that the partitions $\mathbf{E^{(1)}}, \ldots, \mathbf{E^{(k)}}$ are P_i-independent for $i = 1, \ldots, n$. We wish to construct a probability measure Q such that (1) Q incorporates the weights in a transparent way, and (2) the partitions $\mathbf{E^{(1)}}, \ldots, \mathbf{E^{(k)}}$ are Q-independent. The following procedure seems like a natural way to accomplish this:

(I.) Define the probability measure P by $P := w_1P_1 + \cdots + w_nP_n$.

(II.) For each *nonempty* event of the form $E_{1,j_1} \cap E_{2,j_2} \cap \cdots \cap E_{k,j_k}$ (the family of all such events constituting the so-called *cross partition of* $\mathbf{E^{(1)}}, \ldots, \mathbf{E^{(k)}}$, denoted $\mathbf{E^{(1)}} \otimes \cdots \otimes \mathbf{E^{(k)}}$) let

$$\mu(E_{1,j_1} \cap E_{2,j_2} \cap \cdots \cap E_{k,j_k}) := P(E_{1,j_1})P(E_{2,j_2}) \cdots P(E_{k,j_k}) \tag{3.1}$$

(III.) Let Q be the revision of P by Jeffrey conditionalization[13] on the (obviously countable) partition $\mathbf{E^{(1)}} \otimes \cdots \otimes \mathbf{E^{(k)}}$, with

$$Q\left(E_{1,j_1} \cap E_{2,j_2} \cap \cdots \cap E_{k,j_k}\right) = \mu\left(E_{1,j_1} \cap E_{2,j_2} \cap \cdots \cap E_{k,j_k}\right). \tag{3.2}$$

It is straightforward to verify the Q-independence of the partitions $\mathbf{E^{(1)}}, \ldots, \mathbf{E^{(k)}}$. Moreover, among all probability measures R on **A** satisfying $R\left(E_{1,j_1} \cap E_{2,j_2} \cap \cdots \cap E_{k,j_k}\right) = \mu\left(E_{1,j_1} \cap E_{2,j_2} \cap \cdots \cap E_{k,j_k}\right)$, and hence preserving the independence of the partitions $\mathbf{E^{(1)}}, \ldots, \mathbf{E^{(k)}}$, Q is nearest to P on several measures of closeness, including the variation distance, the Hellinger distance, and Kullback–Leibler divergence, in the latter two cases, uniquely so (see Diaconis and Zabell 1982).

Given the fixed finite family $\mathbb{E} = \{\mathbf{E^{(1)}}, \ldots, \mathbf{E^{(k)}}\}$ of countable, **A**-measurable partitions of Ω, let $\Pi_{\mathbb{E}}$ denote the set of all probability measures P on **A** for which the partitions $\mathbf{E^{(1)}}, \ldots, \mathbf{E^{(k)}}$ are P-independent, and $\Pi_{\mathbb{E}}^n$ its n-fold Cartesian product.

[13] If P is a probability measure on a sigma algebra **A** of subsets of any set Ω, $\mathbf{E} = \{E_i\}$ is a countable, **A**-measurable partition of Ω, and (μ_i) is a sequence of nonnegative real numbers summing to 1, then the probability measure Q, defined for all A in **A** by $Q(A) = \Sigma_i \mu_i P(A|E_i)$ (†) is said to come from P by *Jeffrey conditionalization* (or *probability kinematics*) on **E**. In the above formula it is assumed that if $P(E_i) = 0$, then $\mu_i = 0$, and that the term $\mu_i P(A|E_i) = 0$ in that case, notwithstanding the fact that $P(A|E_i)$ is undefined. See Jeffrey (1965). When Ω is equal to the countable set S, and $\mathbf{A} = 2^S$, (†) may be expressed in the form $Q(s) = \mu_i P(s)/P(E_i)$, where $s \in E_i$. (‡)

Setting $T(P_1,\ldots.P_n) = Q$, as constructed above in steps (I.)–(III.), defines a pooling operator $T\colon \Pi_E^n \rightarrow \Pi_E$. By restricting the domain of T to the (still infinite) set of profiles Π_E^n we are thus able to define a pooling operator that preserves the independence of the family $\mathtt{E}$. Moreover, $T(P_1,\ldots.P_n)(A) = 0$ whenever $P_1(A) = \cdots = P_n(A) = 0$. If, as in Sect. 2, Ω is equal to the countable set S and $\mathbf{A} = 2^S$, it is meaningful to ask whether T is constructible by means of normalized state-wise aggregation. In view of our earlier comments, it will come as no surprise to learn that this is not in general the case.[14]

3.2 Independent Random Variables

The above approach applies in particular to the case of independent random variables. As above, let $(\Omega, \mathbf{A}, P)$ be a probability space. Suppose that $\Theta_1 = \{\theta_{1,1}, \theta_{1,2},\ldots\}$, $\Theta_2 = \{\theta_{2,1}, \theta_{2,2},\ldots\},\ldots$, and $\Theta_k = \{\theta_{k,1}, \theta_{k,2},\ldots\}$ are countable sets, and that the mappings $X_i : \Omega \rightarrow \Theta_i$ are $\mathbf{A}$- measurable for $i = 1,\ldots,k$ (i.e., for each i and j, the set $E_{i,j} := \{\omega \in \Omega\colon X_i(\omega) = \theta_{i,j}\} \in \mathbf{A}$). Recall that the discrete random variables $X_1,X_2,\ldots,X_k$ are said to be P-*independent* if, for every $\left(\theta_{1,j_1}, \theta_{2,j_2},\ldots, \theta_{k,j_k}\right) \in \Theta_1 \times \Theta_2 \times \cdots \times \Theta_k$,

$$\begin{aligned} &P\left(X_1 = \theta_{1,j_1},\ X_2 = \theta_{2,j_2},\ldots,\ X_k = \theta_{k,j_k}\right) \\ &\quad = P\left(X_1 = \theta_{1,j_1}\right)P\left(X_2 = \theta_{2,j_2}\right)\cdots P\left(X_k = \theta_{k,j_k}\right) \end{aligned} \tag{3.3}$$

i.e., if

$$P\left(E_{1,j_1} \cap E_{2,j_2} \cdots \cap E_{k,j_k}\right) = P\left(E_{1,j_1}\right)P\left(E_{2,j_2}\right)\cdots P\left(E_{k,j_k}\right) \tag{3.4}$$

So the P-independence of $X_1, X_2,\ldots,X_k$ is equivalent to the P-independence of the partitions $\mathbf{E}^{(1)} = \{E_{1,1}, E_{1,2}, \ldots\}$, $\mathbf{E}^{(2)} = \{E_{2,1}, E_{2,2}, \ldots\}$, …, and $\mathbf{E}^{(k)} = \{E_{k,1}, E_{k,2}, \ldots\}$. Thus if $X_1, X_2,\ldots, X_k$ are P_i -independent for $i = 1,\ldots,k$, we can apply the

[14] Let $(P_1,\ldots,P_n)$ and $(P_1^*,\ldots,P_n^*)$ be profiles of probability distributions on S, and $s_1,s_2 \in S$. Suppose that $P_i(s_1) = P_i^*(s_1)$ and $P_i(s_2) = P_i^*(s_2)$ for $i = 1,\ldots,n$, that $T(P_1,\ldots,P_n) = Q$, and $T(P_1^*,\ldots,P_n^*) = Q^*$. Note that if T is constructible on the basis of normalized state-wise aggregation, then by formula (2.2) it must be the case that $Q(s_1)/Q(s_2) = Q^*(s_1)/Q^*(s_2)$. Now let $S = \{a,b,c,d,e,f\}$, $A = \{a,b,c,d\}$, and $B = \{a,b,e\}$, and consider the following scenarios, where $P = ½(P_1 + P_2)$ and $P^* = ½(P_1^* + P_2^*)$:

	a	b	c	d	e	f
P_1	.06	.06	.40	.08	.08	.32
P_2	.20	.20	.04	.06	.40	.10
P	.13	.13	.22	.07	.24	.21
P_1^*	.06	.08	.40	.16	.06	.24
P_2^*	.20	.12	.04	.04	.48	.12
P*	.13	.10	.22	.10	.27	.18

Note that the partitions $\{A, A^c\}$ and $\{B, B^c\}$ are P_i- and P_i^*-independent for i = 1, 2 (though, as expected, neither P- nor P*-independent). Suppose that Q comes from P, and Q* from P*, by Jeffrey conditioning on the partition $\{AB, AB^c, A^cB, A^cB^c\}$, with $Q(AB) = P(A)P(B)$, $Q(AB^c) = P(A)P(B^c)$, $Q(A^cB) = P(A^c)P(B)$, $Q(A^cB^c) = P(A^c)P(B^c)$, $Q^*(AB) = P^*(A)P^*(B)$, $Q^*(AB^c) = P^*(A)P^*(B^c)$, $Q^*(A^cB) = P^*(A^c)P^*(B)$, and $Q^*(A^cB^c) = = P^*(A^c)P^*(B^c)$, whence $T(P_1, P_2) = Q$ and $T(P_1^*, P_2^*) = Q^*$. An easy computation shows that $Q(a)/Q(c) = 29/44 \neq 208/253 = Q^*(a)/Q^*(c)$. So by the previous remark, T is not constructible on the basis of normalized state-wise aggregation.

method described above in Sect. 3.1 to construct a probability measure Q on $\mathbf{A}$ that preserves this independence. From Q we can of course determine the joint density function q of the random variables $X_1, X_2, \ldots, X_k$. If $\left(\theta_{1,j_1}, \theta_{2,j_2}, \ldots, \theta_{k,j_k}\right) \in \Theta_1 \times \Theta_2 \times \cdots \times \Theta_k$, then

$$\begin{aligned} q\left(\theta_{1,j_1}, \theta_{2,j_2}, \ldots, \theta_{k,j_k}\right) &:= Q\left(X_1 = \theta_{1,j_1}, X_2 = \theta_{2,j_2}, \ldots, X_k = \theta_{k,j_k}\right) \\ &= Q\left(E_{1,j_1}\right) Q\left(E_{2,j_2}\right) \cdots Q\left(E_{k,j_k}\right) \end{aligned} \tag{3.5}$$

Remark 1 In the above analysis, we have followed the practice of probability theorists in construing random variables as $\mathbf{A}$-measurable mappings on a measurable space (Ω, $\mathbf{A}$), and so our focus has naturally been on pooling various probability measures on $\mathbf{A}$. A full determination of the consensual measure Q will in general require carrying out all three of the steps described in Sect. 3.1, including the final application of Jeffrey conditionalization. There are many reasons for wanting a *full* determination of Q (rather than just of the values of Q on events in the cross-partition $\mathbf{E}^{(1)} \otimes \cdots \otimes \mathbf{E}^{(k)}$, which only requires carrying out the first two of those steps[15]). There may, for example, be random variables other than $X_1, X_2, \ldots, X_k$ defined on Ω whose consensual densities or expectations we would like to determine. We may also be interested in consensual values of certain conditional probabilities, including those that require the values of Q on a sub-algebra of $\mathbf{A}$.

Remark 2 If our sole interest is in determining the consensual joint density function q of $X_1, X_2, \ldots, X_k$, then, as is clear from formula (3.5), we need only determine Q on the events in $\mathbf{E}^{(1)} \otimes \cdots \otimes \mathbf{E}^{(k)}$. By (3.1) and (3.2) this only involves the values taken by $P_1, \ldots, P_n$ on the events in the family $\mathbf{E}^{(1)} \cup \cdots \cup \mathbf{E}^{(k)}$, which is to say the values of the marginal densities of each of the variables $X_1, X_2, \ldots, X_k$, as assessed by each of the n individuals. Then formula (3.5) takes the form

$$q\left(\theta_{1,j_1}, \theta_{2,j_2}, \ldots, \theta_{k,j_k}\right) = q_1\left(\theta_{1,j_1}\right) q_2\left(\theta_{2,j_2}\right) \cdots q_k\left(\theta_{k,j_k}\right) \tag{3.6}$$

where

$$q_i = w_1 p_{1,i} + w_2 p_{2,i} + \cdots + w_n p_{n,i}, i = 1, \ldots, k, \tag{3.7}$$

and

$$p_{r,i}\left(\theta_{i,j_i}\right) = P_r\left(X_i = \theta_{i,j_i}\right), \ r = 1, \ldots, n. \tag{3.8}$$

Remark 3 In many applications, random variables are *sui generis*, i.e., defined by direct specification of their density functions, rather than as mappings on a measurable space.[16] Then (3.6) and (3.7) furnish a method of pooling the density

[15] If it should happen that all of the events in this cross partition are atomic, then of course Q will be fully determined after steps (I.) and (II.). Applying Jeffrey conditioning in this situation simply amounts to determining Q(A) by summing the probabilities of all of the atomic events contained in A.

[16] Discrete random variables $X_1, X_2, \ldots, X_k$ defined directly by their joint density function q may of course be conceptualized as mappings on a measurable space by taking $\Omega = \Theta_1 \times \Theta_2 \times \ldots \times \Theta_k$, $\mathbf{A} = 2^{\Omega}$,
$P(A) = \Sigma(\theta_1, \theta_2, \ldots, \theta_k)$, taken over all $(\theta_1, \theta_2, \ldots, \theta_k) \in A$, and
$X_i(\theta_1, \theta_2, \ldots, \theta_k) = \theta_i$.

functions of n individuals who agree on the independence of the random variables $X_1, X_2, \ldots, X_k$.[17]

Remark 4 The notion of conditional independence plays an important role in probability theory, especially in the construction of Bayesian networks, where conditional independence is exploited to avoid storing the entire joint density function of a family of random variables (Pearl 1988). Suppose then that n individuals agree on a conditional independence structure associated with the variables $X_1, X_2, \ldots, X_k$, but disagree on the values of certain conditional densities. It would clearly be desirable to devise a pooling method that *exactly preserves* the agreed-upon conditional independence structure, but this appears to be a very difficult problem. It might be tempting to think that a simple variant of (3.6) and (3.7) will do the trick here, namely, pooling the relevant conditional densities instead of the marginal densities. But the following example shows that this will not suffice: Suppose that you and I are epistemic peers, and we agree that the ternary random variables $X_1, X_2, \ldots, X_k$ constitute a Markov chain. But our transition matrices M and M* differ, with

$$M = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 5/6 & 1/12 & 1/12 \\ 1/3 & 1/4 & 5/12 \end{bmatrix} \quad \text{and} \quad M* = \begin{bmatrix} 1/3 & 1/6 & 1/2 \\ 1/6 & 1/4 & 7/12 \\ 2/3 & 1/12 & 1/4 \end{bmatrix}. \tag{3.9}$$

Then

$$½(M + M*) = \begin{bmatrix} 1/2 & 1/6 & 1/3 \\ 1/2 & 1/6 & 1/3 \\ 1/2 & 1/6 & 1/3 \end{bmatrix}. \tag{3.10}$$

converting what we agreed was a Markov chain into a sequence of independent random variables, with $X_2, \ldots, X_k$ being identically distributed.[18]

References

Abou-Zaid, S. H. (1984). *On functional equations and measures of information*. M.Phil. Thesis, University of Waterloo.

Arrow, K. (1951). *Social choice and individual values*. Cowles Commission Monograph 12. New York: Wiley (second edition, 1963).

Barlow, R., Mensing, R. & Smiriga, N. (1985). Combination of experts' opinions based on decision theory. In *Proceedings of the International conference on reliability theory and quality control*, University of Missouri.

Bradley, R., Dietrich, F., & List, C. (2006). *Aggregating causal judgements*. Choice Group working papers, Vol. 2, London School of Economics.

[17] In this special case, our proposal coincides with one appearing in an unpublished paper of Bradley et al. (2006).

[18] The consensual marginal density of X_1, calculated as the average of our respective marginal densities of X_1, may of course differ from the consensual marginal densities of $X_2, \ldots, X_k$.

Christensen, D. (2007). Epistemology of disagreement: The good news. *Philosophical Review, 116*, 187–217.

Diaconis, P., & Zabell, S. (1982). Updating subjective probability. *Journal of the American Statistical Association, 77*, 822–829.

Elga, A. (2007). Reflection and disagreement. *Nous, 41*, 478–502.

Feldman, R. (2007). Reasonable religious disagreements. In L. Anthony (Ed.),*Philosophers without gods* (pp. 194–214). Oxford: Oxford University Press.

Genest, C., & Wagner, C. (1987). Further evidence against independence preservation in expert judgement synthesis. *Aequationes Mathematicae, 32*, 74–86.

Genest, C., & Zidek, J. (1986). Combining probability distributions: a critique and annotated bibliography. *Statistical Science, 1*, 114–148.

Gutting, G. (1982). *Religious belief and religious skepticism*. Notre Dame: University of Notre Dame Press.

Jeffrey, R. (1965). *The logic of decision*. New York: McGraw-Hill.

Kelly, T. (2005). Peer disagreement and higher order evidence. In R. Feldman & T. Warfield (Eds.), *Disagreement.* Oxford: Oxford University Press.

Laddaga, R. (1977). Lehrer and the consensus proposal. *Synthese, 36*, 473–477.

Laddaga, R., & Loewer, B. (1985). Destroying the consensus. *Synthese, 62*, 79–96.

Lehrer, K., & Wagner, C. (1981). *Rational consensus in science and society*. Philosophical studies series in philosophy 24. Dordrecht: D. Reidel Publishing Company.

Lehrer, K., & Wagner, C. (1983). Probability amalgamation and the independence issue: A reply to Laddaga. *Synthese, 55*, 339–346.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Mateo: Morgan Kaufmann Publishers.

Schmitt, F. (1985). Consensus, respect, and weighted averaging. *Synthese* (special issue—*Rational Consensus*), *62.1*, 25–46.

Shogenji, T. (2007). *A conundrum in Bayesian epistemology*. Paper presented at the Formal Epistemology Workshop, Carnegie Mellon University.

Sundberg, C., & Wagner, C. (1987). A functional equation arising in multi-agent statistical decision theory. *Aequationes Mathematicae, 32*, 32–37.

Wagner, C. (1982). Allocation, Lehrer models, and the consensus of probabilities. *Theory and Decision, 14*, 207–220.

Wagner, C. (1984). Aggregating subjective probabilities: some limitative theorems. *Notre Dame Journal of Formal Logic, 25*, 233–240.

Wagner, C. (1985). On the formal properties of weighted averaging as a method of aggregation. *Synthese, 62*, 97–108.