What is the 'Equal Weight View'?

David Jehle
dj55@cornell.edu

&

Branden Fitelson
branden@fitelson.org

- We'll be talking about a pair of (Bayesian) agents S_1 and S_2 who are **epistemic peers**, regarding some propositions P.
- We'll call such p's in P **peer-propositions** (for S_1 and S_2).
- We will not assume any general "laws" of the **logic of peer-proposition-hood** (but, we will presuppose that $\ulcorner p \,\&\, q \urcorner$ is a peer-proposition $\not\Rightarrow$ $\ulcorner p \urcorner$ is a peer-proposition).
- We'll be exploring various **peer-update rules** (PURs) for such pairs of Bayesian agents. These PURs are *rules for updating credences*, upon learning (***exactly!***) the information $[D(P)]$ that S_1 and S_2 disagree on some set of peer propositions P.
- The idea behind **Equal Weight Rules** (EWRs) is that, when S_1 and S_2 learn $D(P)$, they should adopt consensus credences on P that "roughly split the difference" [1] in credence on P.
- We will begin with the simplest EWR. Then, using various constraints from the literature on Bayesian judgment aggregation [5], we will develop more sophisticated EWRs.

- Let $\Pr_i^0(p)$ be the credence S_i assigns to p at t_0, and let $\Pr_i^1(p)$ be the credence S_i assigns to p at t_1, where $t_1 > t_0$.
- We will assume that S_1 and S_2 learn ***exactly*** (important caveat for *any* update rule!) the following between t_0 and t_1:

  $[D(P)]$ For each $p \in P$, $\Pr_1^0(p) \neq \Pr_2^0(p)$.

  [Note: we *do* mean to assume here that $D(P)$ includes the *numerical values* of the $\Pr_i^0(p)$, but *that* information is only relevant to our PUR's *to the extent that it informs about the disagreement* ***qua disagreement***. If the $\Pr_i^0(p)$ are *also* relevant (in the context) to the determination of $\Pr_i^1(p)$ for *other* reasons, our PUR's will ignore these other relevancies.]
- A PUR will just be a rule, which, for each $p \in P$, prescribes how the credences of S_1 and S_2 should be updated, so as to properly respond to credal disagreements $D(P)$.
- For simplicity, we'll assume that S_1 and S_2 share a sentential language $\mathcal{L}$ with just two atomic sentences A and B.

- Our first PUR is **straight averaging** ("split the difference"):

$$\text{(SA)} \qquad \Pr_{SA}^1(p) = \frac{\Pr_1^0(p) + \Pr_2^0(p)}{2}$$

- This naïve, *exact* "split the difference" PUR (SA) may sound appealing, but it is *under-specified*, as it stands. Example:

A	B	$\Pr_1^0(\cdot)$	$\Pr_2^0(\cdot)$	$\Pr_{SA}^1(\cdot)$
$\top$	$\top$	**0.1**	**0.55**	**0.325**
$\top$	$\bot$	**0.2**	**0.25**	**0.225**
$\bot$	$\top$	0.3	0.15	??
$\bot$	$\bot$	0.4	0.05	??

- Here, $A \,\&\, B$ and $A \,\&\, {\sim}B$, are peer-props (bold), but ${\sim}A \,\&\, B$ and ${\sim}A \,\&\, {\sim}B$ are not. So, (SA) prescribes new credences for $A \,\&\, B$ and $A \,\&\, {\sim}B$, but not for ${\sim}A \,\&\, B$ and ${\sim}A \,\&\, {\sim}B$.

☞ *Neither S_1 nor S_2 can keep their old credences in both ${\sim}A \,\&\, B$ and ${\sim}A \,\&\, {\sim}B$ — on pain of synchronic incoherence*!

Probabilism (P). $\Pr_1^1(\cdot)$ and $\Pr_2^1(\cdot)$ should be *probabilities*.

- (SA) must be revised, so as to tell agents what to do when (**P**) + (SA) *forces* changes to credences on *non*-peer p's.
- **Informal Idea**: revise (SA) to (SA**MC**), which recommends that (in such cases) each agent makes **minimal** (forced) **changes** to their credences on non-peer propositions.

(SAMC) Upon learning (exactly) $D(P)$, S_1 and S_2 should (*i*) obey (SA) for peer-propositions P, and (*ii*) if (**P**) should force additional revisions, then each agent should revise their credences by moving to *a closest probability function* compatible with both (SA) and (**P**).

- In our example above, (SAMC) entails these *unique* $\Pr_i^1(\cdot)$'s (assuming a *Euclidean distance metric* [3] on credence f's):

A	B	$\Pr_1^0(\cdot)$	$\Pr_2^0(\cdot)$	$\Pr_1^1(\cdot)$	$\Pr_2^1(\cdot)$
⊤	⊤	**0.1**	**0.55**	**0.325**	**0.325**
⊤	⊥	**0.2**	**0.25**	**0.225**	**0.225**
⊥	⊤	0.3	0.15	0.175	0.275
⊥	⊥	0.4	0.05	0.275	0.175

- (SAMC) ensures that the updates prescribed by (SA) will obey *probabilism* (**P**). What about **conditionalization**?
- We would like our PUR to *commute with conditionalization*.
- Let $\Pr_i^{0+p}(q) = \Pr_i^0(q \mid p)$ be the degree of belief an agent i should assign to q, upon learning (exactly) p, after $t = 0$.
- And, let $\overline{\Pr_i^0}(\cdot)$ be what our (PUR) prescribes for the agent i's credence function, upon learning (exactly) $D(P)$, after $t = 0$.

Conditionalization (C). Suppose p, q, and $p \,\&\, q$ are peer-propositions for S_1 and S_2 (at both t_0 and t_1), and also that q remains a peer-proposition for S_1 and S_2 (at t_0) on the supposition that p is true. Then, we should have:

$$\overline{\Pr_i^{0+p}}(q) = \overline{\Pr_i^0}(q \mid p)$$

- $\overline{\Pr_i^{0+p}}(q)$ *conditionalizes on p first* and *then* peer-updates.
- $\overline{\Pr_i^0}(q \mid p)$ *peer-updates first*, and *then* conditionalizes on p.

☞ *The order in which we conditionalize/PU shouldn't matter.*

- Suppose A, B, and $A \,\&\, B$ are peer-propositions for S_1 and S_2 (at both t_0 and t_1), and that B remains a peer-proposition (at t_0) on the supposition that A is true. Then, our example entails a *unique* (SAMC)-distribution for both agents at t_1:

A	B	$\Pr_1^0(\cdot)$	$\Pr_2^0(\cdot)$	$\Pr_i^1(\cdot)$
⊤	⊤	0.1	0.55	0.325
⊤	⊥	0.2	0.25	0.225
⊥	⊤	0.3	0.15	0.225
⊥	⊥	0.4	0.05	0.225

- First, let's calculate the $\Pr_i^{0+A}(B)$ values for the two agents:

$$\Pr_1^{0+A}(B) = \Pr_1^0(B \mid A) = \frac{\Pr_1^0(A \,\&\, B)}{\Pr_1^0(A)} = \frac{0.1}{0.3} = 0.3333$$

$$\Pr_2^{0+A}(B) = \Pr_2^0(B \mid A) = \frac{\Pr_2^0(A \,\&\, B)}{\Pr_2^0(A)} = \frac{0.55}{0.8} = 0.6875$$

- Second, apply (SAMC) to these (peer) $\Pr_i^{0+A}(B)$ values:

$$\overline{\Pr_i^{0+A}}(B) = \frac{0.3333 + 0.6875}{2} = .5105$$

- Finally, let's calculate the value of $\overline{\Pr_i^0}(B \mid A)$. This can be done uniquely here, since — in this example — (SAMC) entails a unique (SAMC)-distribution for both agents, at t_1:

$$\overline{\Pr_i^0}(B \mid A) = \frac{\overline{\Pr_i^0}(B \,\&\, A)}{\overline{\Pr_i^0}(A)} = \frac{\Pr_i^1(B \,\&\, A)}{\Pr_i^1(A)} = \frac{0.325}{0.55} = 0.5909$$

- This is a counterexample to (**C**) for *any* PUR that *exactly* "splits the difference" on P — including (SA) *and* (SAMC).
- Moreover, this is also an (SAMC)-counterexample to:

Preservation of Conditional (In)dependencies (PCI): $\Pr_1^1(\cdot)$ and $\Pr_2^1(\cdot)$ should neither reverse initially agreed-upon assessments of conditional (in)dependence, nor force new disagreements about relations of conditional (in)dependence, among the set of peer-propositions P.

- To see why this is an (SAMC)-counterexample to (**PCI**), note:

$$\Pr_1^0(B) = 0.4 > \Pr_1^0(B \mid A) = 0.3333, \text{ and}$$
$$\Pr_2^0(B) = 0.7 > \Pr_2^0(B \mid A) = 0.5909, \text{ but}$$
$$\Pr_i^1(B) = 0.55 < \Pr_i^1(B \mid A) = 0.5909.$$

- At t_0, S_1 and S_2 agree that A and B are *negatively* correlated. But, at t_1, S_1 and S_2 both reverse their assessments, and come to agree that A and B are *positively* correlated.
- In the literature on Bayesian consensus, (**P**) and (**C**) are usually taken as basic desiderata for any adequate PUR [9].
- (**PCI**), on the other hand, is far more controversial [6, 8].
- We won't take a stand on these controversies here.
- Rather, we'd like to explore a natural weakening of (SAMC), which is capable of satisfying (**P**) and (**C**), as well as (**PCI**) and many other possible sets of constraints besides.
- Recall our talk at the beginning of "roughly" splitting the difference. That's going to be the guiding informal idea…

Approximate **SAMC** (ASAMC): Upon learning (exactly) $D(P)$:

$$\Pr_1^1(p) \approx \Pr_2^1(p) \approx \frac{\Pr_1^0(p) + \Pr_2^0(p)}{2},$$

where $\Pr_i^1(p)$ is *strictly between* $\Pr_1^0(p)$ and $\Pr_2^0(p)$.

And, where the update is done so as to satisfy (**P**) *and* (**C**). If changes to non-peer credences are forced in order to ensure (**P**) and (**C**), then the other changes should be made so as to minimize the distance of $\Pr_1^1(\cdot)$, $\Pr_2^1(\cdot)$ from $\Pr_1^0(\cdot)$, $\Pr_2^0(\cdot)$, while maintaining (**P**) and (**C**). Finally, if the satisfaction of further constraints C (*e.g.*, **PCI**) is desired, then these should be added to the constraint satisfaction problem (in both the initial and "minimal change" steps), so that the update process respects these additional constraints as well.

- Note: $x \approx y \overset{\text{def}}{=} |x - y| < \epsilon$. Other definitions could be used. We assume a *single* ϵ for *all* $p \in P$. This could be relaxed.
- Note: we require that $\Pr_i^1(p)$ be *strictly between* $\Pr_1^0(p)$ and $\Pr_2^0(p)$, so as to *rule-out* a **dictatorial** update, which adopts *one* of the agent's credence in p as the "consensus" value.

- As stated, (ASAMC) is *ambiguous* between two readings:
  1. $\Pr_1^1(p)$ *must equal* $\Pr_2^1(p)$. Here: (**A**) *exact credal agreement* is reached on each peer-proposition. But, on this reading, the consensus value $\Pr_c^1(p)$ will be closer to one of the initial credences $\Pr_i^0(p)$ than it is to the other. This violates a condition we call "equal credence Δ's" (**ECΔ**). One might maintain that (**ECΔ**) is central to any "equal weight" view.
  2. $\Pr_1^1(p)$ and $\Pr_2^1(p)$ *may remain unequal*. Here, exact consensus need *not* be reached on all peer-propositions [that is, (**A**) may be violated]. But, this reading can be further precisified, so as to ensure that each updated credence $\Pr_i^1(p)$ is *equally far* from the halfway point between the initial credences $\Pr_i^0(p)$ [(**ECΔ**)]. So, this reading may be closer, in spirit, to the "equal weight" idea.
- We won't take a stand here on which of these precisifications of (ASAMC) is preferable, as an (EWR).
- Rather, we will instead discuss some interesting formal properties that are common to both readings of (ASAMC).

- On either reading of (ASAMC), we have the following:

  Theorem. (ASAMC) is compatible with (**P**), (**C**), and (**PCI**). That is, we can *always* find consensus credence functions that obey (ASAMC) as well as (**P**), (**C**), and (**PCI**). [The only question will be *how large a value of* ϵ will be required.]
- Sometimes, non-trivially "large" values of ϵ are required in order to yield (ASAMC)-updates satisfying (**P**), (**C**) & (**PCI**).
- *E.g.*, in our last example above (Table 3 on slide 7), there do exist (ASAMC) updates satisfying (**P**), (**C**), and (**PCI**), but they all require a threshold value of $\epsilon > \frac{1}{16}$, for all $p \in P$.
- In the MATHEMATICA notebook for this paper [4] (which has all technical results), we have examples in which $\epsilon > \frac{1}{10}$ is forced by (ASAMC) in order to ensure (**P**), (**C**) & (**PCI**).
- We will not take a stand here on how large ϵ should be allowed to get, in various contexts (or whether different p's in P should require different ϵ's, depending on context, *etc.*).
- Our aim here is merely to survey the landscape of EWRs.

- It is worth noting that *all* of the PUR's we discuss satisfy the following constraint, which is *strictly weaker* than (**A**):

  Unanimity (**U**): $\Pr_1^1(\cdot)$ and $\Pr_2^1(\cdot)$ should not force new point-wise disagreements about credence values concerning peer-propositions on which S_1 and S_2 already agree (at t_0).
- (**U**) is perhaps *the most basic* of all desiderata for PUR's.
- Another constraint that people have often discussed in the historical literature on judgment aggregation [7] is:

  Irrelevance of Alternatives (**IA**): $\Pr_1^1(p)$ and $\Pr_2^1(p)$ should each be *functions* of $\Pr_1^0(p)$ and $\Pr_2^0(p)$. That is, for each peer-proposition p, $\Pr_1^1(p) = f_1[\Pr_1^0(p), \Pr_2^0(p)]$, and $\Pr_2^1(p) = f_2[\Pr_1^0(p), \Pr_2^0(p)]$, for some *functions* f_1 and f_2.
- While (**IA**) may make some sense in a full belief/inference context (as in traditional judgment aggregation [7]), it makes much less sense in a probabilistic/Bayesian context.
- *E.g.*, we conjecture that any remotely plausible EWR/PUR which satisfies (**IA**) must fail to satisfy either (**P**) or (**C**) [2].

- Note that of all the conditions usually discussed in the literature (see below), (ASAMC) fails to satisfy *only* (**IA**).
- This is because "approximate splittings" can be achieved *in multiple ways, for the same pair of initial credence values.*
- As such, there will (in general) be no *function*(s) of said credence values that yields the (ASAMC)-updated values.
- One could try to state (ASAMC) as function of the initial credences *plus* ϵ. But, since ϵ may itself vary with context, this could still (strictly speaking) lead to violations of (**IA**).

	Can Rule (Always) Satisfy Condition?						
Rule	(**P**)	(**C**)	(**U**)	(**A**)	(**EC∆**)	(**IA**)	(**PCI**)
(SA)	No*	No	Yes	Yes	Yes	Yes	No
(SAMC)	Yes	No	Yes	Yes	Yes	Yes	No
(ASAMC$_1$)	Yes	Yes	Yes	Yes	No	No	Yes
(ASAMC$_2$)	Yes	Yes	Yes	No	Yes	No	Yes

☞ By "going approximate", one can avoid all of the probative triviality results in the Bayesian literature on consensus.

[1] Christensen, D. (2007) "Epistemology of Disagreement: The Good News", *Philosophical Review* 119: 187-217.

[2] Dalkey, N. (1972). *An impossibility theorem for group probability functions.* P–4862, The Rand Corporation ⟨http://www.rc.rand.org/pubs/papers/P4862/⟩.

[3] Diaconis, P. and Zabell, S. (1982) "Updating Subjective Probability", *Journal of the American Statistical Association* 77: 822-830.

[4] Jehle, D. and Fitelson, B. (2009) "What is the 'Equal Weight View'?", manuscript ⟨http://fitelson.org/ew_episteme.pdf⟩. The Mathematica notebook for this paper is at ⟨http://fitelson.org/ew.nb⟩.

[5] Lehrer, K. and Wagner, C. (1981) *Rational Consensus in Science and Society: A Philosophical and Mathematical Study*. Dordrecht-Boston: Reidel.

[6] ______ (1983) "Probability Amalgamation and the Independence Issue: A Reply to Laddaga", *Synthese* 55: 339-346.

[7] List, C. and Puppe, C. (2009) "Judgement aggregation: a survey," in Anand, P., Pattaniak, P. and Puppe, C. (*eds.*) *Oxford handbook of rational and social choice.*

[8] Loewer, B. and Laddaga R. (1985) "Destroying the Consensus", *Synthese* 62: 79-95.

[9] Shogenji, T. (2007) "A Conundrum in Bayesian Epistemology of Disagreement", presented at FEW 2007 ⟨http://www.fitelson.org/few/few_07/shogenji.pdf⟩.