Additivity and the opinion-set

Catrin Campbell-Moore*

April 27, 2020

WARNING: Updates coming soon!

1 Agenda and probabilities

Definition 1.1 (Worlds, propositions and Boolean algebras)**.** There are different possible setups:

1. • Start with a collection, W.
   - Propositions, $A \in \mathcal{P}$ are conceived of as $A \subseteq W$.
   - $\mathcal{P}$ forms a Boolean algebra if
     - $W \in \mathcal{P}$,
     - $\varnothing \in \mathcal{P}$,
     - $A \in \mathcal{P} \implies W \setminus A \in \mathcal{P}$,
     - $A, B \in \mathcal{P} \implies A \cap B \in \mathcal{P}$,
   - To link to the sentential framework, we define $w(A) = \mathsf{t}$ if $w \in A$, and $w(A) = \mathsf{f}$ if $w \notin A$.
2. • We start with $\mathcal{P}$ a collection of sentences in a language.
   - $\mathcal{P}$ forms a Boolean algebra if
     - $\top \in \mathcal{P}$,
     - $\bot \in \mathcal{P}$,
     - $A \in \mathcal{P} \implies \neg A \in \mathcal{P}$,
     - $A, B \in \mathcal{P} \implies A \wedge B \in \mathcal{P}$.
   - W is collection of classically consistent functions $w : \mathcal{P} \to \{\mathsf{t}, \mathsf{f}\}$ (it might be a subset of the collection of all classical logic models).
3. [Stating the previous one slightly more generally]
   - Start with $\mathcal{P}$ simply as a non-empty set.
   - W is some collection of functions $w : \mathcal{P} \to \{\mathsf{t}, \mathsf{f}\}$.

*

- $\mathcal{P}$ forms a Boolean algebra if
  - There is some $A \in \mathcal{P}$ such that $w(A) = \mathsf{t}$ for all $w \in W$. Let $\top$ denote some such A.
  - There is some $A \in \mathcal{P}$ such that $w(A) = \mathsf{f}$ for all $w \in W$. Let $\bot$ denote some such A.
  - If $A \in \mathcal{P}$ then there is some $B \in \mathcal{P}$ such that $w(A) = \mathsf{t}$ iff $w(B) = \mathsf{f}$. Let $\neg A$ denote some such B.
  - If $A, B \in \mathcal{P}$ then there is some $C \in \mathcal{P}$ such that $w(C) = \mathsf{t}$ iff $w(A) = \mathsf{t}$ and $w(B) = \mathsf{t}$. Let $A \wedge B$ denote some such C.

For now, assume that W is finite.

I typically like to think in the sentential framework, but it shouldn't matter for this discussion. We can move back and forth between the frameworks. I think in some sense they're equivalent. There will be different once we consider, e.g., weakening the logic (allowing for more truth values). But that's a different paper.

Do we need to assume that $\mathcal{P}$ is a Boolean algebra? This allows us to extend any agenda, $\mathcal{A}$, to its Boolean closure. We need this to make use of some of the definition of probabilities (see below comment). We might need it to be able to define expectations relative to any p (which are defined using an extension of p to its Boolean closure). In fact, there might be multiple extensions, so really the notion of expectation relative to p isn't defined. Instead we quantify over all such extensions to define strict propriety. However in certain cases, e.g., when Acc is additive, the choice of extension doesn't in fact matter so expected accuracy relative to p was well defined.

Your total credence function is a function from some subset of $\mathcal{P}$ to $[0, 1]$. We call this subset your "opinion set", $\mathcal{O}$. It is the collection of all propositions that you have an opinion about. Your total credence function is then some function $c : \mathcal{O} \to [0, 1]$. Formally $\mathcal{O}$ is simply a fixed subset of $\mathcal{P}$.

Definition 1.2 (Opinion set)**.** An opinion set, $\mathcal{O}$, is a (fixed) collection of propositions, i.e., $\mathcal{O} \subseteq \mathcal{P}$.

We then have:

Definition 1.3 (Agenda)**.** An agenda, $\mathcal{A}$, is any collection of propositions that you have opinions about, i.e., $\mathcal{A} \subseteq \mathcal{O}$.

Formally, $\mathcal{A}$ and $\mathcal{O}$ are just the same sorts of objects: collections of propositions. So all our definitions, e.g., of credences, are defined for either of these. I don't know anyone else distinguishing them this sort of way, I'm playing with doing it. The idea of why I've kept them separate is that they seem to go with different pictures: An agenda is simply the propositions that I happen to be considering right now. Whereas an "opinion set" is the collection of all propositions that I have an opinion about. An accuracy measure that measures the accuracy of my (total) opinions, should rather work with $\mathcal{O}$. Pettigrew (2016) works with what he calls an "opinion set", but I think he rather uses the picture behind the agenda as I pictured it here. Fitelson uses "agenda". A number of authors call (either of) these an "algebra". But this really implies that they're Boolean algebras, which I don't want to imply.

We here allow them to be infinite. But we can't actually do accuracy stuff with infinite agendas yet.

Definition 1.4 (Credences)**.** A credence function on $\mathcal{A}$ is a function $c : \mathcal{A} \to [0, 1]$. $\mathsf{Creds}_{\mathcal{A}}$ is the collection of all credence functions on $\mathcal{A}$.

Note I've put $[0, 1]$ into the definition of credence function. One needn't do that. But perhaps I like Richard's idea that the numbers are arbitrary bounds, so it'd be meaningless to be outside them. He talks about this in a blog. I'm not sure if there's a better reference.

This defines the range of possible credence functions. *Your* $\mathcal{A}$ credence function is your total credence function, which is an $\mathcal{O}$ credence function, restricted to $\mathcal{A}$.

Definition 1.5 (Probabilities)**.** For $c \in \mathsf{Creds}_{\mathcal{A}}$, c is *probabilistic* iff c is in the convex hull of the $w : \mathcal{A} \to \{0, 1\}$, i.e. the smallest convex set containing such w. Let $\mathsf{Probs}_{\mathcal{A}}$ denote the collection of such c.

I.e. there are $\lambda_1, \ldots, \lambda_n$ positive summing to 1 and $w_1, \ldots, w_n \in W$ such that $c(A) = \sum_i \lambda_i w_i(A)$.

This definition is one that's also been used in non-classical probabilism (see, e.g., Williams, 2014) (although I disagree with some of that: I think the model of belief also needs to be altered).

Note that when we allow W to be infinite, we will be looking at *finitely additive* probabilities (and we will keep our attention on *Boolean* algebras rather than σ-algebras). We can define this as: c is probabilistic iff c is in the *closed* convex hull of the $w : \mathcal{A} \to \{0, 1\}$. (Note that for finite W, the convex hull is closed, so this doesn't alter the previous definition in the finite case.) Such finitely additive probabilities have the property that if p is probabilistic then for every finite $\mathcal{A}' \subseteq \mathcal{A}$, $p{\restriction}_{\mathcal{A}'}$ is probabilistic.

1.1 Other characterisations of being probabilistic

We also know we can characterise it as:

Proposition 1.6 (Paris, 2001)**.** *$c \in \mathsf{Probs}_{\mathcal{A}}$ iff c is not Dutch bookable, i.e. there is no finite $A_1, \ldots, A_n \in \mathcal{A}$ and $s_1, \ldots, s_n \in \mathbb{R}$ such that $\sum_i s_i(w(A_i) - c(A_i)) < 0$ for all w.*

The advantage of these characterisations is we didn't then need to initially assume a $\mathcal{P}$ forming a Boolean algebra. We might just start with a fixed $\mathcal{P}$ which is the collection of propositions I have opinions about, call this $\mathcal{O}$. We don't need to assume any closure principles. And we define probabilistic directly using this structure. I actually think that's a nicer way to go.

But some people will be unhappy with that as the primitive definition of probabilistic. Here's an alternative definition, which Pettigrew (2016) uses.

Proposition 1.7 (Probability)**.** *Assume that $\mathcal{P}$ forms a Boolean algebra (thus any $\mathcal{A}$ can be extended to a Boolean algebra)...*

$c \in \mathsf{Creds}_{\mathcal{A}}$ is probabilistic iff there is Boolean algebra, $\mathcal{B} \subseteq \mathcal{P}$ which extends $\mathcal{A}$ and with $p : \mathcal{B} \to [0,1]$ extending c, with p a finitely additive probability function which extends c. I.e.:

- $p(W) = 1$.
- $p(A) \geqslant 0$ *for all* A.
- $p(A \cup B) = p(A) + p(B)$ *if* $A \cap B = \varnothing$.

As an interesting sidenote, perhaps there are more general axiomatisations.For example:

Proposition 1.8. *If $\mathcal{A}$ is closed under conjunctions and containing $\varnothing$ and W, then p is extendable to a probability function iff:*

- *Normalization:* $p(W) = 1$
- *Emptyset:* $p(\varnothing) = 0$
- *∞-valuation (via-entailments): For C in $\mathcal{A}$ and K finite $\subseteq \mathcal{A}$,*

$$\text{If } C \supseteq \bigcup K, \text{ then } p(C) \geqslant \sum_{\varnothing \neq J \subseteq K} (-1)^{\#J+1} p(\bigcap J)$$
$$\text{If } C \subseteq \bigcup K, \text{ then } p(C) \leqslant \sum_{\varnothing \neq J \subseteq K} (-1)^{\#J+1} p(\bigcap J)$$

What about when $\mathcal{A}$ is not closed under conjunctions? I think this axiomatisation won't be enough, but I haven't actually got a counterexample yet.

2 Accuracy

2.1 Definition and some comments

Definition 2.1. An $\mathcal{A}$ accuracy measure is a function $\mathsf{Acc} : \mathsf{Creds}_{\mathcal{A}} \times W_{\mathcal{A}} \to \mathbb{R} \cup \{-\infty\}$.[1]

(Some people might think that accuracy is only something we apply to coherent agents, so it's $\mathsf{Acc} : \mathsf{Probs}_{\mathcal{A}} \times W \to \mathbb{R} \cup \{-\infty\}$. Of course such accuracy measures won't be useful for dominance arguments for probabilisim, basically by definition.)

[1] $W_{\mathcal{A}}$ is $\{w \restriction_{\mathcal{A}} \mid w \in W\}$.

Note that I've also put it into the definition of an accuracy measure that it only depends on one's credences in $\mathcal{A}$ propositions, and the truth values of those propositions. One might want to include that as a separate axiom of accuracy, but it seems so basic I've taken it as part of the definition.

Because of this assumption of extensionality we can see:

Remark. For $\mathcal{A}$ a singleton, $\mathcal{A} = \{B\}$, where B is contingent, $\mathsf{Creds}_{\{B\}} = [0, 1]$, and $W_{\{B\}} = \{\mathsf{t}, \mathsf{f}\}$, so a $\{B\}$ accuracy measure is a function from $[0, 1] \times \{\mathsf{t}, \mathsf{f}\}$ to $\mathbb{R} \cup \{-\infty\}$.

We can define axioms, or properties, of such $\mathcal{A}$ accuracy measures, for example:

Definition 2.2 (Strict propriety)**.** Let Acc be an $\mathcal{A}$ accuracy measure. Acc is strictly proper if for all $p \in \mathsf{Probs}_{\mathcal{A}}$ and $c \in \mathsf{Creds}_{\mathcal{A}} \setminus \{p\}$,

$$\mathrm{Exp}_p\mathsf{Acc}(p) > \mathrm{Exp}_p\mathsf{Acc}(c).$$

Actually for $\mathcal{A}$ non-Boolean, we need to be a little careful about defining expectation and how this works. We should probably actually define it as Richard does in Pettigrew (2016, Def.2.2.1). Though note that for additive Acc we don't need to be careful (Campbell-Moore and Levinstein, Forthcoming, ftnte.5).

2.2 Which agendas get to count?

One might only think that accuracy is meaningful if it is applied to one's total credence function. In which case we should simply focus on an $\mathcal{O}$ accuracy measure.

On the other hand, a number people think that's not right. It does make sense to talk about the accuracy of your credence in A in isolation? I think this probably is meaningful. In which case what we're working with is not simply a measure of accuracy but various accuracy measures for various agendas. We can make this precise:

Definition 2.3. For $\mathfrak{A}$ a collection of agendas, an $\mathfrak{A}$ accuracy notion defines an $\mathcal{A}$ accuracy measure for any $\mathcal{A} \in \mathfrak{A}$.

Most people working with additive accuracy measures probably think that what they're claiming is that the accuracy of a credence function is the sum of the accuracy of the individual credences that constitute it. This isn't that there's some functions acc that add to make the overall accuracy, but that these local function are the measures of accuracy of the individual credences. So this then states a notion of coherence between the accuracy measures on $\{B\}$ and the global accuracy measures on $\mathcal{O}$ (or other $\mathcal{A}$). It is possible to just work with accuracy as a formal constraint — there are some functions, call them acc, which sum to make the accuracy of one's total credences — but I think this picture is unusual.

However, even if one works with a more general accuracy notion and think that local accuracy — the accuracy of one's credences in individual propositions

— is meaningful, we should be careful about our justifications of properties of accuracy.

Is the argument for weak propriety really compelling if we are focusing on an agenda $\mathcal{A} \subset \mathcal{O}$. For example, suppose on a single proposition I think that some other credence value would be accuracy-theoretic better, but nonetheless when we take all propositions into account ($\mathcal{O}$), I think my total credence function is the best. Then that doesn't seem like such a bad epistemic failing. So this only motivates rational immodesty wrt $\mathcal{O}$ not $\mathcal{A}$ accuracy more generally. In which case we only really get a justification for weak propriety of the $\mathcal{O}$ accuracy measure. But, we can obtain the weak propriety of the local accuracy measures by going via justification of global weak propriety:

Proposition 2.4. *If an $\mathcal{O}$ accuracy measure is weakly proper and additive, then so are the constituent local measures.*

And in fact, furthermore justify the weak propriety on any $\mathcal{A} \subseteq \mathcal{O}$ where we take accuracy to be meaningful by use of the converse:

Proposition 2.5. *For an additive accuracy notion, if the local accuracy measures are weakly proper, then so is the $\mathcal{A}$ accuracy measure.*

I think this is particularly problematic for propriety principles. Note that Pettigrew (2016) doesn't rely on propriety principles in his characterisation of accuracy. One should go through his justifications with this sort of picture in mind and check if they're really compelling.

(What about truth-directedness? Again I think we should focus on directly justifying global truth directedness and showing that local truth directedness follows from it.)

So we should be careful about direct justifications of local accuracy theoretic principles.

Similarly, we should be careful about using the results of accuracy theoretic arguments with respect to partial agendas. Suppose my partial $\mathcal{A}$ credences are accuracy dominated, so there's some other $\mathcal{A}$ credence function which is better whatever the world is like (with respect to the $\mathcal{A}$ propositions), but that there is no other total credence function (on $\mathcal{O}$) which dominates my total credence function. This wouldn't give us a justification of being probabilistic on $\mathcal{A}$. Sure I could do accuracy theoretically better on $\mathcal{A}$ by moving to those other credences, but that would make my total credence function accuracy-theoretically worse. That doesn't mean I should thus change my credences.

Again, for additive accuracy measures this cannot happen.

Proposition 2.6. *Suppose we have an additive accuracy notion. Let c be a total credence function (defined on $\mathcal{O}$). If $\mathcal{A} \subset \mathcal{O}$ is such that $c\restriction_{\mathcal{A}}$ is accuracy theoretically dominated, then so c with respect to $\mathcal{O}$ accuracy.*

We also have a tie between $\mathcal{O}$ probabilities and $\mathcal{A}$ probabilities.

Proposition 2.7. *If an $\mathcal{O}$ credence function c is probabilistic then $c\restriction_{\mathcal{A}}$ is also probabilistic.*

(Note that in the case of infinitely many worlds, we are also interested in the more general claim: An $\mathcal{O}$ credence function c is probabilistic iff for every finite $\mathcal{A} \subseteq \mathcal{O}$, $c\restriction_{\mathcal{A}}$ is probabilistic.)

More generally, we should make our definitions of properties on the local and global side such that for additive accuracy measures we have equivalences between the local and global properties.

So I think that we should ultimately be interested in $\mathcal{O}$ accuracy, and any local or more general $\mathcal{A}$ accuracy considerations are intersting by virtue of their relationship to $\mathcal{O}$ accuracy.

(Note this is interestingly different to Dutch book considerations. It is an epistemic failure to be Dutch bookable on $\mathcal{A} \subset \mathcal{O}$.)

One consequence of this line of thought is that we really need to put more work into thinking about accuracy in infinite settings. Maybe I do actually have opinions about infinitely many propositions, i.e. $\mathcal{O}$ is infinite. So we should really be interested in how to apply accuracy considerations to this infinite set. If $\mathcal{O}$ is countable, then we can define a notion of global accuracy which is (weightedly) additive (see Walsh, Kelley, and SSK) and all our equivalences between global and local accuracies should go though. And all our probability notions are also global and locally equivalent (see Proposition 2.7) because we focused on finite additivity. So everything should go through nicely. But what happens when $\mathcal{O}$ is uncountable? (See Kelley.) We'll have to be very careful.

2.3 Examples

Example 2.8 (Brier)**.** The local Brier score is:[2] $\mathsf{B}(x,\mathsf{t}) = 1-(1-x)^2$, $\mathsf{B}(x,\mathsf{f}) = 1-x^2$.

For any finite $\mathcal{A} \subseteq \mathcal{P}$, we have an $\mathcal{A}$ accuracy measure: $\mathsf{B}(c,w) = \sum_{A\in\mathcal{A}} \mathsf{B}(c(A), w(A))$.

B is: strictly truth directed, strictly probabilistic proper.

More generally, whenever we have a function $\mathsf{acc} : [0,1] \times \{\mathsf{t},\mathsf{f}\} \to \mathbb{R} \cup \{-\infty\}$, we can generate accuracy measures for every finite $\mathcal{A} \subseteq \mathcal{P}$ just using additivity.

This is a reason that people working with additive measures usually aren't careful about what they're assuming $\mathcal{A}$ to be.

Here's another example:

Example 2.9 (Additive log)**.** The additive log score is generated by the local measure: $\mathsf{ALog}(x,\mathsf{t}) = \ln(x)$, $\mathsf{ALog}(x,\mathsf{f}) = \ln(1-x)$. So for any $\mathcal{A}$, we have an $\mathcal{A}$ accuracy measure: $\mathsf{ALog}(c,w) = \sum_{A\in\mathcal{A}} \mathsf{ALog}(c(A), w(A))$.

ALog is: strictly truth directed, strictly probabilistic proper.

Example 2.10 (Global log)**.** For a partition, $\mathcal{A}$, the global log score is: $\mathsf{GLog}(c,w) = \ln(c(E_w))$ for the $E_w \in \mathcal{A}$ which is true at w.

In fact, I think that people usually think that the global log measure only measures the accuracy of *probability* functions. In this sense it's not an $\mathcal{A}$ accuracy measure at all, as I defined that. Instead $\mathsf{GLog} : \mathsf{Probs}_{\mathcal{A}} \times W_{\mathcal{A}} \to \mathbb{R} \cup \{-\infty\}$.

[2] I added 1− to make it an accuracy measure rather than inaccuracy.

Pettigrew (2016) considers a slightly more general global log score. If $\mathcal{A}$ is such that any probabilistic c is uniquely extendable to the Boolean closure of $\mathcal{A}$ (in particular we get $c(E_w^{\mathcal{A}})$ well defined for $E_w^{\mathcal{A}} = \{v \in W \mid v(A) = w(A)$ for all $A \in \mathcal{A}\}$ the atom of the Boolean algebra containing w). Then we can put $\mathsf{GLog}(c, w) = \ln(c(E_w^{\mathcal{A}}))$ when c is probabilistic.

This is a way of extending the lattice notion of global log to a Boolean algebra. Some agendas, though, won't suffice. For example, $\{A \cap B, A \cup B\}$.

Note that the singleton does suffice. There is a unique extension of every $x \in [0, 1]$ to the Boolean closure $\{B, \neg B\}$. And in that sense we get $\mathsf{GLog}(x, \mathsf{t}) = \ln(x)$, $\mathsf{GLog}(x, \mathsf{f}) = \ln(1 - x)$.

However, note that GLog is not the addition of these, that generates the additive log score, which is different. So this is not additive.

If you think that the Global log score is only defined on a partition, then there is a sense in which we can say it is additive: we can find some local functions which sum to generate the global log score: $\mathsf{LocalPartitionalGLog}(x, \mathsf{t}) = \ln(x)$, $\mathsf{LocalPartitionalGLog}(x, \mathsf{f}) = 0$. We could then extend this to any agenda using additivity. This would generate something that differs from the previous definition. Is it interesting?

Officially this extension of the global log was only defined on probability functions. It could of course be extended to a more general accuracy measure by setting, e.g., $\mathsf{Acc}(c, w) = 0$ if c is not probabilistic. But we shouldn't then expect e.g. truth directedness to hold. Any of our properties only get to be justified when all credence functions under consideration are probabilistic. So it'd justify *probabilistic* strict propriety, and *probabilistic* truth directedness. And these global properties do not entail their local analogues.

GLog is: weakly truth directed, strictly probabilistic truth directed, strictly probabilistic proper.

$\mathsf{LocalPartitionalGLog}$ is not (probabilistic) truth directed or (probabilistic) weakly proper.

References

Catrin Campbell-Moore and Benjamin A Levinstein. Strict propriety is weak. *Analysis*, Forthcoming.

Jeff B. Paris. A note on the dutch book method. In *Proceedings of the Second International Symposium on Imprecise Probabilities and their Applications, ISIPTA*, pages 301–306, 2001.

Richard Pettigrew. *Accuracy and the Laws of Credence*. Oxford University Press, 2016.

J. Robert G. Williams. Probability and non-classical logic. In Christopher Hitchcock and Alan Hájek, editors, *Oxford Handbook of Probability and Philosophy*. Oxford University Press, 2014.