Additivity and the opinion-set

Catrin Campbell-Moore*

April 24, 2020

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.

I think we might also need to require that W is closed (using the powerset topology on $\{\mathsf{t}, \mathsf{f}\}^{\mathcal{P}}$).

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. But that's not a job for now.

We need to assume that there's a Boolean structure around in the background to make use of the usual definition of probabilities (see below comment). If we want to use alternative definitions of probability we can completely delete that requirement that $\mathcal{P}$ forms a Boolean algebra.

Definition 1.2 (Agenda)**.** An agenda, $\mathcal{A}$, is any collection of propositions, i.e., $\mathcal{A} \subseteq \mathcal{P}$.

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

Many people call this the "algebra", but that's because they're assuming it is a Boolean algebra. This is called an "opinion set" by Richard. Fitelson uses the "agenda" term. I think the two terms, "agenda" and "opinion set", come along with different pictures of what it's doing. 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.

Definition 1.3 (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.

I'm not sure this is the best definition to take as the primary one, but here's what it is to be probabilistic (eg Pettigrew, 2016)

Definition 1.4 (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 $A \cap B = \varnothing$.

1.1 Other characterisations of being probabilistic

We also know we can characterise it as:

- $c \in \mathsf{Probs}_{\mathcal{A}}$ iff c is in the convex hull of the $w : \mathcal{A} \to \{0, 1\}$. If $\mathcal{A}$ is finite, this is just that there are $\lambda_1, \ldots, \lambda_n$ positive summing to 1 and $w_1, \ldots, w_n \in W$ such that $c(\varphi_i) = \sum_i \lambda_i w_i(A)$. If $\mathcal{A}$ is infinite, then ...

To do. Remember we ensured W closed.

- $c \in \mathsf{Probs}_{\mathcal{A}}$ 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. We'd need to justify it more. Anyway, since I already wrote it up with the Boolean extensions, I'll leave it for now.

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

Proposition 1.5. *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.

2 Accuracy

There are two ways of thinking about accuracy and agendas.

Let $\mathcal{O}$ be a fixed agenda. We will call this one my "opinion set". We think of it as: all the propositions I have opinions about. If we had a setup where $\mathcal{P}$ wasn't required to be a Boolean algebra, we can just put $\mathcal{P}$ as $\mathcal{O}$: all the propositions that there are are the ones that I have opinions about.

Definition 2.1. An $\mathcal{O}$ accuracy measure is a function $\mathsf{Acc} : \mathsf{Creds}_{\mathcal{O}} \times W \to \mathbb{R} \cup \{-\infty\}$.

In fact some people (perhaps eg the global log score) might not think of accuracy as defined on $\mathsf{Creds}_{\mathcal{O}}$ but only on $\mathsf{Probs}_{\mathcal{O}}$. I ignore that right now.

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

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

Definition 2.3. An $\mathcal{O}$ accuracy measure is (weightedly-)additive if there are functions $\mathsf{acc}_B : [0,1] \times \{\mathsf{t}, \mathsf{f}\} \to \mathbb{R} \cup \{-\infty\}$ for every $B \in \mathcal{P}$ such that

$$\mathsf{Acc}(c, w) = \sum_{B \in \mathcal{A}} \mathsf{acc}_B(c(B), w(B)).$$

A feature of this picture is that acc is merely a formal tool: it's just the function that generates Acc. We directly justify propriety of Acc. Any properties of acc are formal tools. We don't get to justify propriety of acc by thinking about goodness of individual credence functions. Instead we justify strict propriety of Acc, and if we want to use any properties of acc we should show they follow from properties of Acc. So we are then interested in proving results like

Proposition 2.4. *If* Acc *is additive, and is generated by* acc_B*, and* Acc *is strictly proper, then so is* acc_B*, where this means:*

> *for all* $p \in \mathsf{Probs}_{\mathcal{O}}$ *and* $c \in \mathsf{Creds}_{\mathcal{O}} \setminus \{p\}$*,* $\mathrm{Exp}_p\mathsf{acc}(p) > \mathrm{Exp}_p\mathsf{acc}(c)$*.*

which is equivalent to:

> *For any* $x \in [0,1]$ *and* $y \neq x$*,*
>
> $$x \times \mathsf{acc}(x, \mathsf{t}) + (1-x) \times \mathsf{acc}(x, \mathsf{f}) > x \times \mathsf{acc}(y, \mathsf{t}) + (1-x) \times \mathsf{acc}(y, \mathsf{f})$$

I think I like this way of going. The only issue is that accuracy when $\mathcal{O}$ is infinite becomes difficult. And I think it really might be that our actual opinion set is infinite. In which case, how should I think about it?

Alternatively, we could think, we only get to consider accuracy when we restrict $\mathcal{O}$ to a finite collection, $\mathcal{A}$. So we have $\mathcal{A}$ accuracy measures for a range of finite collections of propositions. And those are what we should theorise about.

Here's the alternative picture: the notion of accuracy gives you an $\mathcal{A}$ accuracy measure for a range of $\mathcal{A}$. Which range? Perhaps: any $\mathcal{A} \subseteq \mathcal{O}$, (we might for now just think $\mathcal{O} = \mathcal{P}$ and ignore Boolean worries). Or perhaps just the $\mathcal{A}$ which form Boolean algebras, or perhaps just the $\mathcal{A}$ which are partitions.

This seems to fit with Pettigrew (2016) as he seems to actually justify properties of Acc by considering the local accuracy. He thinks that local accuracy is a meaningful thing that can be philosophised about.

Perhaps this goes along with an alternative picture. One might think accuracy is a more general thing: it gives you an $\mathcal{A}$ accuracy measure for a *range of* $\mathcal{A}$. Which range? Perhaps: any $\mathcal{A} \subseteq \mathcal{P}$, (or any $\mathcal{A} \subseteq \mathcal{O}$, where $\mathcal{O}$ contains all the propositions that I have opinions about). Or perhaps just the $\mathcal{A}$ which form Boolean algebras, or perhaps just the $\mathcal{A}$ which are partitions.

Definition 2.5. For $\mathfrak{A}$ a collection of agendas, an $\mathfrak{A}$ accuracy notion defines an $\mathcal{A}$ accuracy measure for any $\mathcal{A} \in \mathfrak{A}$.

If we have a more general picture of accuracy, an $\mathfrak{A}$ accuracy notion, where the singletons are in $\mathfrak{A}$, we might define additivity as:

Definition 2.6. If $\mathfrak{A}$ contains all the singletons $\{B\}$ for $B \in \mathcal{O}$. For an $\mathfrak{A}$ accuracy notion, we can let $\mathsf{acc}_B : [0,1] \times \{\mathsf{t}, \mathsf{f}\} \to \mathbb{R} \cup \{-\infty\}$ be the $\{B\}$ accuracy measure given by the accuracy notion.[1] Then we can say that this accuracy notion is additive if for every $\mathcal{A} \in \mathfrak{A}$,

$$\mathsf{Acc}(c, w) = \sum_{B \in \mathcal{A}} \mathsf{acc}_{\varphi}(c(B), w(B)).$$

With this picture we then get to justify constraints directly on acc.

2.1 Examples

Whenever we define a function $\mathsf{acc} : [0,1] \times \{\mathsf{t}, \mathsf{f}\} \to \mathbb{R} \cup \{-\infty\}$, we can generate a $\wp(\mathcal{P})$ accuracy notion just using additivity. So, for example:

Example 2.7. The Brier score is $\mathsf{BS}(x, \mathsf{t}) = 1 - (1-x)^2$, $\mathsf{BS}(x, \mathsf{f}) = 1 - x^2$.[2]

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

Example 2.8. The global log score can be though of this way: $\mathsf{GlLog}(x, \mathsf{t}) = \ln(x)$, $\mathsf{GlLog}(x, \mathsf{f}) = 0$. For any $\mathcal{A}$, we have an $\mathcal{A}$ accuracy measure: $\mathsf{GlLog}(c, w) = \sum_{A \in \mathcal{A}} \mathsf{GlLog}(c(A), w(A))$.

[1] Officially I defined a $\{B\}$ accuracy measures as $B : \mathsf{Creds}_{\{B\}} \times W \to \mathbb{R} \cup \{-\infty\}$. Now, $\mathsf{Creds}_{\{B\}}$ are simply values $x \in [0,1]$. We also impose the additional constraint of "extensionality": the accuracy can only depend on the truth value of B itself. So we can thus legitimately describe such acccuracy measures as functions $\mathsf{acc}_B : [0,1] \times \{\mathsf{t}, \mathsf{f}\} \to \mathbb{R} \cup \{-\infty\}$.

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

But this isn't how people usually conceive of it. I think people only think: the global log score is a score that's only defined when $\mathcal{A}$ is a partition. And in that case, $\mathsf{GlLog}(c, w) = \ln(c(A_w))$ for the $A_w \in \mathcal{A}$ which is true at w.

In fact, I think the Global Log score is usually thought of not as an $\mathcal{A}$ accuracy measure in my sense at all, they think: it's only defined over *probability* functions.

References

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