Believing Probabilistic Contents: On the Expressive Power and Coherence of Sets of Sets of Probabilities

Catrin Campbell-Moore and Jason Konek

November 1, 2019

Moss (2018) argues that rational agents are best thought of not as having degrees of belief in various propositions. Instead, they are best thought of as having *beliefs in probabilistic contents*, or *probabilistic beliefs*. Probabilistic contents are sets of probability functions. Probabilistic belief states, in turn, are modeled by sets of probabilistic contents, or sets of sets of probability functions. We argue that this Mossean framework is of considerable interest quite independently of its role in Moss' account of probabilistic knowledge or her semantics for epistemic modals and probability operators. It is an extremely general model of uncertainty. Indeed, it is at least as general and expressively powerful as every other current imprecise probability framework, including lower probabilities, lower previsions, sets of probabilities, sets of desirable gambles, and choice functions. In addition, we partially answer an important question that Moss leaves open, *viz.*, why should rational agents have consistent probabilistic beliefs? We show that an important subclass of Mossean believers avoid Dutch bookability iff they have consistent probabilistic beliefs.

1 The Mossean Framework

On the traditional Bayesian view, rational agents do not simply categorically believe or disbelieve propositions. Rather, they have *degrees of belief* or *credences* in those propositions. For example, rather than categorically believing that you will make it through security at the airport in under 30 minutes, you might be 0.9 confident that you will make it. Having a degree of belief, on this view, is a matter of taking a complex attitude—0.9 confidence—toward a simple possible-worlds proposition: something that is either true or false at a world.[1]

Moss agrees that rational agents do not just categorically believe or disbelieve simple possible-worlds propositions. But on the Mossean view, they do not bear complex attitudes toward those propositions either. Rather, they bear a simple attitude toward a complex content. More specifically, they believe *probabilistic contents*, which, formally, is a set of probability functions, $p : \mathcal{F} \to \mathbb{R}$, where $\mathcal{F}$ is a σ-algebra on a domain of possibilities Ω.[2]Moss calls such attitudes *probabilistic beliefs*.

[1]For a detailed introduction to precise and imprecise credences, see Titelbaum (2019), Mahtani (2019) and Konek (2019).

[2]In fact, Moss takes probabilistic contents to be sets of *probability spaces*. A probability

This framework allows us to model our opinions in a natural way. For example, you might think that it is more likely to rain than not. What you think expresses a constraint on probabilities, *viz.*, that the probability of rain is greater than the probability of no rain: $p(\mathit{Rain}) > p(\mathit{No\ Rain})$. And on Moss's view, we model your opinion by saying that you believe the probabilistic content

$$\{p \mid p(\mathit{Rain}) > p(\mathit{No\ Rain})\}.$$

We can model your opinion that the Greenland ice sheet is between 60% and 80% likely to melt by saying that you believe the probabilistic content

$$\{p \mid 0.6 \leqslant p(\mathit{Greenland\ ice\ sheet\ melts}) \leqslant 0.8\}.$$

We can model your opinion that the recent downturn in the stock market provides no evidence one way or the other about whether dinosaurs were wiped out by asteroids by saying that you believe the probabilistic content

$$\{p \mid p(\mathit{Dinos\text{-}asteroids} \mid \mathit{Stocks\ down}) = p(\mathit{Dinos\text{-}asteroids})\}.$$

Finally, we model your total doxastic state by collecting all of the probabilistic contents that you believe into a big set: a set of probabilistic contents, or a set of sets of probability functions.

Moss provides a number of reasons to prefer this framework to the traditional Bayesian framework where you model an agent's credences in simple possible-worlds propositions by either a single probability function (if the agent has precise credences) or a set of probability functions (if she has imprecise credences). For example, it seems *irrational* to both be 80% confident that it will rain tonight and 80% confident that it will not. Proponents of the Mossean framework can tell a simple story about why this is so. Being 80% confident that it will rain tonight is simply a matter of believing the probabilistic content

$$P_R = \{p \mid p(\mathit{Rain}) = 0.8\}$$

Likewise, being 80% confident that it will not rain tonight is a matter of believing the probabilistic content

$$P_N = \{p \mid p(\neg \mathit{Rain}) = 0.8\}$$

But these two contents are *inconsistent*. They are inconsistent because they are disjoint sets of probabilities: $P_R \cap P_N = \varnothing$. And plausibly it is irrational to believe inconsistent contents (Moss, 2018, p. 11, this is something we will revisit in section 3). On the traditional Bayesian account, in contrast, we must explain why having a credence of 0.8 in a proposition and its negation is irrational by telling some more complicated, perhaps more controversial story—*e.g.*, by showing that having such credences renders you Dutch bookable (see, e.g., Pettigrew, 2019), or accuracy-dominated (Joyce, 1998, 2009; Pettigrew, 2019), etc.

Proponents of the Mossean framework can also tell a simply story about apparent ordinary language quantification over contents (Moss, 2018, p. 12). For example, suppose that we both think that the Greenland ice sheet is between

space is a triple consisting of a domain of possibilities Ω, a σ-algebra $\mathcal{F}$ on Ω, and a probability measure p on $\mathcal{F}$. We will assume a fixed *finite* domain Ω and σ-algebra $\mathcal{F}$ in what follows. This also means we do not have to deal with issues of countable vs finite additivity.

60% and 90% likely to melt. Then after a little research you come to think that it is between 70% and 80% likely. It is natural to say that you have come to believe something that I do not. Proponents of the Mossean framework can explain this by saying that you have come to believe the probabilistic content

$$\{p \mid 0.7 \leqslant p(\textit{Greenland ice sheet melts}) \leqslant 0.8\}$$

I have not. On the traditional Bayesian account, in contrast, we must say that this sort of quantified claim is elliptical for the claim that you have adopted a new imprecise credal attitude toward a simple possible-words proposition—the Greenland ice sheet will melt—which I have not adopted.

2 Comparison with other Imprecise Probability Frameworks

Whether or not you are swayed by these considerations, there is good independent reason to explore the Mossean framework. Sets of probabilistic contents, or sets of sets of probabilities provide an extremely general model of uncertainty. Indeed, the Mossean framework is at least as general and expressively powerful as every other current imprecise probability framework, including lower probabilities, lower previsions, sets of probabilities, sets of desirable gambles, and choice functions. Drawing on the work of Walley (2000), Seidenfeld et al. (2010), Quaeghebeur et al. (2015), De Bock and de Cooman (2018) and Van Camp and Miranda (2019) we can rank these frameworks in order of increasing generality:

Coherent lower and upper probabilities

⇓

Coherent lower and upper previsions / Closed convex sets of probabilities

⇓

Coherent sets of desirable gambles / Partial preference relations

⇓

Coherent sets of desirable gamble sets / Choice functions

⇓

Moss's framework: Sets of sets of probabilities

2.1 Lower and upper probabilities

Let's begin with lower and upper probabilities. An agent's lower and upper probabilities do not encode *precise* degrees of confidence or credence. Rather, they encode the agent's *minimum* and *maximum* degrees of confidence. For example, rather than being exactly 70% confident that you will make it through security at the airport in under 30 minutes, you might be at least 50% and at most 90% confident. If there is nothing stronger to about how confident you are, then your lower probability that you will make it through in under 30 minutes is 0.5 and your upper probability is 0.9.

We can model these lower and upper probabilities by a pair of functions, $\underline{p} : \mathcal{F} \to \mathbb{R}$ and $\overline{p} : \mathcal{F} \to \mathbb{R}$. In our example, we have

$$\underline{p}(\textit{Make it in under 30 min}) = 0.5, \qquad \overline{p}(\textit{Make it in under 30 min}) = 0.9.$$

We call lower and upper probabilities coherent iff they are given by a set of probabilities, P (Walley, 1991, p. 135):[3]

$$\underline{p}(X) = \inf \{p(X) \mid p \in P\}, \qquad \overline{p}(X) = \sup \{p(X) \mid p \in P\}$$

These probabilities provide a more general model of uncertainty than precise probabilities. But they nevertheless fail to capture various types of opinions that are important for inference and decision-making. For example, they cannot capture *comparative beliefs* (Walley, 2000, p. 130). Comparative beliefs are opinions of the form *X is at least as likely as Y*, which we write as $X \succcurlyeq Y$. (Strictly speaking, these are *weak* comparative beliefs. See section 2.2.) To illustrate, suppose that you have the following comparative beliefs about whether your favourite team will win, draw or lose the upcoming match.

$$\mathit{Lose} \vee \mathit{Draw} \succcurlyeq \mathit{Win} \succcurlyeq \mathit{Draw} \succcurlyeq \mathit{Lose}$$

The maximum and minimum levels of confidence consistent with these comparative beliefs are given by:

	Lose ∨ *Draw*	*Win*	*Draw*	*Lose*
$\overline{p}$	2/3	1/2	1/2	1/3
$\underline{p}$	1/2	1/3	1/4	0

Notice though that $\underline{p}$ and $\overline{p}$ leave open whether *Win* is more likely than *Draw*, or *Draw* is more likely than *Lose*. For example, the following probability function is consistent with the constraints given by $\underline{p}$ and $\overline{p}$:

	Lose ∨ *Draw*	*Win*	*Draw*	*Lose*
q	2/3	1/3	1/2	1/6

According to q, *Win* is less likely than *Draw*. The upshot: comparative beliefs cannot in general be recovered from their corresponding lower and upper probabilities. These are thus not sufficiently expressively powerful to model comparative beliefs.

The Mossean framework, in contrast, is powerful enough to model both coherent lower and upper probabilities, as well as comparative beliefs. You have the coherent lower and upper probabilities, $\underline{p}$ and $\overline{p}$, in this framework, just in case for all $X \in \mathcal{F}$ you believe

$$\{p \mid \underline{p}(X) \leqslant p(X) \leqslant \overline{p}(X)\}$$

and you do not believe any other probabilistic contents which would determine a narrower interval.[4] You have the comparative beliefs given by $\succcurlyeq$ just in case

$$X \succcurlyeq Y \ \textit{iff you believe} \ \{p \mid p(X) \geqslant p(Y)\}.$$

[3] However, different sets of probabilities can give rise to the same lower and upper probabilities: sections 2.2 and 2.3 note that lower probabilities cannot capture weak comparative beliefs but sets of probabilities can. (See also Walley, 2000, §4-5.)

[4] I.e. you do not believe any Q such that $\underline{Q}(X) := \inf \{q(X) \mid q \in Q\} > \underline{p}(X)$, or $\overline{Q}(X) := \sup \{q(X) \mid q \in Q\} < \overline{p}(X)$.

2.2 Lower and upper previsions / Closed convex sets of probabilities

The expressive limitations of lower and upper probabilities motivated proponents of imprecise probabilities to move to a more general framework: lower and upper previsions. Linear previsions are precise estimates of quantities of interest. For example, your prevision or estimate of the price of a given stock 12 months hence might be £220 per share. *Lower and upper previsions* capture bounds on best estimates. For example, you might think that the best estimate of the price of the stock in question is at least £150 and at most £300 per share. If there is nothing stronger to say about how high or low you think that best estimate is, then these are your lower and upper previsions.

Lower and upper previsions are given by a pair of functions $\underline{\text{est}} : \mathcal{G} \to \mathbb{R}$ and $\overline{\text{est}} : \mathcal{G} \to \mathbb{R}$ defined on the space $\mathcal{G}$ of all gambles which are bounded functions $G : \Omega \to \mathbb{R}$.

Coherence constraints are given on lower and upper previsions, which makes them equivalent to closed convex sets of probabilities (Walley, 2000, p. 133). The coherent lower and upper previsions given by a set of probabilities, P, are

$$\underline{\text{est}}(G) = \inf \left\{ \text{Exp}_p[G] \mid p \in P \right\}, \qquad \overline{\text{est}}(G) = \sup \left\{ \text{Exp}_p[G] \mid p \in P \right\}$$

And for each coherent lower and upper previsions there is a unique such closed and convex set of probabilities.

These provide a more general model of uncertainty than lower and upper probabilities.[5] For example, they are general enough to capture comparative beliefs. You have the comparative beliefs given by $\succcurlyeq$ just in case for all $X, Y \in \mathcal{F}$[6]

$$X \succcurlyeq Y \textit{ iff } \underline{\text{est}}(X - Y) \geqslant 0$$

But lower and upper previsions are still not general enough. They fail to capture some important types of opinions. We have focussed thus far on *weak* comparative beliefs, *i.e.*, opinions of the form X *is at least as likely as* Y ($X \succcurlyeq Y$). But rational agents also have *strict* comparative beliefs, *i.e.*, opinions of the form X *is strictly likelier than* Y ($X \succ Y$). Lower and upper previsions cannot distinguish weak comparative beliefs from strict comparative beliefs (Walley, 2000, p. 135).

On the face of it, they seem to have the flexibility to distinguish between weak and strict comparative beliefs. Why not say, for example, that

$$X \succcurlyeq Y \textit{ iff } \underline{\text{est}}(X - Y) \geqslant 0 \quad \textit{and} \quad X \succ Y \textit{ iff } \underline{\text{est}}(X - Y) > 0.$$

To see why this proposal will not work, consider the following example. Suppose that you are throwing darts, and suppose that your next dart is equally likely to hit each of the uncountably many points on the dartboard. Consider the proposition that it will hit the point-sized bullseye, B, and the proposition that it will hit exactly the line-sized edge of the inner and outer ring, E. You think that E is strictly more likely than B, $E \succ B$; but there is no amount by which it is more likely: it is not 0.1 more likely, or 0.01, or any positive real. Since you will thus have $\underline{\text{est}}(E - B) < \epsilon$ for all positive reals, $\epsilon > 0$, the structure

[5] We can capture any lower probability function by a lower prevision function restricted to indicator variables.

[6] Here we identify X with its indicator gamble.

of the reals will require that $\underline{\text{est}}(E - B) = 0$. Thus, your strict comparative opinion cannot be captured by your lower and upper previsions.

The Mossean framework is powerful enough to model both coherent lower and upper previsions, as well as weak and strict comparative beliefs. You have the lower and upper previsions given by $\underline{\text{est}}$ and $\overline{\text{est}}$, in Moss' framework, just in case for all G you believe

$$\left\{p \mid \underline{\text{est}}(G) \leqslant \text{Exp}_p[G] \leqslant \overline{\text{est}}(G)\right\}$$

and you do not believe any other probabilistic content which would determine a narrower interval.

Moss's framework can also capture both weak comparative beliefs (section 2.1) and strict comparative beliefs. You have the strict comparative beliefs given by $\succ$ just in case

$$X \succ Y \textit{ iff you believe } \{p \mid p(X) > p(Y)\}.$$

In the darts example, you have the strict comparative belief $E \succ B$ by virtue of believing that E is more likely than B but not by any positive amount. So you believe $\{p \mid p(E) > p(B)\}$ but also $\{p \mid p(E) - p(B) < \epsilon\}$ for each $\epsilon > 0$. In this example you have *infinitesimal probability* in both E and B:

$$\textit{you believe } \{p \mid p(X) > 0\} \textit{ and for every } \epsilon > 0 \textit{ you believe } \{p \mid p(X) < \epsilon\}.$$

Note, though, that this collection of beliefs is infinitely inconsistent: no (real-valued) probability function is compatible with everything that you believe.

2.3 Sets of desirable gambles

The expressive limitations of lower and upper previsions motivated the shift to sets of desirable gambles (Walley, 2000, §6). An agent's desirable gamble set $\mathcal{D}$ rounds up all of the gambles G that she strictly prefers to the status quo—the constant gamble $\mathbf{0}$ that takes the value 0 at every world $w \in \Omega$.

Sets of desirable gambles provide a more general model of uncertainty than coherent lower and upper previsions. For example, sets of desirable gambles are general enough to capture both weak and strict comparative beliefs by:[7]

$$X \succ Y \textit{ iff } X - Y \in \mathcal{D}.$$
$$X \succcurlyeq Y \textit{ iff } X - Y + \epsilon \in \mathcal{D} \textit{ for all } \epsilon > 0.$$

A set of desirable gambles $\mathcal{D}$ is said to be coherent iff it satisfies:

1. $\mathbf{0} \notin \mathcal{D}$
2. If $G(w) \geqslant 0$ for all $w \in \Omega$, and $G(w) > 0$ for some $w \in \Omega$, then $G \in \mathcal{D}$.[8]
3. If $G \in \mathcal{D}$ and $\lambda > 0$, then $\lambda G \in \mathcal{D}$
4. If $F, G \in \mathcal{D}$, then $F + G \in \mathcal{D}$

[7] As Walley (1991, p. 151) notes, sets of desirable gambles are *equivalent* to partial preference relations, or comparative prevision/estimation relations. These latter models are more general than comparative belief relations.

[8] We think this axiom is too strong. We conjecture that we can still show representability when we replace it with the two axioms: If $G(w) > 0$ for all $w \in \Omega$, then $G \in \mathcal{D}$. If $G \in \mathcal{D}$ and $F(w) \geqslant G(w)$ for all $w \in \Omega$, then $F \in \mathcal{D}$. And that this corresponds to dropping axiom 4. for probabilistic beliefs.

Once again, the Mossean framework is powerful enough to model coherent sets of desirable gambles. Probabilistic beliefs determine a set of desirable gambles by:

$$G \in \mathcal{D} \textit{ iff you believe } \{p \mid \mathrm{Exp}_p[G] > 0\}$$

and every coherent set of desirable gambles is generated by some probabilistic beliefs.[9] We can ensure that the resultant set of desirable gambles is coherent if your probabilistic beliefs satisfy:

1. You believe Ω and do not believe $\emptyset$.
2. If you believe P and Q then you believe $P \cap Q$.
3. If you believe P and $Q \supseteq P$ then you believe Q.
4. For each $w \in \Omega$, you believe $\{p \mid p(w) > 0\}$.[10]

So the Mossean framework is at least as general as coherent sets of desirable gambles. But once again sets of desirable gambles are not general enough. For example, coherent sets of desirable gambles are incapable of capturing the types of opinions modelable by non-convex sets of probability functions. Consider an example from (Van Camp and Miranda, 2019, p. 419). Suppose you know that a coin is double-sided but have no idea whether it has two heads or two tails. The gambles that are desirable to you are just those that have positive payout if it lands heads and positive payouts if it lands tails. I.e. your set of desirable gambles is just

$$\mathcal{D} = \{G \mid G(\textit{Heads}) > 0 \textit{ and } G(\textit{Tails}) > 0\}.$$

But this is exactly the same set as someone who has no information about the bias of the coin. So sets of desirable gambles fail to distinguish between importantly distinct doxastic states.

The Mossean framework handles this example straightforwardly: You believe the non-convex probabilistic content

$$\{p \mid p(\textit{Heads}) = 1 \textit{ or } p(\textit{Tails}) = 1\}.$$

2.4 Sets of desirable gamble sets / Choice functions

Limitations of this sort spurred Seidenfeld et al. (2010), De Bock and de Cooman (2018), and Van Camp and Miranda (2019) to adopt an even more general model of uncertainty: sets of desirable gamble sets, or equivalently, choice functions or rejection functions.[11] These are the most general models of uncertainty currently under investigation in the imprecise probability community. They can represent any set of probabilities, including non-convex ones,[12] as well as any set of desirable gambles.

To see the expressive power of this new framework, consider the following gambles:

[9] Suppose $\mathcal{D}$ is coherent and $G \notin \mathcal{D}$. Use the hyperplane separation theorem to show that for each $F_1, \ldots, F_n \in \mathcal{D}$ we can find probabilistic p with $\mathrm{Exp}_p[F_i] > 0$ and $\mathrm{Exp}_p[G] \leqslant 0$. So we can find probabilistic beliefs (indeed, ones satisfying the principles stated) that generate $\mathcal{D}$.

[10] This is required to satisfy axiom 2.

[11] Choice or rejection functions choose, for each collection of gambles, a subset of the better/worse gambles in the set. This is equivalent to sets of desirable gamble sets. (De Bock and de Cooman, 2018)

[12] See Seidenfeld et al. (2010, Theorem 2).

	Heads	Tails
G_H	£1	−£10
G_T	−£10	£1

If you knew the coin has two heads, then G_H would be desirable. If you knew it had two tails, then G_T would be desirable. If you know that it is double-sided, then although neither of these individual gambles is desirable, you know that the collection of gambles $\{G_H, G_T\}$ contains at least one desirable gamble. Someone who knows nothing about the coin will not think that this necessarily contains a desirable gamble, for example, if the coin is fair then neither of the gambles is desirable.

We therefore model your doxastic state with a set of desirable gamble sets, $\mathcal{K}$, which rounds up all of the set of gambles A that contain at least one desirable gamble G.

Beliefs in probabilistic contents determine a set of desirable gamble sets by:

$$A \in \mathcal{K} \textit{ iff you believe } \left\{p \mid \mathrm{Exp}_p[G] > 0 \textit{ for some } G \in A\right\}.$$

We say a set of desirable gamble sets is coherent if:

1. $\emptyset \notin \mathcal{K}$
2. $A \in \mathcal{K}$ then $A \setminus \{0\} \in \mathcal{K}$.
3. If $G(w) \geqslant 0$ for all $w \in \Omega$, and $G(w) > 0$ for some $w \in \Omega$, then $\{G\} \in \mathcal{K}$.[13]
4. If $A, B \in \mathcal{K}$, and $\lambda_{G,F}, \mu_{G,F} \geqslant 0$ with at least one > 0, then
$$\{\lambda_{G,F} G + \mu_{G,F} F \mid G \in A,\ F \in B\} \in \mathcal{K}$$
5. If $A \in \mathcal{K}$ and $B \supseteq A$, then $B \in \mathcal{K}$.
6. If the closed convex hull of A is $\in \mathcal{K}$, then $A \in \mathcal{K}$.[14]

Any coherent set of desirable gamble sets can be recovered from probabilistic beliefs satisfying our conditions in section 2.3.[15] Moreover, any probabilistic beliefs satisfying those conditions generate a coherent set of desirable gamble sets.

The Mossean framework can also capture some believers who are not represented in the set of desirable gamble sets framework. Imagine, for example, that you are throwing darts on a board with a special line-sized "bonus ring." Hitting the ring wins the game automatically. But it is only in play at certain times. The Mossean framework can distinguish between someone who believes that the probability of hitting the ring and winning automatically is 0 and someone who leaves open whether it is 0 or infinitesimal. The set of desirable gamble sets framework cannot.[16]

We have thus seen that Moss's framework provides a very natural and expressively powerful model of uncertainty.

[13] As in footnote 8, we think this should be replaced with weaker axioms and conjecture that representability still holds.

[14] This final axiom is not included in De Bock and de Cooman (2018). It is analogous to Axiom 2b from Seidenfeld et al. (2010), and it is needed for the representation by probabilistic beliefs.

[15] This requires the additional mixtures axiom 6..

[16] This is so even when axiom 4. is weakened. Extensions of the desirability framework, though, will probably be able to deal with it.

3 Consistency

One important project in epistemology is to determine what it takes for your probabilistic beliefs to be rational or irrational. Now that we have an alternative framework for representing your doxastic state, we have to reconsider these questions. Moss can say quite a lot about rationality if she can make use of the following norm:

> **Consistency Norm:** It is irrational to have beliefs whose probabilistic contents are inconsistent.

This identifies a whole class of Mossean believers as irrational. Moss proposes that this is an advantage of her account: it can explain the irrationality of your probabilistic beliefs directly in terms of the contents of your beliefs rather than the traditional Bayesian approach which would have to give a story in terms of, e.g., Dutch book or accuracy-based arguments.

However, more needs to be done to justify the consistency norm: Is it irrational to believe inconsistent probabilistic contents? If so, why? Easwaran and Fitelson (2015) provide us with a cautionary tale in the case of full beliefs towards simple possible-worlds propositions. They argue that plausible normative principles only justify a coherence constraint that is weaker than the consistency norm. The same might hold for Moss's framework too. Moreover, in representing your opinions regarding the dart hitting a point-sized bullseye we represented her as having inconsistent probabilistic beliefs and it wasn't obviously irrational.[17]

To attempt to justify the Consistency Norm, we will turn to the sorts of resources that are given for justifying probabilism for the traditional Bayesian epistemologist. In particular we will consider Dutch book style arguments.[18]

Consider a concrete case: suppose someone thinks both that it's 80% likely to rain tomorrow and that it's 80% likely not to rain tomorrow, i.e. she holds the following two probabilistic beliefs:

$$P_R = \{p \mid p(\mathit{Rain}) = 0.8\}$$
$$P_N = \{p \mid p(\neg\mathit{Rain}) = 0.8\}$$

Her beliefs are inconsistent, and the consistency norm tells us that she is thus irrational. But what is it that makes her irrational? Why is it bad for her to have such inconsistent probabilistic beliefs? Here is one answer: she is Dutch bookable. Consider the gambles with the following pay-outs:

	Rain	No rain
G_R	90p	−£1.10
G_N	−£1.10	90p
Total	−20p	−20p

[17] Moss might maintain infinite consistency by expanding the underlying space to include infinitesimal probabilities. The point, though, is that we ought to provide rigorous justifications for coherence norms like the Consistency Norm.

[18] Whilst we think that the accuracy-based approach is ultimately the way that we should go, a lot of work needs to be done to determine how accuracy applies when your total doxastic state is modelled by beliefs towards probabilistic contents. This will combine work on accuracy for imprecise probability Konek (ms) with accuracy for full-beliefs (Easwaran and Fitelson, 2015). So, whilst we think that that is a good way of going, we leave it for future work.

Her belief in P_R makes her desire G_R, and her belief in P_N makes her desire G_N, but taken together, this would result in a guaranteed loss. This sort of phenomena is, it is claimed in the Dutch book discussions, enough to show irrationality.

We can summarise our attempted argument for the Consistency Norm as follows:

- Definition: For a gamble G, a bounded function from Ω to $\mathbb{R}$, if you believe probabilistic contents P and every $p \in P$ has $\text{Exp}_p[G] > 0$, then G is **desirable**.
- Definition: Someone is **Dutch bookable** if there is a finite collection of gambles $G_1 \ldots G_n$, such that each G_i is desirable, but the combined result of taking all the G_i is a guaranteed loss; i.e. for all $w \in \Omega$, $\sum_i G_i(w) < 0$.
- Premise: It is irrational to be Dutch bookable.
- Desired Theorem: If someone has inconsistent probabilistic beliefs, they are Dutch bookable. [This has to be slightly weakened]

∴ Conclusion: The Consistency Norm. It is irrational to have beliefs whose probabilistic contents are inconsistent.
[Given the weakening of Theorem, our conclusion will be slightly weaker]

Whilst there are debates in the literature about the plausibility of the premise linking Dutch bookability to irrationality, we nonetheless think it is progress if we are able to provide a Dutch book argument for the consistency norm.

This argument still hangs on what we labelled as a 'desired theorem': inconsistency entails Dutch bookability. In fact this this theorem does not hold in general. But it does hold in the special case where finitely many of your probabilistic beliefs in closed and convex sets of probabilities are inconsistent. More generally:

Theorem 1. *If you have some finite collection of probabilistic beliefs whose closed convex hulls are jointly inconsistent, then you are Dutch bookable.*[19]

And conversely, if there is no such finite inconsistency in the closed convex hulls of your beliefs, then you are not Dutch bookable.

This will provide an argument for a norm that is weaker than the consistency norm. For example, suppose you believe that two coins are probabilistically independent whilst simultaneously believing that they're probabilistically dependent. Your probabilistic beliefs have inconsistent probabilistic contents, but this isn't enough to give Dutch bookability because neither of these beliefs commits you to any non-trivial desirability of gambles.

We might be able to obtain something slightly stronger by weakening our notion of Dutch bookability. For example, it is plausibly irrational to judge a collection of gambles which are guaranteed to pay out the status quo to be (strictly) desirable. Suppose that you believe that your train is both *more* likely than not to be on time (probability $> 1/2$) and also *less* likely than not to be on time (probability $< 1/2$). Then you are Dutch bookable in this weaker sense.

[19] The closed convex hull of a set of probabilities is the smallest closed and convex set containing it.

Whatever norm this weaker notion of Dutch bookability vindicates, it will fall short of the full consistency norm that Moss identifies due to cases like the two coin probabilistic dependence example above.[20]

The results here represent significant progress towards justifying the Consistency Norm. We have shown that a large class of failures of consistency result in Dutch bookability, and can thus be deemed irrational.

4 Conclusion

We have argued that Moss's framework provides a very natural and expressively powerful model of uncertainty. It is at least as general as sets of desirable gamble sets (or equivalently choice/rejection functions)—the most general model of uncertainty currently under investigation in the imprecise probability community. We have also made some progress towards justifying Moss's Consistency Norm. We have shown that a large range of failures of consistency will lead you to be Dutch bookable. In particular, these failures are when the closed and convex hull of your beliefs are finitely inconsistent.

References

Border. Separating hyperplanes theorems, 2010.

Jasper De Bock and Gert de Cooman. A desirability-based axiomatisation for coherent choice functions. In S. Destercke, T. Denoeux, M.A. Gil, P. Grzergorzewski, and O. Hryniewicz, editors, *Uncertainty Modelling in Data Science: Proceedings of SMPS*, pages 46–53. Springer, 2018.

Kenny Easwaran and Branden Fitelson. Accuracy, coherence, and evidence. *Oxford studies in epistemology*, 5:61–96, 2015.

James M Joyce. A nonpragmatic vindication of probabilism. *Philosophy of Science*, 65(4):575–603, 1998.

James M Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In Franz Huber and Christoph Schmidt-Petri, editors, *Degrees of Belief*, volume 342. Springer, Dordrecht, 2009.

Jason Konek. Comparative probabilities. In J Weisberg and R Pettigrew, editors, *The Open Handbook of Formal Epistemology*, pages 267–348. Philpapers, 2019.

Jason Konek. Accuracy and imprecise probability. ms.

Anna Mahtani. Imprecise probabilities. In J Weisberg and R Pettigrew, editors, *The Open Handbook of Formal Epistemology*, pages 107–130. Philpapers, 2019.

Sarah Moss. *Probabilistic knowledge*. Oxford University Press, 2018.

Richard Pettigrew. *The Dutch Book Argument*. Cambridge University Press, 2019.

[20] We conjecture that if for each probabilistic belief we look at the largest set of probabilities with the same resultant (strictly) desirable gambles, and that these are finitely inconsistent, then the probabilistic belief state in question is Dutch bookable in this weaker sense.

Erik Quaeghebeur, Gert de Cooman, and Filip Hermans. Accept & reject statement-based uncertainty models. *International Journal of Approximate Reasoning*, 57:69–102, 2015.

T. Seidenfeld, M.J. Schervish, and J.B. Kadane. Coherent choice functions under uncertainty. *Synthese*, 172(1):157–176, 2010.

Michael Titelbaum. Precise credences. In J Weisberg and R Pettigrew, editors, *The Open Handbook of Formal Epistemology*, pages 1–56. Philpapers, 2019.

Arthur Van Camp and Enrique Miranda. Irrelevant natural extension for choice functions. *Proceedings of Machine Learning Research*, 103:414–423, 2019.

Peter Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, New York, 1991.

Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2-3):125–148, 2000.

A Proof of the Dutch book result

We assume that Ω is finite.[21]

Theorem 2. *If you have some finite collection of probabilistic beliefs in closed and convex sets of probabilities which are jointly inconsistent, then you are Dutch bookable.*

Proof. We consider probabilities as vectors in $\mathbb{R}^\Omega$, i.e. $p : \Omega \to \mathbb{R}$,[22] and use norms and dot-product as expected.

Lemma 3. *We can choose some e so that for b_i the closest in P_i to e, $e = \frac{\sum_i b_i}{n}$.*[23] *Moreover, by possibly ignoring some of the initial P_is, we can do this so that $e \notin P_i$ for any P_i.*

Proof. Note first, that for any collection $b_1, \ldots, b_n$, $\sum_{i=1}^n \frac{\|e-b_i\|^2}{n}$ is minimized by $e = \frac{\sum_i b_i}{n}$. But we cannot simply choose our b_i and define e this way because we needed to show that e is in an average of the b_i which are chosen as the members of P_i minimizing the distance between e and itself.

So, we specify e to be what minimizes $\sum_i \|e - \mathsf{closest}(e, P_i)\|^2$, where $\mathsf{closest}(e, P_i) = \arg\min_{b_i \in P_i} \|b_i - e\|$. Since P_i is convex and closed, there will be a unique such minimum. We can then show that taking $b_i = \mathsf{closest}(e, P_i)$, we have $e = \frac{\sum_i b_i}{n}$.

To ensure that $e \notin P_i$, we note that we can delete any P_j with $e \in P_j$ from consideration and we will still have e is the average of the remaining b_i because if $e \in P_j$ then $e = b_j$.[24] □

[21] This is for mathematical convenience, the results might be generalisable. Note it suffices to assume $\mathcal{F}$ is finite and take Ω to be the atoms of the Boolean algebra generated by $\mathcal{F}$.

[22] This determines p defined on $\mathcal{F}$ by the assumption that p is probabilistic and that Ω is finite.

[23] Our result would still work if e is a weighted average of the b_i by appropriately weighting each G_i in the final collection of gambles under consideration.

[24] And then $e = \frac{\sum_i b_i}{n} = \frac{\sum_{i \neq j} b_i}{n-1}$.

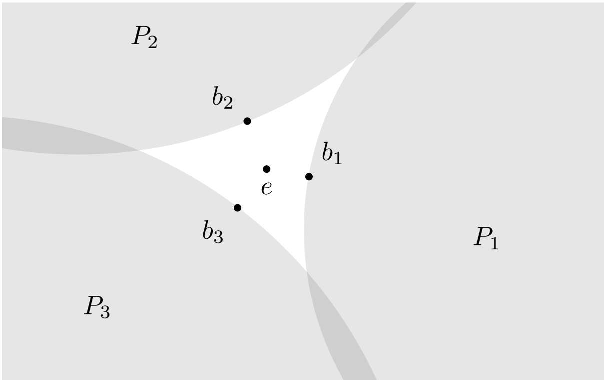

Figure 1: Illustration of our choice of b_i and e

Assume we've deleted the P_i from consideration where $e \in P_i$. Now focus on an individual probabilistic content P_i.

Lemma 4. *There is some* $\epsilon_i > 0$ *such that for each* $p \in P_i$, $\text{Exp}_p[b_i - e] > \text{Exp}_e[b_i - e] + \epsilon_i$. *I.e. the gamble*

$$G_i = (b_i - e) - (\text{Exp}_e[b_i - e] + \epsilon_i)$$

is desirable.

Proof. $\{e\}$ and P_i are disjoint convex sets. So we can use the strong hyperplane separation theorem (see Border, 2010). This gives us some linear functional F_i, in fact $F_i = b_i - e$, with some $\epsilon_i > 0$ where:

$$p \cdot F_i > e \cdot F_i + \epsilon_i \text{ for all } p \in P_i.$$

Thus,[25] $\text{Exp}_p[b_i - e] > \text{Exp}_e[b_i - e] + \epsilon_i$ for all $p \in P_i$, as required. □

Lemma 5. *The collection of gambles* $G_i = (b_i - e) - (\text{Exp}_e[b_i - e] + \epsilon_i)$ *when taken together lead to a guaranteed loss.*

Proof. By recalling that $e = \frac{\sum_i b_i}{n}$ we see that $\sum_i (b_i(w) - e(w)) = 0$; but that $\sum_i (\text{Exp}_e[b_i - e] + \epsilon_i) = \sum_i \epsilon_i > 0$. □

This suffices for theorem 2. □

Theorem (1)**.** *If you have some finite collection of probabilistic beliefs whose closed convex hulls are jointly inconsistent, then you are Dutch bookable.*

And conversely, if there is no such finite inconsistency in the closed convex hulls of your beliefs, then you are not Dutch bookable.

Proof. The result that such inconsistency entails Dutch bookability follows immediately from theorem 2 by noting that if someone believes P, with $Q \supseteq P$, and Q desires G, then so does P.

[25] As $\text{Exp}_p[F] = \sum_w p(w)F(w) = p \cdot F$.

We now prove the converse result: Consider a collection of gambles $G_1, \ldots G_n$ each of which is desirable. This desirability comes from some P_i. So $\text{Exp}_p[G_i] > 0$ for all $p \in P_i$. Thus, $\text{Exp}_p[G_i] \geqslant 0$ for all $p \in Q_i$, where Q_i is the smallest closed and convex set of probabilities extending P_i. If $Q_1, \ldots, Q_n$ is consistent, we can take some $p^* \in Q_1 \cap \ldots \cap Q_n$. So, $\text{Exp}_{p^*} G_i \geqslant 0$ for all i; and thus $\text{Exp}_{p^*} \sum_i G_i = \sum_i \text{Exp}_{p^*} G_i \geqslant 0$. So it cannot be that $\sum_i G_i(w) < 0$ for all $w \in \Omega$ as p^*'s expectation of a guaranteed loss would be negative. □