Please check for updates (here) before citing.

Draft October 11, 2019

Risk and Accuracy: Prospects for an Alethic Epistemology of Risk-Sensitive Partial Belief

Catrin Campbell-Moore, Ben Levinstein*

October 11, 2019

1 Introduction

It is well known that expected utility theory doesn't capture many decisions that people actually take. In particular, people often take account risk in a way that is impermissible according to expected utility theory. Lara Buchak (2013) has recently argued that such risk aware behaviour is in fact rational and has developed a sophisticated decision theory to account for this.

In Buchak's decision theory, agents have, in addition to their utilities and degrees of belief, a *risk profile*. This risk profile affects how much the utilities of various outcomes affect the agent's evaluation based on how risky the outcomes are.

In this paper we investigate an epistemic analog of Buchak's account of practical rationality. In particular we investigate whether there is a notion of 'accuracy', or closeness-to-truth, that can be made to work with Buchak's account.

Considerations of accuracy have been used to justify a whole host of rationality constraints, the most prominent of which is probabilism: the thesis that one's degrees of beliefs should satisfy the axioms of probability. It would be excellent for Buchak if she could justify probabilism through appeal to accuracy and potentially worrying for her account if she cannot. In other words, we wish to know whether or under what conditions risk-aware agents who care about accuracy should have credences that obey the axioms of probability theory.

In both the risk-neutral and risk-aware cases, measuring accuracy, or closeness-to-truth, turns out to be more subtle than one initially might have thought. In both cases, the accuracy-firsters want that acceptable measures of accuracy are suited to underwrite a dominance argument for probabilism. That is, by appeal to such measures, we must be able to *derive* that any non-probability function is sure to be less accurate than some probability function, and every probability function itself is always more accurate than any other function at some world. Ideally, the constraints on measures of accuracy that allow one to derive dominance will be independently motivated. Dominance, however, is not enough.

*Many thanks to Branden Fitelson, Bernhard Salow, Jason Konek, Richard Pettigrew, Kenny Easwaran.

In the case of risk-neutrality, every satisfactory measure of accuracy must also be *strictly proper*, meaning each probability function expects itself to be more accurate than any other. By using strictly proper scoring rules, risk-neutral rational agents satisfy a kind of *immodesty* criterion. Supposing that the agent only cares ultimately about how accurate she is, she thinks her own credence has the most epistemic value in expectation.

However, strict propriety is defined based on *expected* accuracy. It is only useful for agents who are risk-neutral and make epistemic decisions based on expected utility (or accuracy) maximisation. If risk-aversion is rational, then we also want these risk-averse agents to be immodest. But they evaluate credences not by expected but by risk-weighted expected utility.

If we are to maintain immodesty for rational agents and allow that risk-aversion is rational, we will have to find measures of accuracy where probabilistic credences will assign themselves highest risk-weighted expected accuracy, and which also allows us to derive the desired dominance result.

Our primary question is whether there are any risk-sensitive proper measures of accuracy at all. That is, are there any measures of accuracy such that any probabilistic credence will assign itself highest risk-weighted expected accuracy?

We can generalise the standard proper measures to risk-sensitive ones for *individual* credences in single propositions. That is, if we want to give a local score to a single credence, we can naturally morph normal proper measures into risk-weighted proper ones.

The first bit of bad news, however, is that there's no single measure that is proper for all risk-profiles. The second is that, unlike in the case of risk-neutrality, we cannot simply add up locally proper measures to generate globally proper ones. In other words, there are no measures that are risk-weighted proper for individual credences such that their sum will generate a risk-weighted proper measure for an entire probability function.

The existence of globally proper rules is important for Buchakians. We wish to establish fundamental norms on credences in a way that accuracy-first epistemologists have done in the risk-neutral case. In particular, we wish to justify probabilism for risk-sensitive agents through appeal to accuracy considerations alone.

Unfortunately, this is a difficult task. We will be able to provide globally risk-sensitive proper measures and show that they allows us to derive dominance, leading us some way along an argument for probabilism. However, the measures we provide are not truth directed. They are thus not legitimate measures of closeness to truth, or accuracy. So to successfully provide an accuracy first foundation for Buchakean epistemology more needs to be done. However we hope that this paves the way for the construction of legitimate measures.

2 Risk aware decision theory

In normal (risk-neutral) decision-theory, to estimate how good an action will be, we take the various possible outcomes and just add up the goodness of each of those outcomes weighted by how likely our agent thinks it is that that outcome obtains.

$$\mathrm{Exp}_b U(A) = \sum_{\text{outcome } O} U(O) \times b(O \text{ obtains})$$

In risk-aware decision theory our agents think that this evaluation of the goodness of an action is missing an important component which allows agents to undertake worst-case-scenario style reasoning. The risk-averse agent thinks the worst-cases should have more weight attached to them than the best-cases. A risk-seeing agent on the other hand will give more weight to the best cases. This results in a risk-averse agent staying closer to the 'best-worst-case' decision rule, i.e. taking safer actions, whereas the risk-seeing agent will go all-in, seeking to maximise the best-case scenario. This style of reasoning can be accounted for by the risk-aware decision theories. Instead of multiplying the utility of various outcomes by the credence that they obtain, in risk-aware decision theory, the weight that they are given should also account for their position in the utility ordering.

Risk-aware decision theory wants to account both for risk-averse and risk-seeking agents, which it does by associating with agents a 'risk-profile'. Formally, this is a continuous, strictly increasing function $r : [0, 1] \to [0, 1]$ such that $r(0) = 0$ and $r(1) = 1$. Consider an agent with risk profile $r(x) = x^2$. This risk profile will be used with the agents beliefs to determine the weight assigned to each outcome as in section 6.

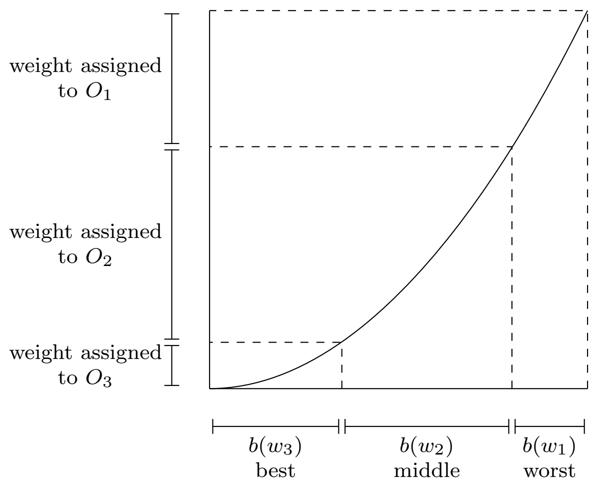

Figure 1: The decision weights for $r(x) = x^2$.

x^2 is a risk-profile that will increase the weight assigned to the worst case.

The risk-weighted expectation can thus be given as:[1]

$$\mathrm{RExp}_b^r U(A) = \sum_{\text{outcome } O} U(O) \times \begin{pmatrix} r(b(\text{something at least as good as } O \text{ obtains})) \\ \hfill -r(b(\text{something better than } O \text{ obtains})) \end{pmatrix}$$

We will summarise this term in brackets as we will use it later.

[1]Our presentation differs from that usually given by Buchak, but follows that of rank-dependent utility theory (Wakker, 2010), and is similar to the presentation of Pettigrew (2015). We chose to do this because this alternative presentation is useful in our technical results. This formulation still works even if there are multiple outcomes that have equal utility (unlike Pettigrew's). See Wakker (2010, Appendix 6.7).

Definition 2.1. For any $c : \mathcal{F} \to \mathbb{R}$,

$$q_c^{<}(w) := r\left(\sum_{v \geqslant w} c(v)\right) - r\left(\sum_{v > w} c(v)\right).$$

And we see that we can concisely write (for probabilistic b)

$$\mathrm{RExp}_b^r U(A) = \mathrm{Exp}_{q_b^{<_U}} U(A).$$

Buchak's theory might be seen as merely descriptive of how agents actually behave. But Buchak argues that her theory is appropriate for normative guidance, i.e. making evaluations according to risk weighted expected utility theory is rational. We will adopt this perspective and assume that there are at least some rational agents who evaluate options according to risk-weighted expected utility theory. We are, however, open to the possibility that this might be rejected, perhaps even on the basis of accuracy considerations. In fact, if there are no accuracy measures that are appropriate for the risk-aware then we would take this as pressure on Buchak's claim that risk-awareness is rationally permissible.

Pettigrew (2015) argued that risk-awareness is irrational because risk-neutrality accuracy dominates it. However Pettigrew uses the familiar risk-neutral accuracy measures, whereas we are more flexible on the way of measuring accuracy. We thus think that his argument would need to be reevaluated once we have properly understood how to think about accuracy for the risk-aware framework.

Now having presented Buchak's decision theory, we move to 'accuracy'.

3 Measuring accuracy

According to accuracy-first epistemology, the higher your credence in truths and the lower your credence in falsehoods the better off you are all epistemic things considered. Accuracy-first epistemology leverages precise measures of accuracy (i.e., proximity of credence function and actual truth-values) to derive important norms on credences, such as probabilism. To execute this project, accuracy-first epistemologists initially place constraints on 'legitimate' measures of accuracy. Such measures function as measures of overall epistemic utility, i.e., how epistemically well-off an agent is in a given world. Accuracy-firsters then exploit norms from decision-theory (such as dominance avoidance) to establish epistemic norms of interest.

The first question, then, is how we can precisely measure accuracy in a way that's adequate for epistemic utility theory. Our accuracy measure will provide for each credence function, c, and world w, a value of accuracy.

In his (1998), Joyce presents a number of constraints on what it is to be a good measure of accuracy. The first is immediate from our informal remarks on accuracy, *viz.*:

Truth-Directedness. If

- For all X, $b(X) \geqslant c(X) \geqslant w(X)$ or $b(X) \leqslant c(X) \leqslant w(X)$
- For some X, $b(X) > c(X) \geqslant w(X)$ or $b(X) < c(X) \leqslant w(X)$

Then $\mathsf{Acc}(b, w) < \mathsf{Acc}(c, w)$.

However, more constraints than mere Truth-Directedness are required to ensure measures of accuracy do the work the accuracy-firster requires. One key goal of accuracy-first epistemology is to vindicate the fundamental norm of *probabilism*. To show this, the accuracy-firsters want to show that for any legitimate way of measuring accuracy, probability functions turn out to be required to successfully pursue accuracy. More precisely, they want:

Dominance. If b is a probability function, then for any credence function c, there exists a world where b is more accurate than c. If b is not a probability function, then there is a probability function c such that c is more accurate than b at every world.

Unfortunately truth-directedness isn't sufficient for this dominance property. There are some measures of accuracy that are truth-directed but allow some non-probability functions to be non-dominated.[2] The accuracy-firsters thus need to justify independently motivated additional constraints on what it is to be a legitimate measure of accuracy, from which dominance can be derived.

What difference does this make for the risk-aware agent? Not much. Dominance results are agreed on by expected utility theory and Buchak's risk weighted expected utility theory alike. So the desire to have dominance as a consequence of a legitimate measure of accuracy doesn't introduce any particularly new problems for the risk-aware.

However, there is another constraint on legitimate measures of accuracy which does apply differently once risk is taken into the picture. The accuracy-firster will need to show that coherent agents think that their own credences are epistemically best according to any legitimate measure. Otherwise, even though they are probabilistically coherent, they would not rationally endorse their own views. Thus, their own probabilistically coherent views would be ruled out as irrational by their own lights no matter what evidence they possessed. In other words, coherent agents must be *immodest*.

In the risk-neutral case, agents evaluate their own credences in terms of expected accuracy. To turn our immodesty criterion into a formal constraint, we then require legitimate measures satisfy:

Strict Propriety. For all probabilistic b, $\mathrm{Exp}_b\mathsf{Acc}(c)$ is (uniquely) maximised at $c = b$.

Indeed, it turns out that if we can justify constraints on accuracy measures that get us strict-propriety, then dominance follows almost for free. I.e. dominance is a *consequence* of strict-propriety, truth-directedness, and some other rather innocuous assumptions.[3]

[2]Consider, for example, the absolute value measure where we simply add up the difference between one's credences and the truth values across the various propositions.

Suppose there is an urn with three balls in it, and it's known that only one will be drawn, each with a $^1/_3$ chance. Let the agenda be: {Ball 1 is drawn, Ball 2 is drawn, Ball 3 is drawn}. If Alice has credence $^1/_3$ in each proposition, she's guaranteed an overall absolute value accuracy score of $-^2/_3 - ^1/_3 - ^1/_3 = -^4/_3$. If Bob has a credence of 0 in each proposition, he's guaranteed an overall score of score of $-1 - 0 - 0 = -1$. So, Bob is guaranteed to be more accurate than Alice even though she's probabilistically coherent.

[3]The additional assumptions he needs are: continuity and boundedness of the measure. In fact Joyce relies on coherent admissibility rather than strict propriety, but coherent admissibility follows from strict propriety.

However, Propriety appeals to standard expectations. We have assumed that risk-awareness is rational. And risk aware agents will evaluate the accuracy of their doxastic state not by its *expected accuracy* but by its *risk-weighted expected accuracy.*

If Buchakeans hope to vindicate probabilism, then, they need their legitimate measures to satisfy:

Strict r-Propriety. For all probabilistic b, $\mathrm{RExp}^r_b\mathsf{Acc}(c)$ is (uniquely) maximised at $c = b$.

Otherwise, they face the same problem where there are probabilistic credences that think some other credence function is better. For the accuracy firster who allows for agents who are rationally risk aware when considering epistemic matters, she will need to provide constraints on legitimate ways of measuring accuracy which ensure that they are r-proper and have the dominance property.

As Pettigrew (2016, §16.4) points out, we cannot simply use our (risk-neutrally) strictly proper scoring rules. In many cases, perhaps all, strictly proper scoring rules will not be strictly r-proper.[4] We thus need to consider whole new kinds of ways of measuring accuracy.

It may appear odd that we would change measures of accuracy merely because of an agent's attitude toward risk. How close one's credence is to the truth should not, it seems, depend on how skittish one is about pursuing truth. Indeed, in standard Buchakean decision theory, the utility function and the risk function are orthogonal. One can vary without affecting the other.

We have two replies. First, this paper is meant to explore Buchak's prospects for a risk-aware, accuracy-first epistemology. Since no measure of accuracy is r-proper for all r, the best we can do is tailor the measures of accuracy based on the risk function. Second, in the risk-neutral case, measures of accuracy are ultimately deemed acceptable only if they are proper.[5] Although many (e.g., Joyce 1998, 2009; Pettigrew 2016) have attempted to impose pre-theoretically acceptable constraints on accuracy measures that entail propriety, at least some of these constraints are backwards-engineered. We leave it to those partial or averse to Buchakian decision theory to determine whether tailoring the utility function to the risk function is acceptable in our case, but if risk-aware accuracy-first epistemology is to succeed, such tailoring is required.

Before beginning our search for r-proper measures, we observe that one could object to the very idea of risk-sensitivity in the epistemic case. That is, one may claim that even supposing that risk-awareness is rational in the instrumental domain, it is not permissible in the epistemic domain. The objector claims: when evaluating the epistemic value of their credences our agents should simply do expected accuracy considerations. We have nothing decisive to say against this option; we just think it is a rather odd picture.

Buchak thinks that risk does not change the value of outcomes, but instead that it is relevant additional factor to credence and utility in the rational pursuit of one's ends. In turn, risk-aware agents generally identify value with risk-weighted expectation. Given the general structure of accuracy-first epistemology, whereby we use decision-theoretic concepts to pursue epistemic value, it would be strange to treat the epistemic case so distinctly from the practical one.

[4] Pettigrew's argument that strictly proper cannot be r-proper doesn't quite work in general. However, the argument is successful if r is such that $r(x) \neq x$ for all (or most) x.

[5] See, e.g. the absolute value measure (footnote 2).

Furthermore, it would be particularly hard for Buchak to reject this claim as she thinks that the risk profile is a single fixed subjective component which is to be used in all domains.

Finally, we think there are independent reasons to develop a risk-aware alethic epistemology. As William James famously said:

> There are two ways of looking at our duty in the matter of opinion.[...] We must know the truth; and we must avoid error, — these are our first and great commandments as would be knowers; but they are not two ways of stating an identical commandment [...] Believe truth! Shun error! — these, we see, are two materially different laws; and by choosing between them we may end by coloring differently our whole intellectual life. We may regard the chase for truth as paramount, and the avoidance of error as secondary; or we may, on the other hand, treat the avoidance of error as more imperative, and let truth take its chance. (1896, §VII)

These commandments are fully alethic, in that they only concern truth and falsity. In the case of full belief, it is easy to satisfy either commandment if one disregards the other. To satisfy *Believe truth!*, one may simply believe every proposition—i.e., always believe X and also $\neg X$. To *Avoid error!*, one may suspend judgment about everything. Rationality requires some way of balancing of these exhortations.

When moving to partial belief, each individual credence strikes a balance between the two commandments. A credence of .7 gives up a chance at perfect accuracy (*Believe truth!*) to ensure that one does not violate the other commandment too severely. Each credence, then, manages risk of error against pursuit of truth.

A risk-aware epistemology may allow us new ways to balance the two commandments. Some agents may care more about avoiding error than they do about believing truth, which results in them caring more about worst-case scenarios, i.e. cases where they will end up with the least accurate credences. A risk-aware epistemology permits agents to be sensitive to questions of epistemic risk that standard accuracy first epistemology is not.

So to apply the accuracy framework for Buchak we need to find entirely new measures of accuracy that are strictly r-proper and then see what happens when we try to pursue these. Unfortunately, finding such measures turns out to be rather more difficult that one might initially have hoped. When just a single proposition is considered we know that they will exist; i.e., we can find local strictly r-proper accuracy measures. But the situation becomes more complicated when we consider the agent's entire credence function. Unfortunately we cannot simply add up the local accuracy measures to come up with a global strictly r-proper accuracy measure. We will, however, show that there is still some hope for finding global strictly r-proper accuracy measures by providing some such measures. However, since they are not truth directed they don't really deserve the term "accuracy measure". They do, however allow us to derive a dominance result. This leaves the risk-aware accuracy firster with some hope of applying the accuracy first project to the risk aware.

4 Local r-proper measures

To start our discussion we will focus on a simpler case. Suppose there is only one proposition at stake, and we want to measure how accurate an agent's credence is in this single proposition.

A *local accuracy measure* is some function $\mathfrak{a} : [0,1] \times \{0,1\} \to \mathbb{R}$, which is truth-directed, i.e.

Truth-Directedness (local). If $x > y$ then $\mathfrak{a}(x,1) > \mathfrak{a}(y,1)$ and $\mathfrak{a}(x,0) < \mathfrak{a}(y,0)$.

We say $\mathfrak{a}$ is strictly r-proper if:

r-Propriety (local). For all probabilistic b, $\mathrm{RExp}_b^r \mathfrak{a}(y)$ is uniquely maximised at $y = b(X)$.

There are three reasons that one might be interested in local r-proper measures. The first is that they are the standard way of generating *global* measures, which take into account the entire doxastic state of an agent. This is usually done by taking the global measure to be a sum of the accuracy in the various propositions. However, we will discuss in section 5 that this strategy doesn't work in the risk-aware setting.

The second reason one might be interested in such local measures is that they are useful if we are interested in the practical benefits of strictly proper accuracy measures: they allow one to elicit beliefs. Using strictly proper accuracy measures we can give a decision problem to an agent and ensure that they report their actual degree of belief. If we give these decisions to the risk-aware agents we will need the scoring rules to be strictly r-proper to ensure that they report their degree of belief. We are often interested in doing this just with respect to a single proposition, such as if there'll be rain tomorrow, and thus are interested in local strictly r-proper accuracy measures. However, it's worth us pointing out that if that's all that you're interested in from the accuracy measures, instead of the additional epistemic benefits such as arguments for probabilism, Kothiyal et al. (2010) already provides measures which allow us to elicit the risk-aware agents credences. Using those measures the risk-aware agents aren't encouraged to directly report their credence, but by knowing what their risk-profile is the elicitor can recover their credence.

Third, measures of local accuracy are important for establishing certain epistemic norms. For instance, Campbell-Moore and Salow (msa) use local measures to justify constraints on gathering evidence.

We can find such local strictly r-proper accuracy measures by generalising the usual representation theorem of Savage. We leave the technical details to the appendix. For this construction we are indebted to Adam Bjorndahl. Roughly, the formal idea is to incorporate the risk-function and credence into Savage's formula for deriving proper scoring rules.

Here we just give an example. This example is the measure that Campbell-Moore and Salow (msa) introduce which is strictly r-proper for $r(x) = x^2$.

$$\mathfrak{a}(y,v) := \frac{-(v-y)^2}{\max\{y, 1-y\}}$$

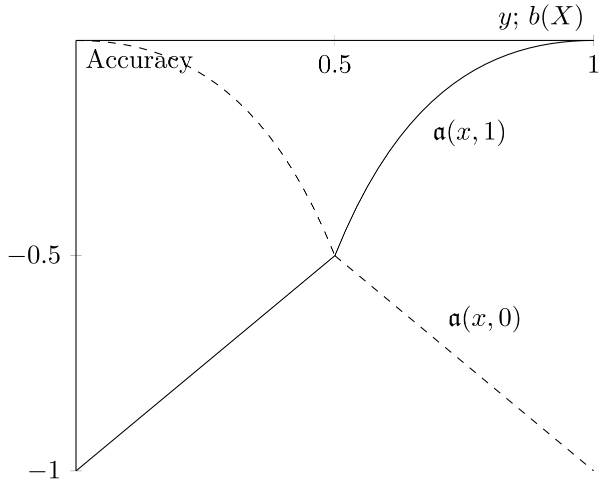

Figure 2: A local r-proper accuracy measure

5 Global Measures and Additivity

Despite their usefulness, local measures can't be used for justifying probabilism: that necessarily involves quantifying an agent's total accuracy over multiple propositions.

So we really need a global measure of accuracy to perform the roles that we want of our accuracy measures.

In the risk-neutral accuracy domain, we usually define global accuracy measures directly from our local ones. We let $\mathsf{Acc}(c,w) := \sum_{X\in\mathcal{F}} \mathfrak{a}(c(X), w(X))$. And so long as our local accuracy measures were strictly proper, the global measures thus defined will also be strictly proper.[6]

It will not be possible for us to apply the analogous trick to find r-proper global measures.

Theorem 5.1. *If* $\mathfrak{a}$ *is strictly* r*-proper, continuous local accuracy measure and* Acc *is the additive global measure defined from it, then* Acc *is not strictly* r*-proper.*

The key reason for this is that when being risk aware, global features become very important.

This idea that global measures are constructed by adding up the local measures is more than simply a mathematically convenient way of creating globally proper rules. An additivity constraint on accuracy measures is often imposed as an axiom on legitimate measures.

Pettigrew (2016, §4.1), for instance, justifies additivity as follows

> When we say that we represent an agent by her credence function, it can sound as if we're representing her as having a single, unified doxastic state. But that's not what's going on. Really, we are just representing her as having an agglomeration of individual doxastic

[6]This argument works because Exp is linear: $\mathrm{Exp}_b \sum_{X\in\mathcal{F}} \mathfrak{a}(c(X), w(X)) = \sum_{X\in\mathcal{F}} \mathrm{Exp}_b \mathfrak{a}(c(X), w(X))$. RExp, however, is not linear.

states, namely, the individual credences she assigns to the various propositions about which she has an opinion. A credence function is simply a mathematical way of representing this agglomeration; it is a way of collecting together these individual credences into a single object.

However, accepting that a credence is agglomeration of various bits doesn't immediately get us that their accuracy should be judged just by adding up the accuracy values on the various propositions. Pettigrew goes on to say

> Of course, there are practical cases in which the outcome of an option is an agglomeration of commodities yet we don't take its value to be simply the sum of the values of the commodities. These are cases in which some of the commodities in question are what economists call dependent goods. A dependent good is a good such that the value that it contributes to the overall value of a commodity bundle of which it is a part depends upon the other commodities in the bundle. For instance, the value contributed to an outcome by a tin of beans depends on whether or not the outcome also includes a tin opener and a stove; the outcome contributed to an outcome by a television remote control depends on whether or not the outcome also includes a television that it can operate; and so on. Therefore, one consequence of [Additivity] is that accuracy is not a dependent good in this sense.

What we learn from our above result is that accuracy must be a dependent good. However, that is very natural in the Buchakean setting. If an agent starts with one credence, then she balances the risk of error against the pursuit of truth. If she then adds a second credence to her agenda, her overall risk-profile changes, as there are now new outcomes available. She could be fairly accurate with respect to both propositions individually, accurate with respect to one and inaccurate with respect to the other, or inaccurate with respect to both. The stakes have changed. Since the risk-profile and since the accuracy measure are not, sadly, orthogonal in the epistemic case, the accuracy measure must change as the risk-profile does.

6 Finding global measures

We will show that we can find some global strictly r-proper accuracy measures. Indeed for every risk profile we can find some such measures. However, the ones that we find will conflict with other requirements one might want to impose on accuracy measures. In particular our accuracy measures that we construct here are not truth directed. They are thus not measures of closeness to truth in any sense, so we should really not use the term 'accuracy measure'. However, for convenience we will continue to do so.

Whereas we show that there are no additive r-proper accuracy measures there is no analogous argument for truth directed r-proper accuracy measures. Indeed we conjecture that they do exist, but we have not yet found any.

We have, however, found *some* accuracy measures. We hopes this paves the way for future work to find more plausible accuracy measures. In the main text, we run through a method for constructing such measures when credence functions are defined over a partition. In appendix C we will show how to generalise these methods to arbitrary agendas.

The measures that we have found start with what we call a *God-given ordering* on worlds. Some worlds are counted as automatically epistemically better than others according to this arbitrary ordering—no matter what your credence function is, you're more accurate in some worlds than in others. Any ordering will do, but it must stay fixed.

Let's consider an example. Suppose there's two possibilities, either it rains or it doesn't. We have a God-given ordering on worlds, so one of these worlds is top. Let's say it's the rain world. Then whatever credence you have, e.g. $b(\text{Rain}) = 0$, you're still epistemically better if it does in fact rain. Epistemically you should hope that it rains, as that will give you more epistemic points, even if you're maximally confident that it will not rain.

The r-proper accuracy measures that we construct are based on our usual accuracy risk-neutral accuracy measures. To construct the risk variant of the accuracy measure we measure the risk-neutral accuracy not of the credence function itself, but of a canonical alternative: the decision weights that are used in a decision problem where the utilities are ordered according to the God-given ordering.

We start with our credence function, c, and transform it into a new function $q_c^<$, based on our God-given ordering, $<$:

$$q_c^<(w) := r\left(\sum_{v \geqslant w} c(v)\right) - r\left(\sum_{v > w} c(v)\right).$$

These are the decision weights that the agent with risk profile r uses in any decision problem where the utilities are ordered according to $<$. We will ensure that Acc orders the worlds according to $<$, and we will thus have

$$\text{RExp}_b^r \mathsf{Acc}(c) = \text{Exp}_{q_b^<} \mathsf{Acc}(c)$$

We can now take a first-stab at our new accuracy measures by using an old-fashioned risk-neutral accuracy measure of the decision weights, e.g. with the Brier score we'd look at taking $\mathsf{Acc}(c, w) = \text{BS}(q_c^<, w)$.

This isn't quite enough for our general accuracy measure, though, because this does not order the worlds according to the God-given ordering. We thus simply add some constants to make sure it does so and we define

$$\mathsf{Acc}(c, w) := \text{BS}(q_c^<, w) + k_w.$$

The constants k_w are chosen so that we have $w < w'$ (according to the God-given ordering) iff $\mathsf{Acc}(c, w) < \mathsf{Acc}(c, w')$.

Using the strict propriety of BS, the equivalence between RExp_b^r and Exp_{q_b} and the fact that $q_c^<$ is a probability if and only if q is, we will be able to show that Acc is strictly r-proper.

Theorem 6.1. *If B is strictly proper and k_w are some constants then*

$$\mathsf{Acc}(c, w) := B(q_c^<, w) + k_w$$

is strictly r-proper so long as the constants are chosen so as to obtain the right God-given ordering $<$.

This can also be easily expanded to deal with the whole algebra by expanding our definition of q_c as spelled out in appendix C

These measures also support a dominance argument for probabilism. I.e. we can show that

- If c is a credence function that is probabilistically incoherent, then there exists a probability function b such that b is strictly more accurate at every world than c.
- If c is a probability function, then there's no alternative credence function that's at least as accurate as c is at every world.

So we have constructed accuracy measures which are strictly r-proper and which support dominance. However, they are not truth directed and so are not legitimate measures of closeness to truth.

Proposition 6.2. Acc *as constructed in theorem 6.1 are not necessarily truth-directed.*

Proof. Suppose our agenda just contains the two propositions p and $\neg p$. And suppose we have a God given ordering which sets p as the best case. If p is false, then moving credence in p lower should get more accurate, by truth directedness.

Consider decreasing $c(p)$ while keeping $c(\neg p)$ fixed. Note then that $q_c^<(p) = c(p)$ decreases, getting more Brier Score accuracy-points, the decrease in p has also altered $q_c^<(\neg p) = r(c(p) + c(\neg p)) - r(c(p))$. Indeed if r is convex, then a fixed increase of $b(\neg p)$ will matter more the worse off it starts; i.e. $r(b(\neg p) + b(p)) - r(b(p)) > r(b(\neg p) + c(p)) - r(c(p))$. So $q_c^<(\neg p) < q_b^<(\neg p)$; and indeed p is true so making $q_c^<(\neg p)$ smaller results in a Brier score dis-improvement. And it might be that this disimprovement outweights the p improvement.

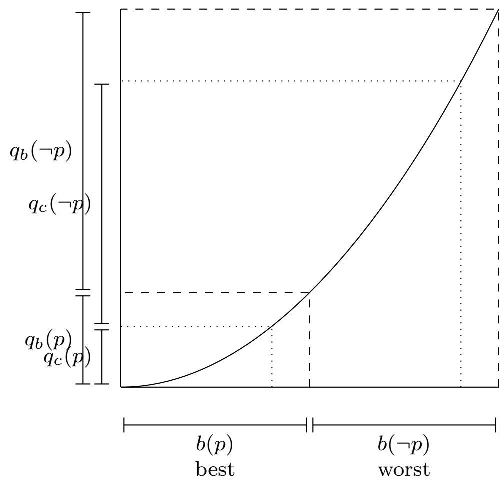

Figure 3: The decision weights for $r(x) = x^2$.

Indeed, this sometimes happens with the Brier score, e.g. $(.5, .5)$ is more accurate than $(.2, .5)$, see `https://www.desmos.com/calculator/gnuvngubtc`. □

Even if they were truth directed, there is another principle that we think is desirable: Extensionality, which roughly says that each world is treated equally. The reason is that we started with an initial God-given ordering.

And furthermore, we can find examples where even over probability functions it's not truth-directed `https://www.desmos.com/calculator/rmf9bzyl0r`

7 Conclusion

So we've managed to give global accuracy measures that are appropriate for the risk-sensitive agent, i.e. which are r-proper. And we have given a dominance argument for probabilism using these measures. So have we given Buchak an epistemic foundation for her agents and a justification of why their beliefs should be probabilistic? Ultimately, no, but we hope to have paved the way for such an argument.

Of course, one might be sceptical about this providing a justification for probabilism for the usual reason that in assuming propriety we essentially already assumed that probability functions were undominated. Whilst that is right, it is what one would need for accuracy measures that might justify probabilism. And as has been done in the risk-neutral case we would now hope that one can replace the r-propriety axiom with alternative directly motivated principles, such as a risk-appropriate version of Pettigrew (2016)'s Decomposition axiom. Though we have no idea what this might look like, we hope it will be useful in future work.

The main reason, though, that we think our story is not a complete success is that the global accuracy measures that we have found are not legitimate measures of closeness to truth: they fail to be truth-directed.

So, we've shown, on the one hand, that it is possible to have risk-sensitive global measures of accuracy that are r-proper and that provide a dominance argument for probabilism. On the other hand, we've shown no additive measures exist and noted that measures of accuracy must be tailored to the risk-function if risk-aware accuracy first epistemology is to succeed. So, while much room for improvement on our measures, we also know that we can't get everything we might have wanted. We've established both new 'upper' and 'lower' bounds on what's achievable for accuracy-first Buchakians.

8 Further work – Other norms

One key advantage of the accuracy measures that we have found is that they are rather easy to work with.

We can show that a Greaves and Wallace style argument will get us to $b_{q_c(\cdot|E)}$ which is not conditionalization. It's conditionalizing the decision weights then 'removing' the r. Note that this depends on the chosen ordering. Campbell-Moore and Salow (msb) argue that this is not what matters, though, for arguments for how one should *actually* update. It's only relevant for the planning.

And note that might lead to a risk-weighted reflection principle: she thinks her current credences are a risk weighted average of her future credences.

Any chance of other norms? Dunno.

Maybe say something about responding to Pettigrew's accuracy argument for ExpU. I presume with our r-measures we get an accuracy argument for RExpU.

Actually that's probably worth checking. It's not clear because of the bloody orderings. Here's another nice thing about fixing our orderings but having that flexibility to change them: we can then always get a nice link with pragmatic cases by choosing our orderings according to the orderings of the practical problem. So, e.g. epistemic gather evidence = pragmatic gather evidence: pick $<$ so that the epistemic stuff = pragmatic stuff. Actually if we suppose the U fixed in each world, then we can just pick the ordering from the utilities. And then we get nice match of epistemic and pragmatic throughout. Why not??

References

Lara Buchak. *Risk and Rationality.* Oxford University Press Oxford, 2013.

Catrin Campbell-Moore and Bernhard Salow. Avoiding risk and avoiding evidence. msa.

Catrin Campbell-Moore and Bernhard Salow. Rational updating for the risk aware. msb.

William James. *The Will to Believe and Other Essays in Popular Philosophy.* Longmans, Green & Company, 1896.

James Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial beliefs. Springer, 2009.

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

Amit Kothiyal, Vitalie Spinu, and Peter P Wakker. Comonotonic proper scoring rules to measure ambiguity and subjective beliefs. *Journal of Multi-Criteria Decision Analysis*, 17(3-4):101–113, 2010.

Richard Pettigrew. Risk, rationality and expected utility theory. *Canadian Journal of Philosophy*, 45(5-6):798–826, 2015.

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

Peter P Wakker. *Prospect theory: For risk and ambiguity.* Cambridge University Press, 2010.

A Local Accuracy Measures

The quick idea: Do a Savage representation theorem but use the 'decision weights' instead of the credences themselves. Here we'll get 0/1 symmetry so we don't pick a God-given ordering.

Theorem A.1 (Bjorndahl). *Let*

$$q_z^{<_z} := \begin{cases} r(z) & z \geqslant 1/2 \\ 1 - r(1-z) & z < 1/2 \end{cases}$$

Note that this corresponds the the earlier definition of $q_c^<$.

Let g be strictly positive function with $g(z) = g(1-z)$.[7] *Let*

$$\begin{aligned}\mathfrak{a}(y,1) &:= \int_y^1 (1-q_z^{<z})g(z)dz \\ \mathfrak{a}(y,0) &:= \int_0^y q_z^{<z} g(z)dz\end{aligned}$$

Big Proof (Proof of theorem A.1). We note:

$$\begin{aligned}\frac{\partial}{\partial y}\mathfrak{a}(y,1) &:= (1-q_y^{<y})g(y) \\ \frac{\partial}{\partial y}\mathfrak{a}(y,0) &:= -q_y^{<y} g(y)\end{aligned}$$

$$\begin{aligned}\mathrm{RExp}_x^r\mathfrak{a}(y) &= q_x^{<y}\mathfrak{a}(y,1) + (1-q_x^{<y})\mathfrak{a}(y,0) \\ \frac{\partial}{\partial y}\mathrm{RExp}_x^r\mathfrak{a}(y) &= q_x^{<y}(1-q_y^{<y})g(y) - (1-q_x^{<y})q_y^{<y}g(y) \quad \text{for } y \in [0,1/2) \cup (1/2,1] \\ &= \left[q_x^{<y} - q_y^{<y}\right] \times g(y)\end{aligned}$$

This is > 0 for $q_x^{<y} > q_y^{<y}$, which is iff $x > y$; $= 0$ if $q_x^{<y} = q_y^{<y}$, which is iff $y = x$; and < 0 if $q_x^{<y} < q_y^{<y}$, which is iff $x < y$. Using the continuity of $\mathrm{RExp}^r\mathfrak{a}(y)$ (which follows from the fact that $\mathfrak{a}(1/2,0) = \mathfrak{a}(1/2,1)$) we get that we thus have a unique maximum at $y = x$.

One needs to also check that this is truth directed and that $\mathfrak{a}(y,1) > \mathfrak{a}(y,0) \iff y > 1/2$. Truth directedness is quite easy using the positivity of g. For the ordering argument, we check that $\mathfrak{a}(1/2,0) = \mathfrak{a}(1/2,1)$ then use the truth directedenss. This is:

$$\int_{1/2}^1 (1-q_z^{<z})g(z)dz = \int_0^{1/2} q_z^{<z} g(z)dz$$

using the fact that $g(z) = g(1-z)$.

B No Additive Global Measures

Quick idea: if Acc is additive from $\mathfrak{a}$, then

$$\mathsf{Acc}(\langle x,\ldots,x\rangle, w) = n \times \mathrm{Exp}_{1/n}\mathfrak{a}(x)$$

which is minimized at some $t \neq 1/n$ where $\mathrm{Exp}_{1/n}\mathfrak{a}(x) = \mathrm{RExp}_t^r\mathfrak{a}(x)$

Theorem B.1. *If $\mathfrak{a}$ is strictly r-proper, continuous local accuracy measure and Acc is the additive global measure defined from it,*[8] *i.e.*

$$\mathsf{Acc}(c,w) = \sum_{A\in\mathcal{F}} \mathfrak{a}(c(A), w(A))$$

then Acc is not strictly r-proper and some probability functions are dominated.

[7]This latter constraint is slightly stronger than we actually need.

[8]We can also add constants for the various propositions.

Big Proof (Proof of theorem B.1).

Lemma B.2. *If $\mathfrak{a}$ is strictly r-proper, continuous local accuracy measure and Acc is the additive global measure defined from it (also with the possible addition of constants for the various worlds),*

- *If* $\mathfrak{a}(r^{-1}(1/n), 1) > \mathfrak{a}(r^{-1}(1/n), 0)$ *then*

$$\max_{\mathfrak{a}(x,1)>\mathfrak{a}(x,0)} \mathsf{Acc}(\langle x, \ldots, x\rangle, w) = r^{-1}(1/n)$$

- *If* $\mathfrak{a}(1 - r^{-1}(1 - 1/n), 1) < \mathfrak{a}(1 - r^{-1}(1 - 1/n), 0)$ *then*

$$\max_{\mathfrak{a}(x,1)>\mathfrak{a}(x,0)} \mathsf{Acc}(\langle x, \ldots, x\rangle, w) = 1 - r^{-1}(1 - 1/n)$$

Proof. Note that:

$$\begin{aligned}\mathsf{Acc}(\langle x, \ldots, x\rangle, w) &= \mathfrak{a}(x, 1) + (n-1) \times \mathfrak{a}(x, 0) + K_w \\ &= n \times \left[\frac{1}{n}\mathfrak{a}(x, 1) + \frac{n-1}{n}\mathfrak{a}(x, 0)\right] + K_w\end{aligned}$$

for K_w taking care of the various constants that might be included. They're not going to affect any maximisation results, though.

- If $\mathfrak{a}(x, 1) < \mathfrak{a}(x, 0)$, then $\mathrm{RExp}^r_t[\mathfrak{a}(x)] = (1-r(1-t))\mathfrak{a}(x, 1)+r(1-t)\mathfrak{a}(x, 0)$
- If $\mathfrak{a}(x, 1) > \mathfrak{a}(x, 0)$, then $\mathrm{RExp}^r_t[\mathfrak{a}(x)] = r(t)\mathfrak{a}(x, 1) + (1 - r(t))\mathfrak{a}(x, 0)$

So by strict r-propriety of $\mathfrak{a}$, we have that

- If $\mathfrak{a}(t, 1) < \mathfrak{a}(t, 0)$, then $\max_{\mathfrak{a}(x,1)<\mathfrak{a}(x,0)}[(1 - r(1 - t))\mathfrak{a}(x, 1) + r(1 - t)\mathfrak{a}(x, 0)] = t$.
- If $\mathfrak{a}(t, 1) > \mathfrak{a}(t, 0)$, then $\max_{\mathfrak{a}(x,1)>\mathfrak{a}(x,0)}[r(t)\mathfrak{a}(x, 1) + (1 - r(t))\mathfrak{a}(x, 0)] = t$.

We choose t so that the relevant $r(t)$ or $1-r(1-t)$ are equal to $1/n$. By using that t we then obtain $\frac{1}{n}\mathfrak{a}(x, 1) + \frac{n-1}{n}\mathfrak{a}(x, 0)$ and can thus deduce facts about the maximisation of Acc. We thus obtain our conditions as in the lemma. □

Lemma B.3. *If $\mathfrak{a}$ and r are continuous*[9] *then there is some n such that one of the following holds:*

- $\mathfrak{a}(r^{-1}(1/n), 1) > \mathfrak{a}(r^{-1}(1/n), 0)$ *and* $\mathfrak{a}(1/n, 1) > \mathfrak{a}(1/n, 0)$*, or*
- $\mathfrak{a}(1 - r^{-1}(1 - 1/n), 1) < \mathfrak{a}(1 - r^{-1}(1 - 1/n), 0)$ *and* $\mathfrak{a}(1/n, 1) < \mathfrak{a}(1/n, 0)$.

(so we have that one of the antecedents of the conditionals in lemma B.2 holds)

Proof. Note that as n goes to ∞, all of $r^{-1}(1/n)$, $1-r^{-1}(1-1/n)$ and $1/n$ converge to 0. Thus for sufficiently large n, the $\mathfrak{a}$ ordering of all of these corresponds to that of 0. So depending on whether $\mathfrak{a}(0, 1) > \mathfrak{a}(0, 0)$ we get one of these above conditions holding for sufficiently large n. □

[9] We only need continuity at 0.

Definition B.4. We say that a risk function r is *appropriately risk aware* if there are arbitrarily large n with $r(1 - {}^1\!/\!_n) \neq 1 - {}^1\!/\!_n$ and $r({}^1\!/\!_n) \neq {}^1\!/\!_n$.

Proof of theorem B.1. By piecing together the above two lemmas we chose the n as above and see that $\langle {}^1\!/\!_n, \dots, {}^1\!/\!_n \rangle$ is Acc dominated by $\langle t, \dots, t \rangle$ for $t = 1 - r^{-1}(1 - {}^1\!/\!_n)$ or $r^{-1}({}^1\!/\!_n)$ as appropriate. □

Observe that this shows that for risk averse profiles, the non-strict-proprietary can already be found on a two-element partition.

C Global Accuracy Measures with God-given Ordering

Quick idea: use the old accuracy of the decision weights (from a fixed God-given ordering). Add constants to make sure the ordering is as it was fixed to be.

We construct a r-proper scoring rule from a neutral scoring rule as follows. We start with a fixed ordering on our worlds.

We will assume that $\mathcal{F}$ contains all singletons $\{w\}$. Note, however, that our argument would work, with some simple modifications if we instead assume that $\mathcal{F}$ contains all $\{v \mid v \geqslant w\}$. We have to ensure it contains enough to simply make sure that RExp^r is well-defined (for probabilistic b).

Theorem C.1. *For each strict ordering on worlds, $<$, there is function $f_r^< : \mathbb{R}^{\mathcal{F}} \to \mathbb{R}^{\mathcal{F}}$ where, for probabilistic b,*

$$\mathrm{RExp}^r_b X = \mathrm{Exp}_{f_r^{<_X}(b)} X.$$

Moreover, we can chose this function so that it is bijective; continuous; and c is probabilistic iff $f(c)$ is probabilistic

We will generally drop the sub- and super-scripts on f. In the main text we called this q_c, but that makes sub and super scripts difficult, and it obviates the fact that it is a function dependent on c.

Proof. We first define f on the worlds:

$$f(c)(w) = r\left(\sum_{v \geqslant w} c(v)\right) - r\left(\sum_{v > w} c(v)\right)$$

(see section 6) as we know that for b probabilistic we then have that

$$\mathrm{RExp}^r_b X = \mathrm{Exp}_{f_r^{<_X}(b)} X.$$

To ensure this is defined for all c, not just the probabilistic ones, we need to extend r to take values outside of the unit interval. We do this by just taking $r(x) = x$ for $x < 0$ and $x > 1$.

We then want to extend f to cases where $\mathcal{F}$ contains some non-worlds; i.e. it is not a partition. To do this we define

$$f(c)(A) := \sum_{w \in A} f(c)(w) + rc(A) - r\left(\sum_{w \in A} c(w)\right)$$

We note that if b is probabilistic, then $b(A) = \sum_{w \in A} b(w)$ and thus

$$f(b)(A) := \sum_{w \in A} f(b)(w).$$

One can check that f is bijective, continuous, and maps probabilities to probabilities, and non-probabilities to non-probabilities. $\square$

Proposition C.2. *Suppose B is a risk-neutral proper scoring rule; then*

$$\mathsf{Acc}(c, w) := B(f_r^{<}(c), w) + k_w$$

is strictly r-proper, so long as k_w are chosen so that $<_{\mathsf{Acc}(c)} = <$ for all c.

Proof. Let $K : w \mapsto k_w$. Note that this does not depend on c.

$$\begin{aligned} \mathrm{RExp}_b^r \mathsf{Acc}(c) &= \mathrm{Exp}_{f(b)} \mathsf{Acc}(c) \\ &= \mathrm{Exp}_{f(b)} (B(f(c)) + K) \\ &= \mathrm{Exp}_{f(b)} B(f(c)) + \mathrm{Exp}_{f(b)} K \end{aligned}$$

By strict propriety of B, this is maximised at $f(c) = f(b)$. This is iff $c = b$ because f is 1-1. So Acc is strictly r-proper. $\square$

Proposition C.3. *If B supports dominance, then so does*

$$\mathsf{Acc}(c, w) := B(f_r^{<}(c), w) + k_w.$$

Proof. Suppose c is non-probabilistic. Then $f(c)$ is also non-probabilistic. So there is some q coherent that B-dominates $f(c)$. Consider $f^{-1}(q)$. We claim that this Acc-dominates c.

$$\begin{aligned} \mathsf{Acc}(f^{-1}(q), w) &= B(f(f^{-1}(q)), w) + k_w = B(q, w) + k_w \\ &> B(f(c), w) + k_w = \mathsf{Acc}(c, w) \end{aligned}$$

$\square$

So we have found some risk-weightedly proper scoring rules that support dominance. But they're not truth directed (see proposition 6.2)