(Rethinking The) Foundations of Subjective Probability

Branden Fitelson[1]

Department of Philosophy
&
Center for Cognitive Science (RuCCS)
Rutgers University

branden@fitelson.org
http://fitelson.org/

---

[1] This talk includes joint work with Rachael Briggs (Sydney/ANU), Fabrizio Cariani (NU), Kenny Easwaran (USC), and David McCarthy (Hong Kong).

- This lecture is (mainly) about the relationship two types of epistemic norms: accuracy norms and coherence norms.
- A simple example that everyone will be familiar with:
  - **The Truth Norm for Belief** (TB). Epistemically rational agents should only believe propositions that are true.
  - **The Consistency Norm for Belief** (CB). Epistemically rational agents should have logically consistent belief sets.
- **Fact.** (TB) entails (CB). Suppose S violates (CB). Then, some of S's beliefs are false. Therefore, S violates (TB). □
- I don't think (TB) is a very interesting "norm". It's of little use in *guiding* our epistemic lives. The fact that (TB) entails (CB) suggests that (CB) isn't a very interesting "norm" either.
- I think (some) *preface cases* are counterexamples to (CB) [and $\therefore$ (TB)], as *norms*. If I'm in a (suitably bad) preface case, you won't be able to show me a consistent belief set that *seems "epistemically better"* to me than my own.

- I'll explore the relationship between *different sorts* of accuracy & coherence norms (which seem more interesting).
- For simplicity, I'll talk about *finite*, *logically omniscient*, *opinionated* agents who make *definite judgments* regarding all propositions (or pairs thereof) in some algebra $\mathcal{B}$.
- I will consider three kinds of judgments:
  - **Qualitative**. S believes p [$B_S(p)$]. S disbelieves p [$D_S(p)$].
  - **Comparative**. S is strictly more confident in p than q [$p \succ_S q$]. S is strictly more confident in q than p [$q \succ_S p$]. S is "doxastically indifferent" between p and q [$p \sim_S q$].
  - **Quantitative**. S's degree of confidence/credence in p is r.
- Time permitting, for each of these, I'll discuss relationships between (analogous) accuracy & coherence norms.
- The first step is to *define inaccuracy* for each of the three types of judgments. Once we've done that, we'll examine some new relationships between (in)accuracy & coherence.

- The *inaccuracies* of S's three types of *judgment sets* are (these get increasingly controversial — more on this below).
  - **Qualitative**. Let $\mathfrak{B}$ be the full set of S's qualitative judgments over $\mathcal{B}$. The *innaccuracy* of $\mathfrak{B}$ at a world w is given by the number of incorrect judgments in $\mathfrak{B}$ at w.
    - $B_S(p)$ is (in)correct in w iff p is true (false) at w.
    - $D_S(p)$ is (in)correct in w iff p is false (true) at w.
  - **Comparative**. Let $\mathfrak{C}$ be the full set of S's comparative judgments over $\mathcal{B} \times \mathcal{B}$. The *innaccuracy* of $\mathfrak{C}$ at a world w is given by the number of incorrect judgments in $\mathfrak{C}$ at w.
    - $p \sim_S q$ is (in)correct at w iff $p \equiv q$ is true (false) at w.
    - $p \succ_S q$ is (in)correct at w iff $p \,\&\, {\sim}q$ is true (false) at w.

  ☞ This is the simplest, 2-valued scoring scheme. I'll discuss more subtle comparative scoring schemes below.
  - **Quantitative**. Let b be S's credence function (b is a *function from $\mathcal{B}$ to the real numbers*). The *degree of inaccuracy* of b at a world w [$I(b, w)$] will be given by some *scoring-rule*.

  ☞ There are various scoring rules that have been proposed in the literature. I'll delve more deeply into scoring rules later.

- Consider this notion of (qualitative) *accuracy-dominance*:
  - One set of qualitative judgments $\mathfrak{B}'$ *accuracy-dominates* another $\mathfrak{B}$ iff (i) $\mathfrak{B}'$ has *strictly fewer* incorrect judgments at *some* possible worlds, and (ii) $\mathfrak{B}'$ contains *at most as many* incorrect judgments as $\mathfrak{B}$ at *every* possible world.
- Next, consider the following qualitative coherence norm:

(QC) S should not have a qualitative judgment set $\mathfrak{B}$ that is (*a priori*) *accuracy-dominated* by some alternative set $\mathfrak{B}'$.

- Why is (QC) compelling? For one thing, it is immune from one analogue of preface cases. Allow me to explain.
- In a (sufficiently bad) preface case, S has a judgment set $\mathfrak{B}$ which is inconsistent, but which is such that no consistent alternative $\mathfrak{B}'$ "looks better" to them, *given their evidence*.
- If we show S an alternative, consistent set $\mathfrak{B}'$, their evidence will suggest — *perhaps non-misleadingly!* — that $\mathfrak{B}'$ contains *more incorrect judgments* than their own set $\mathfrak{B}$.

- However, there can be no analogous cases when it comes to violations of (QC). Such cases would be *paradoxical*. Why?
  1. Suppose S violates (QC). Then, there exists a $\mathfrak{B}'$ which S can know *a priori* accuracy-dominates their judgment set $\mathfrak{B}$.
  2. Suppose I show S such an alternative set $\mathfrak{B}'$. If this were analogous to a preface case, then S's evidence would suggest that $\mathfrak{B}'$ has more incorrect judgments than $\mathfrak{B}$.
  3. But, then, I can run S through the (*a priori*) argument which shows that $\mathfrak{B}'$ *cannot* have more incorrect judgments than $\mathfrak{B}$ — in *any* possible world. This is a new kind of trouble for S.
- So, (QC) is more appropriate for those who want to maintain that there *are some* "coherence"/rationality requirements. [That debate has been "stacked" in Kolodny's [8] favor!]
- But, what in the world is this (QC) norm *like*? Are there independent ways to understand it or get a grip on it? Yes.
- I'll give two characterizations of (QC) — in terms of violation and satisfaction — and then I'll mention some applications.

- The following theorem gives *one* way to understand (QC):
  - **Theorem**. *S violates* (QC) *iff $\mathfrak{B}$ contains a subset B which has a majority of incorrect judgments in every possible world.*
- Of course, (QC) is *strictly weaker* than (CB). That (QC) does *not* entail (CB) can be seen *via* a simple counterexample.
  - S can believe *both P and $\sim P$ without* violating (QC).

	P	$\sim P$	$B_S(P)$	$B_S(\sim P)$	$B_S(P)$	$D_S(\sim P)$	$D_S(P)$	$B_S(\sim P)$
w_1	**F**	**T**	incorrect	correct	incorrect	incorrect	correct	correct
w_2	**T**	**F**	correct	incorrect	correct	correct	incorrect	incorrect

- Here's a revealing conjecture about *satisfying* (QC).
  - **Conjecture**. *S satisfies* (QC) *iff* their $\mathfrak{B}$ can be numerically Pr-represented in the following precise, "Lockean" way:

($\mathfrak{L}_{\mathfrak{B}}$) There exists a probability function Pr such that, $\forall p \in \mathcal{B}$:

$$B_S(p) \text{ iff } \Pr(p) > \tfrac{1}{2}, \text{ and } D_S(p) \text{ iff } \Pr(p) \leq \tfrac{1}{2}$$
OR
$$B_S(p) \text{ iff } \Pr(p) \geq \tfrac{1}{2}, \text{ and } D_S(p) \text{ iff } \Pr(p) < \tfrac{1}{2}.$$

- Next, a concrete example in which (QC) *is* violated.

	$\mathfrak{B}$	$\mathfrak{B}'$	$\mathfrak{L}_{\mathfrak{B}'}$
$\sim X \,\&\, \sim Y$	B	D	$\frac{11}{32}$
$X \,\&\, \sim Y$	B	D	$\frac{7}{32}$
$X \,\&\, Y$	B	D	$\frac{13}{32}$
$\sim X \,\&\, Y$	D	D	$\frac{1}{32}$
$\sim Y$	B	B	$\frac{18}{32}$
$X \equiv Y$	B	B	$\frac{24}{32}$
$\sim X$	D	D	$\frac{12}{32}$
X	B	B	$\frac{20}{32}$
$\sim(X \equiv Y)$	D	D	$\frac{8}{32}$
Y	D	D	$\frac{14}{32}$
$X \vee \sim Y$	B	B	$\frac{31}{32}$
$\sim X \vee \sim Y$	B	B	$\frac{19}{32}$
$\sim X \vee Y$	B	B	$\frac{25}{32}$
$X \vee Y$	B	B	$\frac{21}{32}$
$X \vee \sim X$	B	B	1
$X \,\&\, \sim X$	D	D	0

- S's $\mathfrak{B}$ isn't dominated by any *consistent* set, but $\mathfrak{B}$ is — *uniquely* — dominated by the "coherent" $\mathfrak{B}'$.
- As I mentioned, it is *impossible* for S's evidence to *non-misleadingly* make it appear to S that $\mathfrak{B}'$ contains more incorrect judgments than $\mathfrak{B}$.
- But, it is still possible for there to be a weaker sense in which S's evidence non-misleadingly suggests that her violation of (QC) may be "OK".
- Suppose S's evidence *non-misleadingly* supports the truth of the conjunction $X \,\&\, \sim Y$. Then, S may reason in the following way, when they see $\mathfrak{B}'$.
  - Look, I realize that $\mathfrak{B}'$ cannot have more incorrect judgments than my $\mathfrak{B}$ does.
  - But, *I have good evidence for/know $X \,\&\, \sim Y$*, which (if true) *rules-out* $\mathfrak{B}'$. Since *my* violation of (QC) is *equivalent* to my being dominated by $\mathfrak{B}'$, why should I be *moved* by my violation of (QC)? [*Kolodny's revenge*!?]

- Rachael Briggs, Kenny Easwaran and I are investigating various applications of this accuracy-dominance approach.
- One interesting application is to judgment aggregation. *E.g.*,
  - Majority rule aggregations of the judgments of a bunch of agents — each of whom satisfy (CB) — need not satisfy (CB).
- Question: does majority rule preserve *our* "qualitative coherence", *viz.*, is (QC) preserved by MR? Answer: No!
  - There are sets of judges (minimum # = 5) that (severally) satisfy (QC), while their majority profile *violates* (QC).
  - But, if a set of judges is (severally) *consistent* (in the classical sense), then their majority profile *must* satisfy (QC).
- Another application: define an "entailment" relation ($\Vdash$), as a relation between a set of judgments Γ and a p-judgment:
  - $\Gamma \Vdash B_S(p)$ iff $\Gamma \cup \{D_S(p)\}$ is "*incoherent*" [in the (QC) sense].
  - $\Gamma \Vdash D_S(p)$ iff $\Gamma \cup \{B_S(p)\}$ is "*incoherent*" [in the (QC) sense].
- Next: accuracy & coherence of *comparative* judgments.

- Once upon a time, *comparative* confidence judgments were thought to be more "secure" or "basic" or "fundamental" than *numerical* confidence/credence judgments [7].
- In his watershed essay, de Finetti [2] begins his story about the foundation of subjective probability theory, as follows:

> Let us consider a well-defined event and suppose that we do not know in advance whether it will occur or not; the doubt about its occurrence to which we are subject lends itself to comparison, and, consequently, to gradation. If we acknowledge only, first, that one uncertain event can only appear to us (a) equally probable, (b) more probable, or (c) less probable than another; second, that an uncertain event always seems to us more probable than an impossible event and less probable than a necessary event; and finally, third, that when we judge an event E more probable than an event E', which is itself judged more probable than an event E'', the event E can only appear more probable than E'' (transitive property), it will suffice to add to these three evidently trivial axioms a fourth, itself of a purely qualitative nature, in order to construct rigorously the whole theory of probability.

- What de Finetti is describing here is a "foundationalist" conception of subjective probability, where the *foundation* consists of *relations of comparative confidence* ($\succ_S$).

- First, de Finetti thought that the following two "axioms" for the "strictly more confident in" relation $\succ_S$ were *self-evident*:

(1) $\succ_S$ is transitive, asymmetric, and irreflexive.
  - More precisely, that $\succ_S$ imposes a *strict total order* on $\mathcal{B}$. [Note: de Finetti *assumes*: $p \sim_S q$ *iff* $p \nsucc_S q$ & $q \nsucc_S p$.]

(2) For all p, $p \nsucc_S \top$ and $\bot \nsucc_S p$.

(2.1) S should never be strictly more confident in any p than $\top$.
(2.2) S should never be strictly more confident in $\bot$ than any p.

- Second, de Finetti thought that the following "additivity" axiom — together with axioms (1) and (2) — *suffices* to ensure *numerical probabilistic representatability* of $\succ_S$.

(3) If $\langle p, q\rangle$ and $\langle p, r\rangle$ are both mutually exclusive pairs, then:

$$q \succ_S r \textit{ only if } (p \vee q) \succ_S (p \vee r).$$

- That is, de Finetti believed that the following was true:

(†) If the relation $\succ_S$ satisfies (1)–(3), then there exists a *numerical probability function* b such that:

$$p \succ_S q \textit{ if and only if } b(p) > b(q).$$

- A few decades later, it was discovered [9] that de Finetti was *wrong* about (†). This was a *crack in the foundation.*
- Scott [16] gave *a much stronger* "additivity" axiom (3′), such that (1)–(3′) *does suffice* for numerical Pr-representability of $\succ_S$. [See the Extras slides for a discussion of Scott's Axiom.]
- Here, I'll focus on *justifying de Finetti*'s "intuitive" axioms for $\succ_S$. *Nobody* seems to offer *any* justifications for (1)–(3) [3, 4]. But, some *do* offer (dominance) justifications of the *numerical* Pr-axioms (see Part II). d.F.'s overall strategy was:

(i) First, lay down some "*intuitive*" axioms for $\succ_S$-orderings.
(ii) Then, show that these axioms suffice to ensure numerical probabilistic representability of any "intuitive" $\succ_S$-ordering.
(iii) Finally, justify the axioms of numerical probability theory (*via* the *Brier-dominance approach*, to be discussed Friday).

- If this could be achieved, then one could ground the desired "foundationalist" conception of subjective probability. This project failed at step (ii). But, I think even (i) is dubious…

- Recall our definition of "inaccuracy" of a set of comparative judgments. [Note: other "scoring schemes" are possible for comparative judgments. I will return to this issue below.]
- We can use an analogous notion of "accuracy dominance" to *justify* (some) "axioms" of comparative probability. To wit:
    - $\mathfrak{C}'$ *accuracy-dominates* $\mathfrak{C}$ iff $\mathfrak{C}'$ contains *strictly fewer* incorrect judgments than $\mathfrak{C}$ at *some* w's, and $\mathfrak{C}'$ contains *at most as many* incorrect judgments as $\mathfrak{C}$ at *every* w.

☞ de Finetti's axioms (1) & (2) *can* then (*modulo* a small *caveat* — see below) be given an accuracy-dominance justification.

- **Theorem**. If S's $\succ_S$-ordering (*viz.*, S's comparison set $\mathfrak{C}$) violates *either* (1) *or* (2), then (*modulo* a small *caveat* — see below) there exists a $\mathfrak{C}'$ which *accuracy-dominates* $\mathfrak{C}$.

- As far as I know, we (this is joint work with David McCarthy) are the first to offer *any* justification of (1) and/or (2).

  [And, ours is *natural*, given de Finetti's *numerical* approach [1]. I'll explain what I mean by this later on in today's lecture.]

- But, *all is not beer and skittles...*
- de Finetti's axiom (3) does *not* have an accuracy-dominance justification [$\therefore$ the Scott Axiom (3′), which is required for numerical Pr-representation of $\succ_S$ doesn't have one either].
- This suggests that a de Finetti-style "foundationalist" approach to epistemic probability contains a *logical gap*.

☞ *Why* should we think $\succ_S$ has a numerical Pr-representation?

- It is very interesting to note that the principles of comparative probability that have been uncontroversial (or "self-evident") in the literature [4] are (basically) those with an accuracy-dominance justification in our sense. *E.g.*,

(4) If p entails q, then $p \nsucc_S q$. ["*Monotonicity*" of $\succ_S$.]

- And, those which have been seen as controversial *fail* to have an accuracy-dominance justification. *E.g.*, "additivity" (3), the Scott Axiom (3′), and other principles, such as:

(5) If $p \succ_S q$, then $\sim p \nsucc_S \sim q$. ["*Complementarity*" of $\succ_S$.]

- Consider the following weak form of transitivity.

(WT) If $p \succ_S q$ and $q \succ_S r$, then $r \nsucc_S p$.

- If S violates (WT), then S is accuracy-dominated. Proof:

	P	Q	R	$P \succ_S Q$	$Q \succ_S R$	$R \succ_S P$	$P \sim_S Q$	$Q \sim_S R$	$P \sim_S R$
w_1	**T**	**T**	**T**	incorrect	incorrect	incorrect	correct	correct	correct
w_2	**T**	**T**	**F**	incorrect	correct	incorrect	correct	incorrect	incorrect
w_3	**T**	**F**	**T**	correct	incorrect	incorrect	incorrect	incorrect	correct
w_4	**T**	**F**	**F**	correct	incorrect	incorrect	incorrect	correct	incorrect
w_5	**F**	**T**	**T**	incorrect	incorrect	correct	incorrect	correct	incorrect
w_6	**F**	**T**	**F**	incorrect	correct	incorrect	incorrect	incorrect	correct
w_7	**F**	**F**	**T**	incorrect	incorrect	correct	correct	incorrect	incorrect
w_8	**F**	**F**	**F**	incorrect	incorrect	incorrect	correct	correct	correct

- In fact, this is the *unique* dominating $\mathfrak{C}'$. [*Kolodny's revenge* applies — suppose S *has good reason for/knows* $P \succ_S Q$.]
- The "small caveat" is that *not all* violations of transitivity are dominated! *E.g.*, an S such that $p \sim_S q$, $q \sim_S r$, $p \succ_S r$.
    - This is welcome (to me), since I think there are permissible examples of this kind (*e.g.*, perceptual indiscriminability).
- Next: other consequences of our 2-valued scoring scheme.

- If S violates *Monotonicity* (4), then S is accuracy-dominated.

(4) If p entails q, then $p \nsucc_S q$.

	P	Q	$P \succ_S Q$	$Q \succ_S P$
w_1	**T**	**T**	B	B
w_2	**T**	**F**	—	—
w_3	**F**	**T**	C	A
w_4	**F**	**F**	B	B

- In fact, as this table shows, *any scoring scheme with the above structure* (where $A < C$) entails Monotonicity.
- To see that de Finetti's additivity axiom (3) does *not* have a dominance justification, one must look at *all* the possible ways of "fixing" a violation of (3), and show that *none* of these lead to a comparison set that dominates the original.
- There aren't that many cases to check. [I won't show them.]
- Next: an objection to our simple, 2-valued scoring scheme.

- Recall our definition of "inaccuracy"/"accuracy dominance" for a (complete) set of comparative judgments $\mathfrak{C}$.
  - **Comparative.** Let $\mathfrak{C}$ be the full set of S's comparative judgments over $\mathcal{B} \times \mathcal{B}$. The *innaccuracy* of $\mathfrak{C}$ at a world w is given by the number of incorrect judgments in $\mathfrak{C}$ at w.
    - $p \sim_S q$ is (in)correct at w iff $p \equiv q$ is true (false) at w.
    - $p \succ_S q$ is (in)correct at w iff $p \;\&\; {\sim} q$ is true (false) at w.
  - $\mathfrak{C}'$ *accuracy-dominates* $\mathfrak{C}$ iff $\mathfrak{C}'$ contains *strictly fewer* incorrect judgments than $\mathfrak{C}$ at *some* w's, and $\mathfrak{C}'$ contains *at most as many* incorrect judgments as $\mathfrak{C}$ at *every* w.
- This simple, 2-valed scoring scheme may seem overly simplistic. It is based on the following underlying norm:

  (†) S should be more confident in truths than falsehoods.
- So, if p is **T** and q is **F**, then the judgments $q \succ_S p$ and $p \sim_S q$ are in violation of this basic underlying norm (†).
- But, (†) *alone* does not justify our choice of 2-valued scheme. Indeed, other scoring schemes seem plausible.

- Let's use "+1" to denote *best* epistemic status, "−1" to denote *worst* epistemic status, and "0" to denote "middling" epistemic status. Our simplest, 2-valued scheme is:

	P	Q	$P \succ_S Q$	$Q \succ_S P$	$Q \sim_S P$
w_1	**T**	**T**	−1	−1	+1
w_2	**T**	**F**	+1	−1	−1
w_3	**F**	**T**	−1	+1	−1
w_4	**F**	**F**	−1	−1	+1

- If we're going to use only 2-values ("correct/incorrect"), then it seems to me that this scheme is *forced* on us, by (†).
- But, one might think that a 3-valued scheme makes more sense. David Christensen makes the following observation.

> Suppose I'm going to flip a coin. Can I rationally be indifferent between heads (H) and tails (T)? It seems that $H \sim_S T$ would be dominated by $H \succ_S T$ (or $T \succ_S H$), since $H \sim_S T$ is guaranteed to be "incorrect" and the latter aren't.

- Christensen is right. And, he suggests a 3-valued scheme.

	P	Q	$P \succ_S Q$	$Q \succ_S P$	$Q \sim_S P$
w_1	**T**	**T**	0	0	+1
w_2	**T**	**F**	+1	−1	0
w_3	**F**	**T**	−1	+1	0
w_4	**F**	**F**	0	0	+1

- I agree that D.C.'s scheme does seem superior (intuitively) to our simplest 2-valued scoring scheme (in various ways).
- If we use this (or some other) 3-valued scheme, the obvious way to calculate the score of $\mathfrak{C}$ (at w) is to take the *sum* of these 3-valued scores for all the propositions in $\mathfrak{C}$ (at w).
- Then, we would define accuracy-dominance as follows:
  - $\mathfrak{C}'$ *accuracy-dominates* $\mathfrak{C}$ iff $\mathfrak{C}'$ has a *higher* score than $\mathfrak{C}$ at *some* w, and $\mathfrak{C}'$ doesn't have a lower score than $\mathfrak{C}$ at *any* w.
- In any event, moving to a 3-valued scheme can *not* fill the gap in de Finetti's justification/grounding of subjective probability theory. Indeed, we have an *impossibility result.*

Theorem. *No* 2 *or* 3-valued scoring scheme is such that:

(0) S entails (at least *some* instances of) *both* transitivity *and* additivity as (weak) dominance norms.

and, the the following eight (8) scoring *desiderata* are met:

(1) Having a subset of judgments $\{p \succ_S q, p \succ_S r, q \sim_S r\}$ should not — *in and of itself* — ensure "incoherence".

(2) *Ditto* for subsets of the form $\{p \succ_S q, p \succ_S r, q \succ_S r\}$.

(3) $p \succ_S q$ should get a "worst" score when p is **F** and q is **T**.

(4) $p \succ_S q$ should get the same score when p and q are both **T** as it does when p and q are both **F**.

(5) $p \sim_S q$ should get the same score when p and q are both **T** as it does when p and q are both **F**.

(6) $p \sim_S q$ should get the same score when p is **T** and q is **F** as it does when p is **F** and q is **T**.

(7) The score of $p \succ_S q$ when p is **T** and q is **F** should not be strictly worse than the score of $p \succ_S q$ when p, q are both **T**.

(8) The score of $p \succ_S q$ when p is **T** and q is **F** should be strictly better than the score of $p \succ_S q$ when p is **F** and q is **T**.

- These eight *desiderata* seem to be sacrosanct (Christensen and everyone else I've talked to seems to accept all of them).
- The upshot of our Theorem is that — *it doesn't matter which scoring scheme you use*. *No* scoring scheme can ground *all* of de Finetti's axioms for comparative probability.
  - In fact, our simplest 2-valued scheme *gets as close as any 2 or 3-valued scheme* to grounding all of de Finetti's axioms. [This is why I introduced it first. It is *simple*, and *maximally charitable* to de Finetti (with respect to his project).]
- So, it seems there is no accuracy-dominance justification of all of de Finetti's intuitive axioms (much less the *unintuitive* Scott Axiom — see Extras slides). This *re*-raises a question:

  ☞ *Why* should we think $\succ_S$ has a numerical Pr-representation?
- There seems to be no compelling reason to suppose that our comparative confidence orderings are (numerically) probabilistically representable. This is an important *lacuna*.
- Next: Quantitative judgments (*viz.*, numerical credences).

- Standard arguments for *probabilism* are of the form:
  - An agent S has a non-probabilistic partial belief function b iff ($\Longleftrightarrow$) S has some "bad" property B (*in virtue of* the fact that their c.f. b has a certain "bad" *formal* property F).
- These *arguments* rest on *Theorems* ($\Rightarrow$) and *Converse Theorems* ($\Leftarrow$): b is non-Pr $\Longleftrightarrow$ b has formal property F.
  - **Dutch Book Arguments** [14, 1]. B is *susceptibility to sure monetary loss* (in a certain betting set-up), and F is the formal role played by non-Pr b's in the DBT/Converse DBT.
  - **Representation Theorem Arguments** [15]. B is *having preferences that violate some of Savage's axioms* (and/or *being unrepresentable as an expected utility maximizer*), and F is the formal role played by non-Pr b's in the RT.
- To the extent that we have reasons to avoid these B's, these arguments provide reasons (not) to have a(n) (in)coherent b.
- Joycean [6] arguments for probabilism also fit this pattern.

- According to Joyce [6], if we view credences as "estimates" of (suitable) "numerical representations of truth-values" of propositions, then we can give an argument for probabilism that is based on the "accuracy" of these "estimates".
- Consider a very simple, logically omniscient, opinionated agent S who has only one atomic sentence P in his language.
- All that matters concerning S's *coherence* is whether S's credences $b(P)$, $b(\sim P)$ *sum to one* (*and are non-negative*).
- Following Joyce, let's associate the truth-value **T** (at each world w) with the number 1 and the truth-value **F** with 0.
- The idea will be that $b(p)$ represents the agent S's "estimate" of the truth-value of p. These "estimates" will be subject to an accuracy norm, which will, in turn, give rise to a coherence norm (*viz.*, *probabilism*) for credences.
- Next, measuring the "accuracy" of Joycean "estimates" (b).

- The *inaccuracy* of $b(p)$ at world w will be b's "distance (d) from the number associated with p's truth-value" at w.
- **Example.** Suppose S has just two (contingent) propositions $\{P, \sim P\}$ in their doxastic space. Then, there are two salient possible worlds (w_1 in which P is **T**, and w_2 in which P is **F**). And, the *overall inaccuracy* of b at w [$I(b,w)$] is given by:
  - $I(b,w_1) = d(b(P),1) + d(b(\sim P),0)$.
  - $I(b,w_2) = d(b(P),0) + d(b(\sim P),1)$.
- Various measures (d) of "distance from 0/1-truth-value" have been proposed/defended in the historical literature.
- de Finetti [2] endorsed the following measure of "distance from truth-value" (in one argument for probabilism):
  - $\mathfrak{s}(x,y) = (x-y)^2$.
- The distance measure $\mathfrak{s}$ gives rise to a measure of *overall* inaccuracy ($I_{\mathfrak{s}}$), which is known as the *Brier Score*. In our toy example, the Brier Scores of b in worlds w_1 and w_2 are:

  - $I_{\mathfrak{s}}(b, w_1) = \mathfrak{s}(b(P), 1) + \mathfrak{s}(b(\sim P), 0) = (b(P) - 1)^2 + b(\sim P)^2$.
  - $I_{\mathfrak{s}}(b, w_2) = \mathfrak{s}(b(P), 0) + \mathfrak{s}(b(\sim P), 1) = b(P)^2 + (b(\sim P) - 1)^2$.
- If one adopts the Brier Score as one's measure of b's inaccuracy, then one can give an "accuracy-dominance argument" for the axioms of the probability calculus.
- de Finetti [1] was the first to prove such a *Brier*-dominance theorem. Joyce [6, 5] interprets this as *accuracy*-dominance.
  - **Theorem** (de Finetti). b is *non*-probabilistic *if and only if* there exists a *probabilistic* credence function b' such that (a) *b' has a strictly lower Brier Score than b at some worlds*, and (b) *b' never has a greater Brier Score than b at any world*.

☞ The "bad" B is: *being dominated in accuracy*; and, the "bad" F is: the c.f. b is *Brier-dominated* by some coherent c.f. b'.

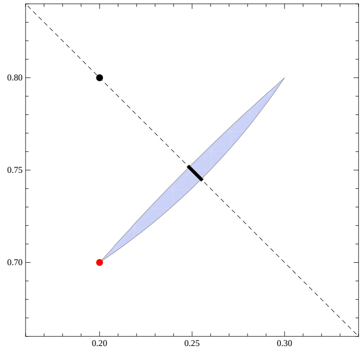
- Finalé: three worries about Joyce-style arguments for probabilism. (1) sensitivity to choice of "accuracy measure", (2) an "evidentialist" worry, (3) language-dependence.

- Recall how inaccuracy measures $I(b, w)$ are defined in terms of measures d of "distance from truth-value".
  - $I(b, w_1) = d(b(P), 1) + d(b(\sim P), 0)$.
  - $I(b, w_2) = d(b(P), 0) + d(b(\sim P), 1)$.
- The Brier score uses the square-of-the-difference measure:
  - $\mathfrak{s}(x, y) = (x - y)^2$.
- And, this choice supports de Finetti/Joyce arguments.
- Interestingly, there are many other (plausible) measures of "distance" that do *not* support de Finetti/Joyce arguments.
- For instance, if $d = \sqrt{\mathfrak{s}}$ (absolute-value-of-the-difference), then the de Finetti/Joyce theorem *fails*. That is, if we adopt:
  - $\mathfrak{a}(x, y) = \sqrt{(x - y)^2} = |x - y|$.

  as our measure of "distance", then the resulting inaccuracy measure does *not* satisfy the de Finetti/Joyce theorem [10].
- Suggestion: *look for agreement between d-measures*. Next: an "evidentialist" worry about de Finetti/Joyce arguments.

- Suppose S adopts the Brier Score as their I-measure, and that S's b is non-probabilistic. Then, there are alternative (coherent) credence functions b' that accuracy-dominate b.
- Intuitively, these b' functions should "look epistemically better" (in a precise sense) than S's current credences b.
- But, a possible "evidentialist" worry remains.
- Consider a very simple toy agent S with one sentence P in their language. And, suppose S's credence function assigns $b(P) = 0.2$ and $b(\sim P) = 0.7$. So, S's b is *non*-probabilistic.
- It follows from de Finetti/Joyce's theorems that there is *a specific set of* credence functions b' that *Brier-dominate* b.
- It seems that this alternative credence function b' should *inevitably* "look epistemically better" to S than her current credence function b. Our worry is that this *needn't* be so.
- Consider the following (toy) illustration of our worry.

- The red dot in the figure is S's credence function b. The shaded region depicts the functions b' that *Brier-dominate* b. [The black dot at $\langle 0.2, 0.8\rangle$ depicts the *only probabilistic* credence function that is compatible with $b(P) = 0.2$.]



- Suppose that S *has good reason* to assign $b(P) = 0.2$ (*i.e.*, S's total evidence E *supports* $b(P) = 0.2$).
- Here, *all* the Brier-dominating functions b' are s.t. $b'(p) \neq 0.2$.
- So, *all the Brier-dominating functions b' may be "ruled-out" by S's evidence.*
- Then, b' *needn't* "look better" than b.

☞ This is analogous to what happens with (bad) preface cases. Evidential norms can sometimes "trump" coherence norms.

- Next: accuracy, language-dependence, and de Finetti/Joyce.

- Suppose we have two *numerical* quantities ϕ and ψ. These might be, for instance, the velocities (in some common units) of two objects, at some time (or some other physical property, like temperature, of two objects at a time).
- Suppose further that we have two sets of (false) predictions concerning the values of ϕ and ψ, which are entailed by two (false) competing hypotheses H_1 and H_2.
- Finally, let's use "T" to denote *the truth* about the values of ϕ and ψ (or, if you prefer, the true hypothesis about their values) — in our standard units. And, let H_1, H_2, and T be:

	ϕ	ψ	α	β
H_1	0.150	1.225	0.925	2.000
H_2	0.100	1.000	0.800	1.700
T	0.000	1.000	1.000	2.000

- It seems clear that the predictions of H_2 are "closer to the truth T about ϕ and ψ" than the predictions of H_1 are.

- However, as Popper [13, Appendix 2] showed (using a recipe invented by David Miller [11]), there exist quantities α and β (as in the table) satisfying both of the following conditions.
    1. α and β are symmetrically inter-definable with respect to ϕ and ψ in the following (linear) way:

$$\begin{array}{ll} \alpha = \psi - 2\phi & \beta = 2\psi - 3\phi \\ \phi = \beta - 2\alpha & \psi = 2\beta - 3\alpha \end{array}$$

    2. The values for α and β entailed by H_2 are strictly "farther from the truth T about α and β" than those entailed by H_1.
- As Miller [11] explains (see [12, Chapter 11] for a recent survey), there is a much more general result in the vicinity.
- For *any* pair of false theories H_1 and H_2 about ϕ and ψ, *many* relations of "closer to the truth" can be *reversed* by looking at what the estimates provided by H_1 and H_2 for ϕ and ψ entail about quantities α and β, which are given by:

$$\begin{array}{ll} \alpha = \mathsf{a}\psi + \mathsf{b}\phi & \beta = \mathsf{c}\psi + \mathsf{d}\phi \\ \phi = \mathsf{a}\beta + \mathsf{b}\alpha & \psi = \mathsf{c}\beta + \mathsf{d}\alpha \end{array}$$

- This is Miller's 2^{nd} (Quantitative) LDP. It *is* a (potential) threat.

- According to Joyce [6], if we view credences as "estimates" of (suitable) "numerical representations of truth-values" of propositions, then we can give an argument for probabilism that is based on the "accuracy" of these "estimates".
- Consider a very simple, logically omniscient, opinionated agent S who has only one atomic sentence P in his language.
- All that matters concerning S's *coherence* is whether S's credences b in P and $\sim P$ sum to one (and are non-negative).
- Following Joyce, let's associate the truth-value **T** (at each world w) with the number 1 and the truth-value **F** with 0. [This is our "numerical representation of the truth-values" of P, $\sim P$.]
- Let ϕ be the *numerical* value we have associated with P's truth-value, and let ψ be the *numerical* value we have associated with $\sim P$'s truth-value (of course, ϕ and ψ will vary in the obvious ways across the two salient possible worlds).
- Next, measuring the "accuracy" of Joycean "estimates" (b).

- Suppose S has a numerical credence function $b : \mathcal{B} \mapsto \mathbb{R}$ (these b functions are *opinionated*, of course, and so we're ignoring suspension of judgment here, once again).
- The *inaccuracy* of $b(p)$ at world w will be b's "distance (d) from the number associated with p's truth-value" at w.
- **Example.** Suppose S has just two (contingent) propositions $\{P, \sim P\}$ in their doxastic space. Then, there are two salient possible worlds (w_1 in which P is **T**, and w_2 in which P is **F**). And, the *overall inaccuracy* of b at w [$I(b, w)$] is given by:
    - $I(b, w_1) = d(b(P), 1) + d(b(\sim P), 0)$.
    - $I(b, w_2) = d(b(P), 0) + d(b(\sim P), 1)$.
- Various measures (d) of "distance from 0/1-truth-value" have been proposed/defended in the historical literature.
- The most popular choice (for giving an accuracy-dominance justification of probabilism) has been the *squared-difference* measure of "distance from 0/1-truth-value", which is:
    - $\mathfrak{s}(x, y) = (x - y)^2$.

- The distance measure $\mathfrak{s}$ gives rise to a measure of *overall* inaccuracy ($I_{\mathfrak{s}}$), which is known as the Brier Score. In our toy example, the Brier Scores of b in worlds w_1 and w_2 are:
  - $I_{\mathfrak{s}}(b, w_1) = \mathfrak{s}(b(P), 1) + \mathfrak{s}(b(\sim P), 0) = (b(P) - 1)^2 + b(\sim P)^2$.
  - $I_{\mathfrak{s}}(b, w_2) = \mathfrak{s}(b(P), 0) + \mathfrak{s}(b(\sim P), 1) = b(P)^2 + (b(\sim P) - 1)^2$.
- If one adopts the Brier Score as one's measure of b's inaccuracy, then one can give an accuracy-dominance argument for the axioms of the probability calculus.
- de Finetti [1] was the first to prove such a *Brier*-dominance theorem. Joyce [6, 5] interprets this as *accuracy*-dominance.
  - **Theorem** (de Finetti). b is *non*-probabilistic *if and only if* there exists a *probabilistic* credence function b' such that (a) *b' has a strictly lower Brier Score than b at some worlds*, and (b) *b' never has a greater Brier Score than b at any world.*
- One can use other underlying measures of distance d here and still preserve a de Finetti-style Theorem. Miller's second LD problem is a potential threat to *any* of these approaches.

- The easiest way to see the (potential) threat posed by Miller's 2^{nd} LD problem is to consider a simple numerical example of a toy agent S of the type we've been discussing.

	ϕ	ψ
b	$\frac{1}{2}$	$\frac{1}{4}$

- Here, S's "estimates" (b) of ϕ and ψ do not sum to one. As a result, there exist alternative "estimates" b' of ϕ/ψ that Brier-dominate b in both of the salient possible worlds.

	ϕ	ψ
b	$\frac{1}{2}$	$\frac{1}{4}$
b'	$\frac{5}{8}$	$\frac{3}{8}$
w_1	0	1
w_2	1	0

- b' is the *Euclidean-closest* (to b) set of "estimates" of ϕ and ψ that Brier-dominate b — with respect to ϕ/ψ "estimation".

- Does Joycean "numerical estimation" face a (*prima facie*) problem analogous to Miller's second LD problem? Yes!

	ϕ	ψ	α	β
b	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{9}{16}$	$\frac{3}{16}$
b'	$\frac{5}{8}$	$\frac{3}{8}$	$\frac{3}{4}$	$\frac{1}{4}$
w_1	0	1	$\frac{7}{16}$	$\frac{9}{16}$
w_2	1	0	$\frac{9}{16}$	$\frac{7}{16}$

- Here, we have numerical quantities α and β, such that:

(i) α/β are symmetrically inter-definable w.r.t ϕ/ψ, *via* ($f^\star$):

$$\alpha = \tfrac{1}{2}\phi + \tfrac{1}{2}\psi + \tfrac{1}{16}\left(\frac{\phi+\psi}{\phi-\psi}\right) \qquad \beta = \tfrac{1}{2}\phi + \tfrac{1}{2}\psi - \tfrac{1}{16}\left(\frac{\phi+\psi}{\phi-\psi}\right)$$

$$\phi = \tfrac{1}{2}\alpha + \tfrac{1}{2}\beta + \tfrac{1}{16}\left(\frac{\alpha+\beta}{\alpha-\beta}\right) \qquad \psi = \tfrac{1}{2}\alpha + \tfrac{1}{2}\beta - \tfrac{1}{16}\left(\frac{\alpha+\beta}{\alpha-\beta}\right)$$

(ii) The "estimates" of α/β entailed (*via* $f^\star$) by b *Brier dominate* the "estimates" of α/β entailed by b' (*via* $f^\star$).

☞ So, we have a Miller-style *reversal* of Brier-domination here!

- Here is a *more general theorem* about our toy agent S.

Theorem. For *any* coherent b' that Brier-dominates S's credence function b with respect to ϕ and ψ, there exist quantities α and β that are symmetrically inter-definable with respect to ϕ and ψ, *via* the transformation $f^\star$ above, such that *b Brier-dominates b'* with respect to α and β.

- It is also noteworthy that the *true* values of α and β "behave like truth-values", in the sense that (a) the true value of α (β) in w_1 (w_2) is identical to the true value of β (α) in w_2 (w_1), and (b) the true values of α and β always *sum to one.*
- Indeed, this transformation $f^\star$ is guaranteed to *preserve coherence* of *all* dominating b''s, *and* the "truth-vectors".
- It is not a coincidence that $f^\star$ is *non*-linear. It can be shown that *no linear f* can play the role that $f^\star$ plays here. There are several reasons for this (some of which I'll mention below).
- Next, two possible responses to this (*prima facie*) threat.

- **Response #1: Naturalness**. The first response is to argue that the ϕ/ψ "numerical representation of the truth-values of $P/{\sim}P$" is somehow "more natural" than the α/β "numerical representation of the truth-values of $P/{\sim}P$".
- I'm not sure how such a "naturalness argument" would go.
- After all, the truth-values of $P/{\sim}P$ are *disanalogous* to *numerical* physical quantities like velocity or temperature.
- In the case of temperature, for instance, the *numerical level* of description is (arguably) the most fundamental/scientific.
- But, in the case of truth-values, their theoretical role seems to be given *fundamentally* at the level of their *algebraic* and *meta-logical* (*viz.*, *logico-structural*) properties.
- The "numerical properties" of the truth-values (*if there be such*) do not seem to be theoretically fundamental. So, it's not clear to me how probative "naturalness" is here.
- I think there is a more promising line of response...

- **Response #2: Disanalogies between "Estimation" & Prediction**.
- The second line of response (which I prefer) is to argue that there are crucial disanalogies between "estimation" (in Joyce's sense) and prediction (in Miller's sense).
- Let $\ulcorner \mathcal{E}_S(x,y) = \langle p,q \rangle \urcorner$ express the claim that $\ulcorner S$ is committed to the values $\langle p,q \rangle$ as their "estimates" of the quantities $\langle x,y \rangle \urcorner$. Our "reversal argument" presupposes the following (as applied to our toy, numerical S, above):

  (†) If $\mathcal{E}_S(\phi,\psi) = \langle p,q \rangle$, then $\mathcal{E}_S(\alpha,\beta) = f^{\star}(p,q)$, where $f^{\star}$ is the symmetric inter-translation function that maps values of $\langle \phi,\psi \rangle$ to/from values of $\langle \alpha,\beta \rangle$ in our Theorem.
- Ultimately, what Joyce needs to argue is that (†) is *false*.
- In order to do this, Joyce needs to tell us more about what "estimation" is. Ideally, he needs to give us a *theory* of $\mathcal{E}_S$.
- Unfortunately, what Joyce *explicitly says* about $\mathcal{E}_S$ is insufficient to explain why (†) should come out false.

- The only *explicit* commitments Joyce has re $\mathcal{E}_S$ seem to be:
  1. Estimates are *not guesses*. Joyce [6, 587] explicitly distinguishes estimation and guessing.
     - This doesn't help me assess (†), so I won't discuss it.
  2. Estimates are *not expectations*. Joyce [6, 587–8] explicitly *disavows* thinking of estimates as expectations. Indeed, Joyce thinks it would be *question-begging* to think of "estimation" as expectation (*e.g.*, expectation *builds-in* too much probabilistic structure from the outset).
     - If estimates *were* expectations, then this would entail that (†) is *false*, since this would rule-out *all non-linear* transformation functions. [A reason to *like* expectation?]
  3. Estimates are *not assertions that* the values of the parameters *are such-and-so*. This is clear, since it's *not* a good idea to assert things that you know (*a priori*) *must be false*. And, this happens whenever you offer "estimates" of "numerical correlates of truth-values" that are *non-extreme*.
     - If estimates *were* assertions, then this would entail that (†) is *true* — assuming a truth/closure norm for assertions.

- de Finetti conjectured that his axioms (1)–(3) *suffice* for:

  (†) $(\exists \Pr)(\forall p)(\forall q)[p \succ_S q \textit{ if and only if } \Pr(p) > \Pr(q)]$.
- But, it was later discovered [9] that this was *false*. That is, there exist $\succ_S$-orderings satisfying de Finetti's axioms (1)–(3), but for which there is no Pr-representation.
- Here is the counterexample reported in [9]. Imagine a Boolean algebra containing 5 states/atoms: $\{s_1, s_2, s_3, s_4, s_5\}$.
- And, suppose we have the following four $\succ_S$ relations:

  (i) $s_4 \succ_S s_1 \vee s_3$.

  (ii) $s_2 \vee s_3 \succ_S s_1 \vee s_4$.

  (iii) $s_1 \vee s_5 \succ_S s_3 \vee s_4$.

  (iv) $s_1 \vee s_3 \vee s_4 \succ_S s_2 \vee s_5$.
- It can be shown that (i)–(iv) are compatible with de Finetti's axioms (1)–(3), *but* (i)–(iv) have no Pr-representation.
- Exercise: verify that this is a counterexample. [Open question: are there any smaller counterexamples?]

- Before stating the Scott Axiom, we'll need one definition:

  Definition. For each state description $\mathfrak{s}$ and each *sequence* (n-tuple) of propositions $\mathbf{Z} = \langle z_1, \ldots, z_n \rangle \in \prod_n \mathcal{B}$, let $\mathfrak{c}(\mathfrak{s}, \mathbf{Z})$ be the number of elements of $\mathbf{Z}$ that are entailed by $\mathfrak{s}$.

- OK, here's the (dreaded) Scott Axiom:

  (SA) Let $\mathbf{X}, \mathbf{Y} \in \prod_n \mathcal{B}$ be (arbitrary) sequences of propositions, each having length $n > 0$. Let $\langle x_1, \ldots, x_n \rangle$ denote the members of $\mathbf{X}$, and $\langle y_1, \ldots, y_n \rangle$ denote the members of $\mathbf{Y}$.

  If the following two conditions are satisfied

  i. For every state description $\mathfrak{s}$, $\mathfrak{c}(\mathfrak{s}, \mathbf{X}) = \mathfrak{c}(\mathfrak{s}, \mathbf{Y})$.

  ii. For all $i \in (1, n]$, $x_i \succ_S y_i$.

  then, the following must also be the case

  iii. $y_1 \succ_S x_1$.

- Not only is (SA) *unintuitive*, it is also *quite strong*. It entails *both* de Finetti's "additivity" (3) *and* (full) transitivity of $\succ_S$.

- I think the best way to grasp the content of (SA) is *via* the following illuminating theorem of Fishburn [4, Ch. 4].

  Theorem (Fishburn). (SA) is true *if and only if* there exists a mass function $\mathfrak{m}$ on $\mathcal{B}$ such that, for all propositions p and q in $\mathcal{B}$, the following *real-valued representation* holds:

  $$(\star) \qquad p \succ_S q \textit{ if and only if } \sum_{\mathfrak{s}_p \vDash p} \mathfrak{m}(\mathfrak{s}_p) > \sum_{\mathfrak{s}_q \vDash q} \mathfrak{m}(\mathfrak{s}_q).$$

  And, given de Finetti's axiom (2), there will always be a *probability* mass function $\mathfrak{m}$ satisfying ($\star$).

- Fishburn's Theorem reveals that (SA) *alone* ensures a real-valued representation ($\mathcal{R}_{\succ_S}$) of the $\succ_S$-ordering.
- Not only does this imply de Finetti's additivity axiom (3), but it also implies axiom (1) as well ($>_{\mathbb{R}}$ is a strict total order).
- Thus, once we have (SA) on board, the only axiom of de Finetti that can do *any* work is his axiom (2), which just ensures that $\mathcal{R}_{\succ_S}$ is a *probabilistic* representation of $\succ_S$.

[1] B. de Finetti, *The Theory of Probability*, Wiley, 1974.

[2] ———, *Foresight: Its Logical Laws, Its Subjective Sources*, in H. Kyburg and H. Smokler (*eds.*), *Studies in Subjective Probability*, Wiley, 1964.

[3] T. Fine, *Theories of Probability*, Academic Press, 1973.

[4] P. Fishburn, *The Axioms of Subjective Probability*, *Statistical Science*, 1986.

[5] J. Joyce, *Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief*, in F. Huber and C. Schmidt-Petri (*eds.*), *Degrees of Belief*, 2009.

[6] ———, *A Nonpragmatic Vindication of Probabilism*, *Philosophy of Science*, 1998.

[7] J.M. Keynes, *A Treatise on Probability*, MacMillan, 1921.

[8] N. Kolodny, *How Does Coherence Matter?*, *Proc. of the Aristotelian Society*, 2007.

[9] C. Kraft, J. Pratt and A. Seidenberg, *Intuitive Probability on Finite Sets*, *The Annals of Mathematical Statistics*, 1959.

[10] P. Maher, *Joyce's Argument for Probabilism*, *Philosophy of Science*, 2002.

[11] D. Miller, *The Accuracy of Predictions*, *Synthese*, 1975.

[12] ———, *Out Of Error: Further Essays on Critical Rationalism*, Ashgate, 2006.

[13] K. Popper, *Objective Knowledge: An Evolutionary Approach*, 2nd edition, 1979.

[14] F. Ramsey, *Truth and Probability*, 1926.

[15] L. Savage, *The Foundations of Statistics*, Dover, 1972.

[16] D. Scott, *Measurement Structures and Linear Inequalities*, *Journal of Mathematical Psychology*, 1964.