Williamson Chapter 10, Evidential Probability

Kevin Dorst
kmdorst@mit.edu

24.805, Fall 2024
Knowledge and its Limits and its Limits

I. Evidential Probabilities

Not chances, frequencies, or credences.

Close to 'rational credence, given your evidence'—but not quite.

Eg we have evidence that there are no ideal agents; so perhaps ideal agent couldn't have our evidence.

Reasonable 'ur-prior' π, conditioned on your total evidence.
$\rightarrow$ Your evidence (so evidential probabilities) varies across worlds.

Given E=K, your total evidence is your total knowledge.

Trivially, $\pi(e|e) = 1$, so your total evidence gets probability one. More carefully: the conjunction of everything entailed by your evidence (e) gets probability 1.

Not: the proposition that e is your total evidence. That often has probability less than 1, as we'll see.

You can lose knowledge, so probability 1 isn't a lifetime commitment.

- Forgetting: coin case.
- Knowledge defeat: urn case.

Why not E=B? TW:

> Given E = B, one can manufacture evidence for one's favourite theories by manipulating oneself into a state of certainty about appropriate propositions—for example, that one has just seen one's guru perform a miracle. That does not capture the spirit of the injunction to proportion one's belief to one's evidence.

We'll come back to this...

II. Modeling evidential probabilities

We can model knowledge with accessibility relations;
K_w is the conjunction of everything you know at w.

xRy iff $y \in K_x$

- Coin case.[1]
- Clock case.[2]

[1] K_w is partitional. Equivalently: R is reflexive, transitive, and symmetric.

[2] K_w is not partitional (not transitive)

We can model evidential probabilities with a (regular) prior π; probabilities at each world w from conditioning π on K_w: $P_w := \pi(\cdot|K_w)$.

When $K_w \neq K_x$, then $P_w \neq P_x$

- Coin case. $P_w(q)$ versus $\langle P(q) = 2/3 \rangle$
- Clock case. (About appearances, if you prefer.)

Posteriors are introspective.

Posteriors are not introspective.

<u>Higher-order uncertainty:</u>
Early = {1,2,3}. At 3, this is 2/3 probable.
But at 3, it's only 2/3-probable *that* it's 2/3-probable.

Where else? $\langle P(Early) = 2/3 \rangle = \{2, 3\}$.

Posteriors are not introspective

<u>Improbable Knowing:</u>
$K(Early) = \{2\}$.

So at 2, $P(Early) = 1$, but $P(P(Early) = 1) = 1/3$.

Self-effacing evidence:
Odd = {1,3,5,7,9,11}.
If *Odd* is true, $P(Odd) = 1/3$; if it's false, $P(Odd) = 2/3$.
So $\langle P(Odd) > 1/2 \rangle \leftrightarrow \neg Odd$ is true everywhere
So it's *known*: $P\Big(\langle P(Odd) > 1/2 \rangle \leftrightarrow \neg Odd\Big) = 1$

Even = ¬*Odd*.

III. Biased Inquiries

Simple creature. Either hot (h), cold (c), or popular (p).
Detector for ¬*hot*; detector for ¬*cold* .
No meta-detector to detect that a detector hasn't been activated.

Upshot:

- If h, knows not cold; leaves open h and p;
- If p, knows not cold nor hot—so knows that p;
- if c, knows not hot; leaves open c and p.

This is a 'sure-win' investigation: starts out $\pi(p) = 1/3$, and will definitely rise to either 1/2 or 1.

Diagram. S4 model.
In fact, S4.2 model (Stalnaker 2006).

Weird!

- Prior doesn't equal expected posterior.
- Prior wants to take a 2-to-1 bet against p; but knows that posterior will pay to call it off.

'Reflection' failure

$\frac{1}{2}(1) + \frac{1}{2}(-2) = -0.50$, so pay up to \$0.50

What to make of this?

- *TW: this is inevitable.*
  Only way to guarantee that $\pi(p) = \mathbb{E}_\pi(P(p))$ is if π is certain that K_w is partitional, i.e. introspective. But that's irrational.
- *TW: How can more knowledge predictably lead astray? Sometimes it can.*
  Inscribed: e_i = *die didn't land i*. Condition π on the 'inscribed truths'.
- *TW: Life is hard.*

> Moral: it is generally a mistake to try to align one's probabilities with what one knows about the results of conditionalizing them on truths with some given property. One instance of this mistake is to try to align our probabilities with what we know about the results of conditionalizing them on truths which we will know in the future. Although we may be made to suffer for the misalignment, it would not be rational to try to avert the suffering by changing our present beliefs. From our present perspective, **the non-partitional structure of our future knowledge is a source of bias**, similar in effect to forgetting although much subtler in its operation. Of course, we shall probably know more tomorrow, and it would be foolish then to disregard the new knowledge.[3] But we cannot take advantage of the new knowledge in advance. We must cross that bridge when we come to it, and accept the consequences of our unfortunate epistemic situation with what composure we can find. Life is hard.

Reject E=K? So long as Ep implies Kp, then anti-luminosity argument still shows that 'evidence entails that' isn't introspective. If it were, $Ep \Rightarrow EEp \Rightarrow KEp$, so Ep must be luminous. But it's not.

Q1: How to feel about argument for factivity of E, or against $E = B$, given this?

Q2: How feel about stipulations of 'no meta-detector'? We *do* have them. Coin case.
Compare: creature that can condition on H but not on $\neg H$?

[3] **Q3:** Would it be?