Higher-Order Uncertainty

Kevin Dorst
kmdorst@mit.edu

24.805, Fall 2023
Imprecision and Irrationality

I. Setup

Cases: Dice vs. socks; people vs. knees; chaotic vs. enigmatic urns.

Features: (1) insensitivity; (2) irresolute assessments; (3) fuzzy boundaries; (4) ambiguity aversion; (5) miscalibration/bias.

Rigidly-designated probability functions (π) vs. descriptions (P).

Question-Reflection vs. Reflection.
Latter drives classic Bayesian results: (i) value of evidence, (ii) no wishful thinking; (iii) convergence to the truth; and (iv) well-calibrated.

Reflection requires *partitional updates* (unlike train case) and *prior clarity* (unlike Williamsonian feeling-cold case).

Combined = Clear Bayes. Pretty much every time you see a Bayesian model, it's a Clear-Bayes model.

II. Clear Bayes can't model ambiguity

Clear Bayes implies introspection on priors and posteriors:

$$[P(q) = t] \to P(P(q) = t) = 1$$
$$[\mathbb{E}_P(X) = t] \to P(\mathbb{E}_P(X) = t) = 1$$

Hierarchical vs. higher-order uncertainty.

Your more-informed self vs. yourself.

Arguments:

1) Insensitivity: the usual.
2) Irresoluteness: why random if known?
3) Fuzzy boundaries: you know where comparisons give out. Eg if $P(q) = 0.4702$, then know that 472 is the first n such that more confident of $a_1 \vee ... \vee a_n$ than q.
4) Ambiguity aversion: either know you prefer betting on q, or know you prefer betting on $\neg q$.
5) Miscalibration: since you reflect your (current and) future opinions, you expect to be calibrated, and (if independent) confident roughly calibrated.

Fluctuating? Fix a time.
Q: Better argument?

What if certain about your probabilities but uncertain about your preferences/utilities? *Bold Claim:* that doesn't make a difference.

> Warm-up. The fire! Suppose secretly, deep down, you prefer money to sentimental value. But you have every reason to think otherwise—every reason to think you prefer the wedding album. What, in the subjectivist sense, should you do?

Take the wedding album! That's what maximizes *expected* value, even if your expectations are mistaken about what your values are.

Now take a case where you're uncertain (50-50) on whether your preferences are represented by U_1 or U_2. Still, you know what your *expected* values are—so you know what, given your uncertainty about what you value, you should do!
⇒ if uncertain about preferences over outcomes but know your credence function, you're still certain about your preferences over options, since those are driven by expected value.
⇒ epistemicist approach needs uncertainty about *probabilities*. Uncertainty about values alone won't do.

Source of uncertainty about $U(a)$ might be the outcome, or it might be how you feel about the outcome. Why matter?

Q: Is this right?
Subtle stuff about "interpersonal" comparisons of utilities and re-scaling.

III. Against imprecision

Insensitivity: ✓

Irresoluteness: ?

Yes for variance; but no explanation for *shape* of variance.

Fuzzy boundaries: ✗

Ambiguity aversion: ✓, on conservative decision rules.
But such rules give stark violations of the value of evidence: so long as you have beliefs about what you will do given imprecision, there is [I think?] *no* ambiguity-inducing update for which there's not *some* decision-problem such that you'll pay money to avoid the update.

Eg Roger's coin and 1:2 bets on p.

Miscalibration: sort of.
If imprecise about q_i, then can avoid being[1] confident that calibrated. But in situations where you're going to *gain* ambiguity (dilate) on the q_i, expect to be calibrated.

[1] determinately

More generally: (1) hard to *gain* ambiguity, and (2) imprecise view is implausible if paired with higher-order *certainty*—going to need higher-order uncertainty anyways. How far can we get with *just* HOU?

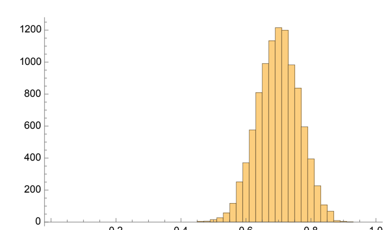
IV. Why higher-order? Noise

Why can't you just find out your credences by writing them down? Because the thing you write down might not be your actual credence.

Noise in cognitive system. Psychologists have known this forever.

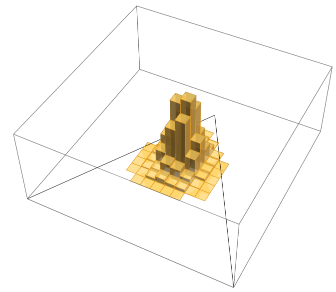
Q1: How can you have precise probabilities/dispositions, and it also be noisy?

- Generative model; credences = sampling dispositions.
  Elicit them by *drawing samples*. (This is how many cognitive scientists think the brain represents probabilities.)
- Think: urn in the head. Suppose 7 red and 3 black marbles. Then $C(red) = 0.7$ (precisely!), but distribution of elicitations when draw 50 samples is: [»]



- Why can't we just *look*? Or if not, why not draw *without replacement*? Because in reality a generative model is more like a computer program: can't be discerned by inspection, and drawing "without replacement" would involve altering it at each step.

Eg $\frac{\textit{RandomReal}[0,100]^{\textit{RandomInteger}[1,4])}}{\textit{RandomReal}[100,1000]}$
Mean? Median?
Sampling $\rightsquigarrow$ 13,000 Median? $\approx$ 10



- How does sampling work when distribution is over more than two possibilities, like a 3-place vector $\pi = (0.2, 0.6, 0.2)$? Sample from (w_1, w_2, w_3) at these rates. Drawing 50 samples and then counting proportions $\rightsquigarrow (0.16, 0.68, 0.16), (0.14, 0.62, 0.24), (0.14, 0.54, 0.32)$,...

Q2: How can you and I have precise probabilities, and you not know what *mine* are?

- Suppose you're unsure of both q and my credence in q, $P(q)$—i.e. the proportion of q-marbles in the urn in my head. How to represent your (noisy) dispositions? (Your credences, C)?
- You have an urn in your head; each marble is labeled with *both* $q/\neg q$, *and* with $P(q) = t$ for some t. (Maybe notecards, not marbles.)
- Eg

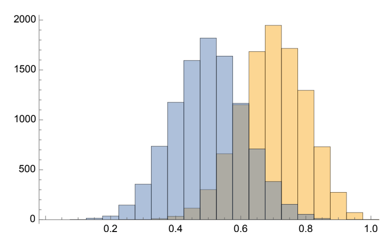
	q	$\neg q$
$P(q) = 0.7$	5	1
$P(q) = 0.5$	2	2

So $C(q) = 0.7$ and $C(P(q) = 0.7) = 0.6$.

Why? Marbles are outcomes in the sample space—answer all relevant Qs. You probably have multiple different small models (urns) that you combine; but if you combine them coherently into a single probability distribution, we can represent with a single urn.
And $C(q|P(q) = 0.7) = \frac{5}{6} \approx 0.83$, so Reflection fails.

- If you ask me my credence, I'll give you a number—but you'll still be unsure what my credence is, due to noise.

  Eg suppose you know I'll draw 20 samples and then announce the mean. I announce 0.65. You know this is about 16.8%-likely if I have $P(q) = 0.7$, and about 7.1% likely if I have $P(q) = 0.5$, so this provides some (*inconclusive*) evidence that I have $P(q) = 0.7$. (If you condition on it—changing the contents of your urn—you'll jump from $C(P(q) = 0.7) = 0.6$ to $C(P(q) = 0.7|\textit{said } 0.65) \approx 0.78$.) You're still uncertain about $P(q)$.



Q3: How can you have precise probabilities, and not know what *yours* are? The same way.

- Suppose you're unsure of both q and *your* credence in q, $C(q)$. How to represent your credences?
- You have an urn in your head; each marble is labeled with both q or $\neg q$, and with $C(q) = t$ for some t.
- Eg

	q	$\neg q$
$C(q) = 0.7$	5	1
$C(q) = 0.5$	2	2

So $C(q) = 0.7$ and $C(C(q) = 0.7) = 0.6$.

And $C(q|C(q) = 0.7) = \frac{5}{6} \approx 0.83$, so Reflection fails.

- You can try to figure out what you think by drawing some samples and announcing the verdict (writing a number down).
  Suppose you write down 0.65. This is *some* evidence that your credence was 0.7, but it's inconclusive—even if you condition on it, you're still unsure what your (updated) credence is.

Since boosts confidence in $C(q) = 0.7$, and $C(q|C(q) = 0.7) > C(q)$ (you trust yourself), sampling provides evidence for q. You're figuring out whether q by figuring out what you think about q!

Note: $C(q)$ is your credence *before* sampling and updating on it; we need a new description (like $C^+(q)$) for your credence after updating. In this case, $C^+(q) = C(q|\textit{said } 0.65)$.

Upshot: In any cognitive system in which credence elicitations are noisy—e.g. if they are generated via sampling from a generative model—then we should expect there to be higher-order uncertainty.

V. How Go Higher-Order?

Warm-up: epistemic logic; how have beliefs and be uncertain what your beliefs are?

$(W, \mathcal{B})$; believe q at w iff $\mathcal{B}_w \subseteq q$. Draw tree model.

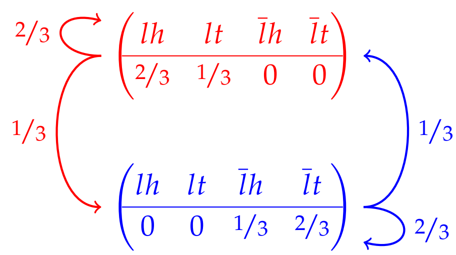
Probability frames: (W, P), where P_w rigidly designates your credence function at w.

Example: The Glance. Either it *looked headsy* (l) or it *looked tailsy* ($\bar{l}$). Conditional l, you're $\frac{2}{3}$-confident of *heads* (h): $P(h|l) = \frac{2}{3}$. Conditional on $\bar{l}$, you're $\frac{1}{3}$-confident of *heads*: $P(h|\bar{l}) = \frac{1}{3}$. You know that much.

The trouble is, you're not sure confident you are that it looked headsy. If it did look headsy, you're $\frac{2}{3}$-confident it did—in which case you're (e.g.) $\frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$-confident it both looked headsy and landed heads.

If it didn't look headsy, you're $\frac{1}{3}$-confident it did.

$$P = \begin{pmatrix} lh & lt & \bar{l}h & \bar{l}t \\ \hline 4/9 & 2/9 & 1/9 & 2/9 \\ 4/9 & 2/9 & 1/9 & 2/9 \\ 2/9 & 1/9 & 2/9 & 4/9 \\ 2/9 & 1/9 & 2/9 & 4/9 \end{pmatrix}$$



<u>Notice:</u>

1) Since how it looks determines your credences, but you're unsure how it looks, you're unsure what your credences are.
   If it *looks headsy*, $P(h) = \frac{5}{9}$. If it *looks tailsy*, $P(h) = \frac{4}{9}$.
   And if l, $P(l) = \frac{2}{3}$ while $P(\bar{l}) = \frac{1}{3}$.
   So if l, $P(P(h) = \frac{5}{9}) = \frac{2}{3}$, while $P(P(h) = \frac{4}{9}) = \frac{1}{3}$.

$\Rightarrow [P(h) = \frac{2}{3}] = \{lh, lt\} = l$, and $[P(h) = \frac{1}{3}] = \{\bar{l}h, \bar{l}t\} = \bar{l}$.

2) *Uncertainty all the way up:*
   And if $\bar{l}$, $P(P(h) = \frac{5}{9}) = \frac{1}{3}$, while $P(P(h) = \frac{4}{9}) = \frac{2}{3}$.
   So, if l, $P(P(P(h) = \frac{5}{9}) = \frac{2}{3}) = \frac{2}{3}$. And so on all the way up.

$\Rightarrow [P(P(h) = \frac{5}{9}) = \frac{2}{3}] = [P(l) = \frac{2}{3}] = l$

3) Due to noise, sampling from your opinions (writing down a number) would reduce but not eliminate this higher-order uncertainty.

Unless took infinitely many samples.

4) *Reflection fails:* $P(h|P(h) = \frac{5}{9}) = \frac{2}{3}$.
   Why? Because learning what your opinions are gives you *new information*—it tells you how things looked!

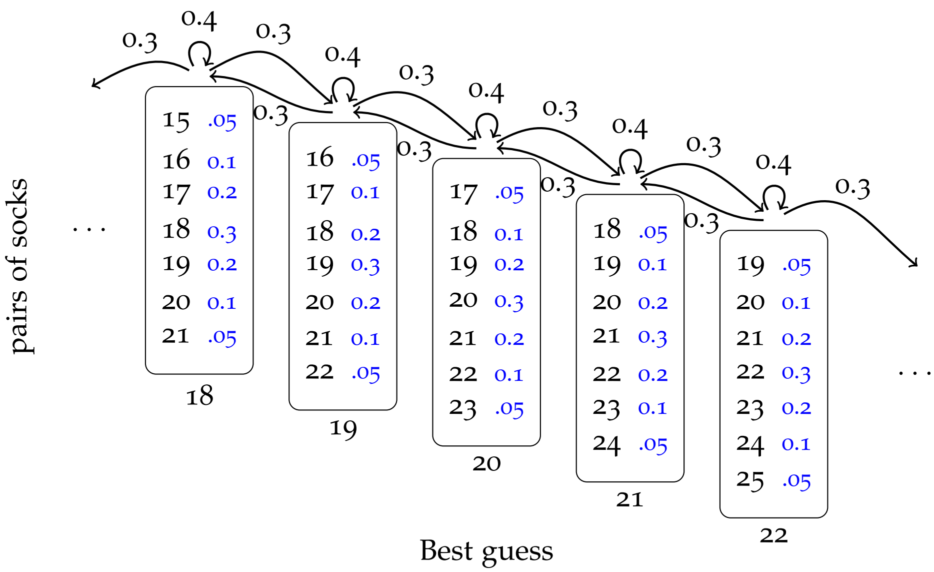
You were $\frac{5}{9}$ only because you were unsure of that fact.

5) *New Reflection holds:* $P(q|P = \pi) = \pi(q|P = \pi)$.
   Conditional on having a given set of opinions, you adopt the opinions *they* would have upon learning as much.

That's why we can "factorize" the frame into cells of"informed" probabilities with weights between them.

VI. Getting our features

Take a simplified version of the socks case: you're unsure what your best guess is:



Numbers in cells = worlds

Blue numbers within cell C represent the *conditional* probabilities of each world, given cell C. (All worlds agree on these conditional probabilities.)

Arrow labeled t from cell x to cell y says: if in cell x, you're t-confident you're in cell y.

⇒ Probabilities of world w in cell C = probability of C (arrow) × probability of w given C (blue number).

Insensitivity: Sort of.
Not strictly, since P (hence $\mathbb{E}_P$) is always precise. So if $\mathbb{E}_P(X) = \mathbb{E}_P(Y)$, then $\mathbb{E}_P(X+1) > \mathbb{E}_P(Y)$.
But implies that the relation "you doubt that your estimate for A is higher than your estimate for B" is intransitive.
E.g. X = *pairs of socks*, Y = *number of 1–4 rolls*[2]. Say best guess is 20.
Then $P(\mathbb{E}_P(X) > \mathbb{E}_P(Y)) = 0.3$,[3] and
$P(E_P(X+1) > \mathbb{E}_P(Y)) = 0.7$; but
$P(E_P(X+1) > \mathbb{E}_P(X)) = 1$.
Also explains "graded incomparability": why you become steadily more confident in your comparisons as Caspar gives me more socks.

Mimic negative intransitivity.

[2] So $P(\mathbb{E}_P(Y) = 20) = 1$

[3] and $P(\mathbb{E}_P(X) < \mathbb{E}_P(Y)) = 0.3$

If made uncertainty about estimates wider, could make jump in probability arbitrarily small

Irresoluteness: ✓
Comes from sampling/noise interpretation.

Might also get directly: if uncertain about your credences, might misname them. $C(q) \neq \mathbb{E}_C(C(q)) \neq \mathbb{E}_C(\mathbb{E}_C(C(q))) \neq \cdots$

Fuzzy boundaries: ✓
If best guess is 20, then "the greatest integer n such that your estimate is higher than n" is in fact 19. But you don't know that—you assign it 40% probability. You're 30%-confident that the answer is 18 and 30%-confident that it's 20.
Likewise generally. If n *is* 18, then wonder whether n is 17, 18, or 19.

With wider uncertainty bands, we could make these probabilities arbitrarily low.

Ambiguity Aversion. Bleh.
If think of decision as "Bet1, or instead Bet2 or Bet3—whichever looks best", and don't trust your judgments, then can be ambiguity averse.
At least this model of ambiguity *allows* the value of evidence.

E.g. in the *looks-headsy* model.

Miscalibration: ✓
Since Reflection will always fail, you generally won't expect your ambiguous judgments to be calibrated.