Upper & Lower Probabilities

Jonathan Weisberg
University of Toronto

Northern Institute of Philosophy
June 14, 2010

Outline

Motivations

Motivating Sets of Measures

Why should partial belief states be represented by sets of probability measures?

- Gaps in our probabilistic judgments
- Empirical evidence about our decision habits
- Gaps in our knowledge about statistics/chances
- Underdetermination by preferences
- Peirce + Carnap = Levi.

Gaps in Probabilistic Judgment

van Fraassen (2006)

"Rain does seem more likely to me than not, but does it seem more than π times as likely as not?"

Jeffrey (1987)

"Can you judge propositions A, B to be irrelevant to each other without having any particular judgmental probabilities for them or their conjunction, AB?"

Jeffrey (1983)

"…data about the orbits of the moons of Uranus are irrelevant to Mendelian genetics, positively relevant to Newtonian dynamics, and positively relevant, but less so, to the conjunction of the two theories."

Gaps in Probabilistic Judgment

A natural way to represent these states of opinion is the set of probability measures that obey just the constraints stated.

van Fraassen (2006)

$$\mathcal{P} = \{p : p(\mathrm{R}) > .5\}$$

Jeffrey (1987)

$$\mathcal{P} = \{p : p(AB) = p(A)p(B)\}$$

Jeffrey (1983)

$$\mathcal{P} = \{p : p(M|D) - p(M) = 0 < p(MN|D) - p(MN) < p(N|D) - p(N)\}$$

Decision Habits: Empirical Evidence

The Ellsberg Problem (Ellsberg 1961; cf. Knight 1921)

An urn contains 30 red balls and 60 balls in some unknown mixture of black and yellow. A ball is chosen randomly.

	Red	Black	Yellow
A	£1	0	0
B	0	£1	0

	Red	Black	Yellow
C	£1	0	£1
D	0	£1	£1

- Most subjects prefer $A \succ B$ and $D \succ C$, but no probability function yields this combination.
- The set $\{p : p(R) = 1/3\}$ is consistent with this combination.

A Simpler Example

Halpern's Ellsberg Problem (Halpern 2003)

A bag contains 100 marbles; 30 are known to be red, and the remainder are known to be either blue or yellow. A ball is chosen randomly. Rank the following three options:

A: £1 if the marble is red, 0 otherwise.

B: £1 if the marble is blue, 0 otherwise.

C: £1 if the marble is yellow, 0 otherwise.

- Most subjects prefer $A \succ B$ and $A \succ C$, but $B \sim C$.
- So their probabilities are $p(B) = p(C) = .35$.
  - But then why don't they prefer $B \succ A$?
- The set $\{p : p(R) = .3\}$ is consistent with these preferences.

Gaps in Our Knowledge of Statistics/Chances

According to many authors (Fisher, Neyman, Kyburg), probability judgments should be grounded in knowledge of objective facts about statistics and/or chances.

- If all you know is that Peter is a Swede and 80% of Swedes are Protestants, you may conclude with 80% certainty that Peter is a Protestant.
- But if all you know is that he is a Swede, you may assign no probability to his being a Protestant.
- The most you can say is that this probability is between 0 and 1.
- A natural representation of this state of opinion is $\{p : p(S) = 1, 0 \leq p(P) \leq 1\}$, i.e. $\{p : p(S) = 1\}$.

Underdetermination by Preferences

According to many authors (Ramsey, Savage, Maher), credences are determined by preferences via a representation theorem.

- Given a preference structure satisfying certain rationality constraints, a unique probability function is determined.

But sometimes, preferences may not determine a unique probability function:

- The Jeffrey-Bolker theorem does not yield a unique probability function.
- Savage's theorem does yield a unique probability function, but only given a "complete" preference structure.
  - A rational agent's preferences may be "fragmentary", i.e. they may not be "complete" (Jeffrey 1987, Joyce 1999).

In such cases, the set of probability functions consistent with the agent's preferences is a natural representation of her credences.

Peirce + Carnap = Levi

The history of Bayesianism according to Levi (1974):

- Carnap tried to find the unique, logically mandated confirmational commitment, but he failed.
- Bayesians responded by saying that any confirmational commitment is reasonable, provided you pick one and stick to it.
- But Peirce pointed out that such arbitrary dogmatism is unwarranted.
- So better to be "ambivalent" over the set of confirmational commitments compatible with your knowledge.

Upper & Lower Probabilities

Definition: *Upper & Lower Probabilities*

$\mathcal{P}$'s upper and lower probabilities are:

$$\begin{aligned} \mathcal{P}_*(A) &= \inf\{p(A) : p \in \mathcal{P}\} \\ \mathcal{P}^*(A) &= \sup\{p(A) : p \in \mathcal{P}\} \end{aligned}$$

Motivation? $\mathcal{P}$ is often "interval-valued" due to:

- Gaps in...
  - our judgments.
  - our statistical knowledge.
  - our preferences.
- The convexity requirement.

When $\mathcal{P}$ is interval-valued, we can think of $\mathcal{P}_*(A)$ as your supremum buying price for a £1 bet on A.

- You should be willing to pay any price below $\mathcal{P}_*(A)$.

Similarly, $\mathcal{P}^*(A)$ is your infimum buying price.

Common Gaps

Why would $\mathcal{P}(A)$ be the interval from $\mathcal{P}_*(A)$ to $\mathcal{P}^*(A)$?

- Often our judgments can only put upper/lower bounds on a probability.
  - I'm 80-90% sure I'll get the promotion.
  - Q: But does your credence actually span that interval?
- Often our statistical knowledge only places upper/lower bounds.
  - The Ellsberg problems are examples.
  - In general, partial knowledge of proportions often yields only upper/lower bounds.
- Underdetermination by preferences.
  - The Jeffrey-Bolker theorem generates interval-valued constraints on credences.
  - Even on Savage's theorem, fragmentary preferences generate interval-valued constraints in many cases, e.g. Ellsberg.

The Convexity Requirement

Definition: Convex Set of Probabilities

$\mathcal{P}$ is convex iff whenever $p, q \in \mathcal{P}$, so is $\alpha p + (1-\alpha)q$ for every $\alpha \in [0,1]$.

Some, especially Levi (1974, 1980), require that $\mathcal{P}$ be convex.

- If $\mathcal{P}$ is convex, $\mathcal{P}(A)$ is always interval-valued.

Why require convexity?

- Levi (1980): if we are undecided between the states of opinion represented by p and by q, we should be open to any compromise between them.

Why not?

- Jeffrey (1987): some intuitively appealing sets are not convex.
- The set $\{p : p(AB) = p(A)p(B)\}$ represents the opinion merely that A and B are independent, but is not convex.

Beware of Underrepresentation

Many sets can have the same upper & lower probabilities:

- Let $\{A, B, C\}$ be a partition and let

$$\begin{aligned} \mathcal{P} &= \{p : p(A) \leq 1/2, p(B) \leq 1/2\} \\ \mathcal{Q} &= \{p : p(A) = p(B)\} \end{aligned}$$

  Then $\mathcal{P}_* = \mathcal{Q}_*$ but $\mathcal{Q} \subset \mathcal{P}$.
- Even two convex sets can even have the same upper & lower probabilities.

Couso, Moral, and Walley (1999) argue that a set of probabilities is behaviourally equivalent to its convex hull, but...

- That depends on our decision theory.
- As Halpern (2003) points out, there are other interesting ways $\mathcal{P}$ can "say more" than its convex hull:
  - Jeffrey's (1987) example treats A and B as independent, but its convex hull includes measures that do not.

Statics

Basic Properties

Some elementary properties of upper and lower probabilities:

1. $\mathcal{P}_*(\Omega) = \mathcal{P}^*(\Omega) = 1$
2. $\mathcal{P}_*(\emptyset) = \mathcal{P}^*(\emptyset) = 0$
3. $\mathcal{P}_*(A) = 1 - \mathcal{P}^*(\overline{A})$
4. $\mathcal{P}_*(A \cup B) \geq \mathcal{P}_*(A) + \mathcal{P}_*(B),\ A \cap B = \emptyset$
5. $\mathcal{P}^*(A \cup B) \leq \mathcal{P}^*(A) + \mathcal{P}^*(B),\ A \cap B = \emptyset$
6. $\mathcal{P}_*(A \cup B) \leq \mathcal{P}_*(A) + \mathcal{P}^*(B) \leq \mathcal{P}^*(A \cup B),\ A \cap B = \emptyset$

Notes:

- Analogues of the inclusion-exclusion rule do not hold!
- (1)–(4) do not entail (5).
- (1)–(5) entail (6).[1]
- (1)–(6) do not axiomatize the theory.

[1] Note: this is contra (Halpern 2003: 31, 58).

Axiomatizations: History

Anger & Lembcke's (1985) axiomatization appears to be the earliest.

- There are earlier axiomatizations in terms of upper and lower *expectations*; the earliest appears to be (Lorentz 1952).
- See (Walley 1991) for a simple and accessible axiomatization in terms of expectations.
- See (Halpern & Pucella 2002) for further references and an excellent discussion of direct axiomatization and its complexity.

Axiomatization: Definition of Covering

The axiomatization requires the following definition:

Definition: *Covering*

A set $\mathcal{A}$ of subsets of Ω *covers A exactly k times* iff every element of A is in exactly k sets in $\mathcal{A}$.

Examples:

- The set $\{\{1, 3, 5\}, \{2, 4, 6\}\}$ covers the set $\{3, 4\}$ exactly once.
- The set $\{\{Alfred, Bonnie, Clyde\}, \{Bonnie, Clyde\}\}$ covers the set $\{Bonnie\}$ exactly twice.
- Exactly how many times does the same set cover $\{Alfred, Bonnie\}$? Undefined.

The Axiomatization

Covering Axiom

If $\mathcal{A} = \{A_1, \ldots, A_k\}$ covers A exactly $m+n$ times and covers $\overline{A}$ exactly m times, then $\sum_{i=1}^{k} \mathcal{P}_*(A_i) \leq m + n\mathcal{P}_*(A)$.

Soundness Theorem

Every set of probability functions $\mathcal{P}$ is such that $\mathcal{P}_*$ satisfies the Covering Axiom and $\mathcal{P}_*(\Omega) \geq 1$.

Completeness Theorem

If f satisfies the Covering Axiom and $f(\Omega) \geq 1$, then $f = \mathcal{P}_*$ for some set of probability measures $\mathcal{P}$ on Ω.[2]

[2] Note: an extra axiom is required if we assume countable additivity.

Dynamics

Dynamics: Obvious Extensions

How should one update a set of probability functions?

- Standard answer: update each member in the usual way and gather up the results.

First the obvious definition:

Definition: *Conditional Probability-Sets*

$\mathcal{P}(\cdot|E) = \{p(\cdot|E) : p \in \mathcal{P} \text{ and } p(E) > 0\}$

Then our rule for updating on propositional evidence is:

Classical Conditioning

If your epistemic state is represented by $\mathcal{P}$ and you receive evidence E, your new epistemic state should be $\mathcal{Q}(\cdot) = \mathcal{P}(\cdot|E)$.

Dynamics: More General Extensions

We can extend more general updating rules from standard probability theory in the same way.

Jeffrey Conditioning

If you receive uncertain evidence in the form of probabilities $q(E_i)$ over the partition $\{E_i\}$, your new probability set should be $\mathcal{Q} = \{q : q = \sum_i p(\cdot|E_i)p(E_i), p \in \mathcal{P}\}$.

Information Minimization

If you receive evidence mandating a probability function in the set S, adopt the probability set:

$$\mathcal{Q} = \{q : q \in S \text{ and } q \text{ minimizes } \sum_{i \in \Omega} p_i \log\left(\frac{p_i}{q_i}\right) \text{ for some } p \in \mathcal{P}\}$$

where $p_i = p(\{i\})$ and similarly for q_i.

Two Problems in Dynamics

We'll look at two problems in dynamics, one more formal and the other more philosophical.

1. Updating *on* sets of probabilities: when your evidence comes in the form of a set of probability measures, how should you update?
   - May be an open problem, I'm not sure.
2. Dilation: extending conditionalization to sets of probabilities in the above way seems to have very unintuitive results.
   - Increasingly receiving attention from philosophers.
   - Appears to be a severe problem.

Updating on Sets of Probabilities

How should you update on a set of probabilities?

- A legitimate question, at least if the questions that motivate Jeffrey Conditioning and Infomin are legitimate.

But the answer isn't obvious.

- Proposal: by analogy with the classical rules, treat the evidence as a constraint and then minimize the distance from your current credences.
  - Trivializes the update.
- Proposal: take $\mathcal{P}_*$ and $\mathcal{E}_*$ and combine them like belief functions, using Dempster's rule.
  - Not every $\mathcal{P}_*$ is a belief function; Dempster's won't always apply?
  - Lower probabilities underdetermine the sets that generate them; will information be lost?

A Crude Proposal

Here's a proposal that at least avoids the above worries: treat the members of $\mathcal{P}$ and $\mathcal{E}$ as Dempster-Shafer mass functions.

- All probability functions are mass functions (Wednesday).
- So the members of $\mathcal{P}$ and $\mathcal{E}$ are mass functions.
- So $p \oplus e$ is defined for all $p \in \mathcal{P}, e \in \mathcal{E}$ (Wednesday).
- So let $\mathcal{Q} = \{q : q = p \oplus e \text{ where } p \in \mathcal{P}, e \in \mathcal{E}\}$.

Dilation

Dilation is a problematic phenomenon that arises when conditioning sets of probabilities. What is dilation?

Definition: *B dilates A*

B dilates A iff $[\mathcal{P}_*(A), \mathcal{P}^*(A)] \subset [\mathcal{P}_*(A|B), \mathcal{P}^*(A|B)]$.

- Intuitively, B spreads one's opinion about A.

Definition: *$\mathcal{B}$ dilates A*

A partition $\mathcal{B} = \{B_i\}$ dilates A iff for all B_i, $[\mathcal{P}_*(A), \mathcal{P}^*(A)] \subset [\mathcal{P}_*(A|B_i), \mathcal{P}^*(A|B_i)]$.

- Intuitively, one's opinion about A will spread no matter what the truth about $\mathcal{B}$ turns out to be.

Dilation: Examples

Seidenfeld & Wasserman (1993)

We're going to flip a fair coin twice, but the flips may not be independent. Your credences are naturally represented by:

$$\mathcal{P} = \{p : p(H_1) = p(H_2) = 1/2,\ p(H_1 \wedge H_2) = x, 0 \leq x \leq 1/2\}$$

It follows that

$$[\mathcal{P}_*(H_1), \mathcal{P}^*(H_1)] = [1/2, 1/2]$$

$$[\mathcal{P}_*(H_1|H_2), \mathcal{P}^*(H_1|H_2)] = [0, 1]$$

$$[\mathcal{P}_*(H_1|\overline{H_2}), \mathcal{P}^*(H_1|\overline{H_2})] = [0, 1]$$

So the partition $\{H_2, \overline{H_2}\}$ dilates H_1.

Dilation: Examples

van Fraassen (2006)

Peter believes that his grade will go up after tomorrow's exam (G). Nothing in his opinion bears on whether the possibility of rain (R) is relevant to his grades. So

$$\mathcal{P} = \{p : p(G) > c\}$$

Thus we have

$$[\mathcal{P}_*(G), \mathcal{P}^*(G)] = [c, 1]$$

$$[\mathcal{P}_*(G|R), \mathcal{P}^*(G|R)] = [0, 1]$$

$$[\mathcal{P}_*(G|\overline{R}), \mathcal{P}^*(G|\overline{R})] = [0, 1]$$

So the partition $\{R, \overline{R}\}$ dilates G.

Dilation: Extent of the Phenomenon

How pervasive is dilation?

- Seidenfeld and Wasserman (1993) provide an extensive and careful disucssion of sufficient and necessary conditions.
  - But the presentation is highly technical, and the punchline gets lost in a welter of notation.
- Seidenfeld (1994) provides a plainer statement of a striking result:
  - Dilation always occurs, unless $\mathcal{P}$ is a singleton or there are no independent propositions.
- Still, van Fraassen's (2006) discussion provides simpler and more direct insight into the cause and pervasiveness of dilation.
  - So we'll follow his presentation.

Dilation: An Elementary Result

Claim: If $\mathcal{P} = \{p : p(A) > a\}$ and A is logically independent of B, then $[\mathcal{P}_*(A|B), \mathcal{P}^*(A|B)] = [0, 1]$.

Proof. Because A, B are logically independent, there is an $r \in \mathcal{P}$ such that $r(B) = 0$.

Now let q be any probability function s.t. $q(B) = 1$. If $q \in \mathcal{P}$, then $q \in \mathcal{P}(\cdot|B)$.

If $q \notin \mathcal{P}$, there are weights a, b such that $p =_{df} aq + br$ is in $\mathcal{P}$; by weighting b heavily enough, we can always bring the weighted average of q and r as close to r as we like, so that $p(A) > a$.

Dilation: An Elementary Result (continued)

Finally, notice that $q(\cdot) = p(\cdot|B)$, since

$$\begin{aligned} p(\cdot|B) &= p(\cdot \wedge B)/p(B) \\ &= \frac{aq(\cdot \wedge B) + br(\cdot \wedge B)}{aq(B) + br(B)} \\ &= \frac{aq(\cdot \wedge B)}{aq(B)} \\ &= q(\cdot|B) \\ &= q(\cdot) \end{aligned}$$

So $q \in \mathcal{P}(\cdot|B)$.

Thus $q \in \mathcal{P}(\cdot|B)$ whenever $q(B) = 1$, i.e. $\mathcal{P}(\cdot|B) = \{p : p(B) = 1\}$.

Dilation: An Elementary Result (continued)

Corollary: if $\mathcal{P} = \{p : p(A) > a\}$ and A is logically independent of B, then $\{B, \overline{B}\}$ maximally dilates A:

$$[\mathcal{P}_*(A|B), \mathcal{P}^*(A|B)] = [0, 1]$$

$$[\mathcal{P}_*(A|\overline{B}), \mathcal{P}^*(A|\overline{B})] = [0, 1]$$

What's going on here intuitively?

- $\mathcal{P}$ is agnostic about the presence/absence of a correlation between A and B.
- So it contains every possible view: from irrelevance to maximal correlation to maximal anti-correlation.
- Conditionalizing each view on B (or $\overline{B}$) covers the range from certainty in A to certainty in $\overline{A}$.

Dilation: The Philosophical Challenge

How serious a problem is dilation? There are at least four ways to press the problem.

- Irrelevant Relevance
- Sturgeon's Character-Matching Thesis
- Violations of Reflection
- White's Coin Puzzle

Irrelevant Relevance

In many cases of dilation, learning seemingly irrelevant information has drastic effects on your opinion.

- Seidenfeld & Wasserman's coin-flip case may not be so bad.
  - Information about one flip affects your opinion about another. . . but they may be related, for all you know.
- Van Fraassen's grade example looks worse.
  - Why should learning that it will rain effect Peter's opinion about whether his grade will improve?
- But Peter's example may be misleading: *we* certainly regard the weather as irrelevant to his grade, but he doesn't. He is agnostic.

Still, it's troubling that agnosticism about relevance always results in your opinion being wiped out by the new information.

Sturgeon's Character-Matching Thesis

This brings us to an idea put forward in (Sturgeon 2008):[3]

Character-Matching Thesis

"Evidence and attitude aptly based on it must match in character."

- If you learn that 1/3 to 1/2 of the balls in the urn are black, your credence should span $[1/3, 1/2]$.
- But if you learn the precise fact that it will rain tomorrow, that should not spread your confidence across $[0, 1]$.

Wheeler (manuscript) challenges the CMT:

- Given imprecise information that a patient more likely drank from bottle A than B, a doctor forms the precise opinion that treatment X is preferable, since $EU(X)$ is maximal on any p such that $p(A) > p(B)$.

[3] See also (Walley 1991: 212-3).

Violations of Reflection

The Reflection Principle (van Fraassen 1984, 1995)

Your current opinion should match the opinions you think you may come to have in the future.

In particular, if you know that you will believe A tomorrow, you should believe A today.

- When $\mathcal{B}$ dilates A, you can know that your opinion tomorrow will span $[0, 1]$.
- So it should do so today.
- But then Peter should be agnostic about his grade today.

Note: while the Reflection Principle has many counterexamples (memory loss, cognitive mishap, etc.), none of them help here.

White's Coin Puzzle

White's Coin Puzzle (White, 2009)

"You haven't a clue as to whether A. But you know that I know whether A. I agree to write 'A' on one side of a fair coin, and '$\overline{A}$' on the other, with whichever one is true going on the Heads side. (I paint over the coin so that you can't see which sides are heads and tails). We toss the coin and observe that it happens to land on 'A'."

The following four conditions should hold:

1. $\mathcal{P}(A) = [0, 1]$
2. $\mathcal{P}(H) = [1/2]$
3. $\mathcal{P}(A|\text{'}A\text{'}) = \mathcal{P}(A)$
4. $\mathcal{P}(H|\text{'}A\text{'}) = \mathcal{P}(H)$

White's Coin Puzzle (continued)

Those four conditions entail $\mathcal{P}(A|\text{`}A\text{'}) = [0, 1]$.

- From (3), every $p \in \mathcal{P}$ has $p(A|\text{`}A\text{'}) = p(A)$, so conditioning leaves the assignments to A unchanged.

They also entail $\mathcal{P}(H|\text{`}A\text{'}) = [1/2]$.

- From (4), every $p \in \mathcal{P}$ has $p(H|\text{`}A\text{'}) = p(H)$, so conditioning leaves the assignments to H unchanged.

But it should be that $\mathcal{P}(H|\text{`}A\text{'}) = \mathcal{P}(A|\text{`}A\text{'})$.

- You know that 'A' is on the heads side iff A, and that it landed on 'A'.
- So you know that H iff A.
- So your attitude to these propositions should be the same.

White's Coin Puzzle (continued)

Apparently, you must dilate your credence in H to $[0, 1]$.

- This requires that we give up (4).
- So we must be agnostic about whether 'A' is relevant to H.
- But you know that the objective chance of H is exactly $1/2$!
- You even know that the objective chance of H given 'A' is $1/2$!

See White (2009) for four more objections to dilating on H.

Decisions

Expected Values for Sets of Probabilities

Definition: Expected Value (Set of Probabilities)

$$E_{\mathcal{P}}(X) = \{v : E_p(X) = v,\ p \in \mathcal{P}\}$$

Example: you know a coin has a heads-bias of either 1/3 or 3/4. What is the expected value of B, a £12 stake on heads?

- $\mathcal{P} = \{p, q\}$ where $p(H) = 1/3$, $q(H) = 3/4$.
- $E_p(B) = 3$, $E_q(B) = 9$.
- So $E_{\mathcal{P}} = \{3, 9\}$.

Upper & Lower Expectations

Definition: Upper & Lower Expected Values

$\underline{E}_{\mathcal{P}}(X) = \inf(E_{\mathcal{P}}(X))$
$\overline{E}_{\mathcal{P}}(X) = \sup(E_{\mathcal{P}}(X))$

Example: in the preceding example $\underline{E}_{\mathcal{P}}(X) = 3$ and $\overline{E}_{\mathcal{P}}(X) = 9$.

Example: in Halpern's Ellsburg problem, what are the upper and lower expected values of a £10 stake on red? Blue?

- Recall that $\mathcal{P} = \{p : p(R) = .3\}$.
- So $\underline{E}_{\mathcal{P}}(R) = \overline{E}_{\mathcal{P}}(R) = .3$.
- $\underline{E}_{\mathcal{P}}(B) = 0$ and $\overline{E}_{\mathcal{P}}(B) = .7$.

Question: which bet is better, R or B?

Decision Rules: Total Domination

Total Domination (cf. Hajek 2000)

If $\underline{E}_{\mathcal{P}}(A) > \overline{E}_{\mathcal{P}}(B)$, you should prefer A to B.

Total Domination may be true as a sufficient condition, but it seems badly incomplete (Weatherson, manuscript):

- Suppose $\underline{E}_{\mathcal{P}}(A) < \overline{E}_{\mathcal{P}}(B)$ and $\underline{E}_{\mathcal{P}}(A) > \underline{E}_{\mathcal{P}}(B)$.
- Let A' be a bet that always pays whatever B does, plus d, where $\underline{E}_{\mathcal{P}}(B) + d = \underline{E}_{\mathcal{P}}(A)$.
- Clearly A' is preferable to B.
- But Total Domination does not require preferring A':

$$\underline{E}_{\mathcal{P}}(A') = \underline{E}_{\mathcal{P}}(A) < \overline{E}_{\mathcal{P}}(B).$$

Decision Rules: Maximinex

Maximinex (Gilboa & Schmeidler 1993)

If $\underline{E}_{\mathcal{P}}(A) > \underline{E}_{\mathcal{P}}(B)$, you should prefer A to B.

Has bad dynamic consequences (Weatherson, manuscript):

- Let $\mathcal{P} = \{p : p(A) = 1/2\}$.
- So you should pay £.40 for a £1 bet on A.
  - $\underline{E}_{\mathcal{P}}(\text{Buy for £.40}) = .1$
  - $\underline{E}_{\mathcal{P}}(\text{Stick}) = 0$.
- Now suppose you learn B or $\neg B$, dilating your credence in A to $[0, 1]$.
- Now you will sell the bet for £.01.
  - $\underline{E}_{\mathcal{P}}(\text{Sell for £.01}) = .01$
  - $\underline{E}_{\mathcal{P}}(\text{Stick}) = 0$.

Decision Rules: Maxi

Maxi

If $\underline{E}_{\mathcal{P}}(A) > \underline{E}_{\mathcal{P}}(B)$ and $\overline{E}_{\mathcal{P}}(A) > \overline{E}_{\mathcal{P}}(B)$, you should prefer A to B.

Maxi has similarly bad dynamic consequences. There are reasonable probability sets with the following features (ibid):

- Maxi recommends buying a bet on A at a certain price.
- Learning B or $\neg B$ then dilates your credence in A.
- Maxi then recommends selling the bet on A for less the amount you bought it for.

Decision Rules: Supervaluationism

Supervaluationism

If $E_p(A) \geq E_p(B)$ for all $p \in \mathcal{P}$, then you should have $A \succeq B$.

How satisfying supervaluationism is depends on your reasons for adopting sets of probabilities.

- If the underdetermination of credences by preferences was your motivation, seems pretty satisfactory.

But if you prefer to think of credences as the guide to life, supervaluationism is a pretty weak guide.

- Supervaluationism looks weak even if you were motivated by the desire to accommodate typical Ellsberg preferences.
  - Typical Ellsberg preferences are compatible with, *but not predicted by*, supervaluationism.

Reminder: Ellsberg Problems

The Classic Ellsberg Problem

30 red balls, 60 black or yellow in unknown proportion.

	Red	Black	Yellow
A	£1	0	0
B	0	£1	0

	Red	Black	Yellow
C	£1	0	£1
D	0	£1	£1

Halpern's Variation

30 red marbles, 70 blue or yellow in unknown proportion.

A: £1 if the marble is red, 0 otherwise.

B: £1 if the marble is blue, 0 otherwise.

C: £1 if the marble is yellow, 0 otherwise.

DUI Supervaluationisms

Maximin Supervaluationism (Levi 1985, 1986)

Do act A if it (a) maximizes E_p relative to some $p \in \mathcal{P}$ and (b) maximizes $\underline{E}_{\mathcal{P}}$ amongst those acts that satisfy (a).

In other words: find those acts that are contenders for maximizing expected utility, then maximize $\underline{E}_{\mathcal{P}}$.

- Predicts typical preferences in the Ellsberg problem:
  - $\underline{E}_{\mathcal{P}}(A) = .3 > \underline{E}_{\mathcal{P}}(B) = 0$, so $A \succ B$.
  - $\overline{E}_{\mathcal{P}}(D) = .7 > \overline{E}_{\mathcal{P}}(C) = .3$, so $D \succ C$.
- Also predicts the preferences in Halpern's version:
  - $\underline{E}_{\mathcal{P}}(A) = .3$ and $\underline{E}_{\mathcal{P}}(B) = \underline{E}_{\mathcal{P}}(C) = 0$.
  - So $A \succ B$, $A \succ C$, and $B \backsim C$.

Obviously, other rules for decision under ignorance (DUI) yield other DUI supervaluationisms.

Levi & Independence of Irrelevant Alternatives

But this proposal contradicts a well-known principle:

Independence of Irrelevant Alternatives

If $A \succ B$ given a choice between A and B, it should not be that $B \succ A$ given a choice between A, B, and C.

An urn contains black and white balls in unknown proportions from 40% to 60% black; one will be chosen randomly.

Choice 1. Win £55 if black and lose £45 if white, or abstain.

Choice 2. Win £55 if white and lose £45 if black, or abstain.

Choice 3. Take the gamble from Choice 1, the gamble from Choice 2, or abstain — choose exactly one.

Levi's proposal recommends abstaining in Choices 1 & 2, but taking the first gamble in Choice 3. (Levi 1997)

- See (Weatherson, manuscript) for further discussion.

Making It Up As You Go

Jeffrey (1987): faced with the first Ellsberg choice, you may choose either way. But once you choose, your second choice may be forced.

- Once $A \succ B$, you commit yourself to $p(R) > p(B) < p(Y)$, and thus to $C \succ D$.

Levi (1985): your first choice doesn't force your second choice.

- "…for the agent does not categorically prefer A over B".

What's at issue appears to be whether the "gaps" in your probabilities/preferences get filled in as you work through a sequence of choices.

Making It Up As You Go (continued)

What is the source of this disagreement?

- According to Jeffrey, you do/should adopt a probabilistic commitment and then make your choice.
  - "I think you do well to find a definite probability function to express your uncertainty, if you can. Of course, if there is no need for definiteness and it is hard to attain, the game is not worth the candle. But in the Ellsberg problems [. . .] I think I would try to express my uncertainty via a single probability assignment — the uniform one, I imagine."
- According to Levi, you need not alter your attitudes, but merely execute a decision rule.
  - Neither $\mathcal{P}$ nor $\succeq$ change as you work through the problem, you just use maximin supervaluationism to choose.

Making It Up As You Go (continued)

Only on Levi's picture is typical Ellsberg behaviour permissible. But is the picture sensible?

- If he restricts himself to bare supervaluationism, this seems not to capture our true preferences in Ellsberg cases.
- On the stronger maximin version, can he say that you do not form "a categorical preference for *A* over *B*"?

On Jeffrey's view, Ellsberg behaviour is hard to make sense of.

- If you adopt the uniform distribution, you should be indifferent between *C* and *D*; but people aren't.

But maybe that's OK:

- Jeffrey's (1983) motivation for adopting sets of probabilities is the problem of old evidence!

So perhaps Jeffrey's view is the more stable one.

References

References I

[1] Bernd Anger and Jörn Lembcke.
Infinitely subadditive capacities as upper envelopes of measures.
Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 68(3):403–414, 1985.

[2] Inés Couso, Serafin Moral, and Peter Walley.
Examples of independence for imprecise probabilities.
In *First International Symposium on Imprecise Probabilities and Their Applications*, 1999.

[3] Daniel Ellsberg.
Risk, ambiguity, and the savage axioms.
Quarterly Journal of Economics, 75(4):643–649, 1961.

[4] Itzhak Gilboa and David Schmeidler.
Updating ambiguous beliefs.
Journal of Economic Theory, 59:33–49, 1993.

[5] Alan Hájek.
Objecting vaguely to pascal's wager.
Philosophical Studies, 98(1):1–14, 2000.

[6] Joseph Y. Halpern.
Reasoning About Uncertainty.
MIT Press, 2003.

[7] Joseph Y. Halpern and Riccardo Pucella.
A logic for reasoning about upper probabilities.
Journal of Artificial Intelligence Research, 17:57–81, 2002.

[8] Richard C. Jeffrey.
Bayesianism with a human face.
In J. Earman, editor, *Testing Scientific Theories*, pages 133–156. University of Minnesota Press, 1983.

References II

[9] Richard C. Jeffrey.
Indefinite probability judgment: A reply to levi.
Philosophy of Science, 54(4):586–591, 1987.

[10] James Joyce.
The Foundations of Causal Decision Theory.
Cambridge University Press, 1999.

[11] Frank H. Knight.
Risk, Uncertainty, and Profit.
Houghton Mifflin, 1921.

[12] Isaac Levi.
On indeterminate probabilities.
Journal of Philosophy, 71(13):391–418, 1974.

[13] Isaac Levi.
The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance.
The MIT Press, 1980.

[14] Isaac Levi.
Imprecision and indeterminacy in probability judgments.
Philosophy of Science, 52(3):390–409, 1985.

[15] Isaac Levi.
The Covenant of Reason.
Cambridge University Press, 1997.

[16] G. G. Lorentz.
Multiply subadditive functions.
Canadian Journal of Mathematics, 4(4):455–462, 1952.

References III

[17] Patrick Maher.
Betting on Theories.
Cambridge University Press, 1993.

[18] Frank Plumpton Ramsey.
Truth and probability.
In D. H. Mellor, editor, *Philosophical Papers*. Cambridge: Cambridge University Press, [1926] 1990.

[19] Leonard J. Savage.
The Foundations of Statistics.
Wiley Publications in Statistics, 1954.

[20] Teddy Seidenfeld.
When normal and extensive form decisions differ.
In Dag Prawitz, Brian Skyrms, and Dag Westerstahl, editors, *Logic, Methodology and Philosophy of Science*, pages 451–463. Elsevier, 1994.

[21] Teddy Seidenfeld and Larry Wasserman.
Dilation for sets of probabilities.
The Annals of Statistics, 21(3):1139–1154, 1993.

[22] Scott Sturgeon.
Reason and the grain of belief.
Noûs, 42(1):139–165, 2008.

[23] Bas van Fraassen.
Belief and the will.
The Journal of Philosophy, 81(5):235–256, 1984.

[24] Bas van Fraassen.
Belief and the problem of ulysses and the sirens.
Philosophical Studies, 77:7–37, 1995.

References IV

[25] Bas van Fraassen.
Vague expectation value loss.
Philosophical Studies, 127(3):483–491, 2006.

[26] Peter Walley.
Statistical Reasoning With Imprecise Probabilities.
Chapman & Hall, 1991.

[27] Roger White.
Evidential symmetry and mushy credence.
In *Oxford Studies in Epistemology*. Oxford University Press, 2009.