Ellsberg paradox, and Choquet and Maxmin expected utility

Rohit Lamba
Pennsylvania State University
rlamba@psu.edu

Febraury 2020

1 The Ellsberg Paradox

There are two urns, each contains 100 balls. Urn I contains 50 red balls and 50 black balls. Urn II contains 100 balls each of which is known to be either red or black but we have no information about how many of them are red and how many of them are black. One ball is drawn from the urns and the DM is asked to bet on the color of the ball drawn. A *red bet* is a bet that yields \$100 if the ball drawn is red. A *black bet* is a bet that yields \$100 if the ball drawn in black.

For each of the urns, which should we prefer - red bet or black bet?

Typical pattern of answers is the following:

- For each urns, subject are indifferent between red and black bets.
- Subjects prefer a red bet on urn I over a red bet on urn II and the same for black bet.

If the agent was probabilistically sophisticated, then the above pattern of behavior would imply that

- The DM considers it more likely that a red ball is drawn from urn I than a red ball is drawn form urn II.
- The DM considers it more likely that a red ball is drawn from urn II than a red ball is drawn from urn I.

This is impossible because for each urn the probabilities have to add up to 1.

Let us write this formally. The states are "colors drawn from the two urns", i.e. $\{RR, RB, BR, BB\}$. Here RB would imply a state where red ball is drawn from ur nI and black ball is drawn from urn II, etc. Consider four bets (Savage acts!)- IR, IB, IIR, IIB. Here IB would imply bet on ball drawn from urn I is black, etc. So we get the following table:

	RR	BB	RB	BR
IR	1	0	1	0
IIB	0	1	1	0
IIR	1	0	0	1
IB	0	1	0	1

Now, recollect the sure thing principle (Axiom 4) from Savage. For acts f, g, h, h', $fAh \succsim gAh \implies fAh' \succsim gAh'$. This implies from our table,

$$\text{I}R \succsim \text{II}B \Leftrightarrow \text{II}R \succsim \text{I}R$$

How? Let $A = \{RR, BB\}$. Then, note that IR and IIR agree on A, and IIBB and IB agree on A as well. Read Axiom 4 carefully, and one can make out that with A well defined as above, $fAh =$ IR, $gAh =$ IIB, $fAh' =$ IIR, $gAh' =$ IB, we get I$R \succsim$ II$B \Leftrightarrow$ II$R \succsim$ IR.

However the pattern of behavior we suggested as plausible, and has been found to be compelling in experimental evidence gives,

$$\text{I}R \sim \text{I}B \succ \text{II}R \sim \text{II}B$$

Hence the paradox.

We can state another paradox with a ingle urn. There is a single urn with three balls. One ball is red and the other two are black or yellow. one ball is drawn from the urn and the decision maker bets on the color drawn. Should we

- prefer a bet on a red or bet on black?
- prefer a bet on not-red or not-black?

	R	B	Y
red	1	0	0
black	0	1	0
not red	0	1	1
not black	1	0	1

The sure thing principle says

$$red \succsim black \;\Leftrightarrow\; not-black \succsim not-red,$$

but we observe

$$red \succ black \;\text{ and }\; not-red \succ not-black.$$

2 Schmeidler's Choquet expected utility

2.1 Capacities and Choquet Integral

If some measure of likelihood is to represent the betting behavior in the Ellsberg setting then it cannot be additive. For example consider the single urn Ellsberg experiment and let μ be our measure of likelihood. Suppose $\mu(R) = \frac{1}{3}$, $\mu(B \cup Y) = \frac{2}{3}$ and $\mu(B) = \frac{1}{4} = \mu(Y)$. The function μ

captures the intuitive (observed) betting behavior in the Ellsberg single urn example. Notice that $\mu(B \cup Y) \neq \mu(B) + \mu(Y)$ and hence is not additive.

We now propose one such "non-additive probability measure", called capacity. Let Ω be a finite set of states and $\mathcal{A}$ be the algebra of all subsets of Ω.

Definition 1. *The function $\mu : \mathcal{A} \to [0, 1]$ is a capacity if*

(i) $\mu(\emptyset) = 0$,

(ii) $\mu(\Omega) = 1$, *and*

(iii) $\mu(A) \geq \mu(B)$ *for all* $A, B \in \mathcal{A}, B \subset A$.

Let f be any real valued function on Ω, that takes on value x_i on E_i, where $\{E_1, \ldots, E_n\}$ is a partition of Ω. Thus, $f = \sum_{i=1}^{m} x_i 1_{E_i}$. Let μ be a capacity on $\mathcal{A}$. How should we define the integral $\int_\Omega f \, d\mu$? The obvious Riemannian way would be

$$J(f, \mu) = \sum_i x_i \mu(E_i),$$

However, this is not well defined. To see this, suppose f is a constant ($f = 1$) on Ω. Then, $J(f, \mu) = 1$. $\mu(S) = 1$ satisfies the condition. But, so does $J(f, \mu) = 1.\mu(A) + 1.\mu(A^c)$. Since we could have $\mu(A) + \mu(A^c) \neq 1$, because μ is a capacity, $J(f, \mu)$ is not well defined. So the stage is set or us to define the Choquet Integral.

Definition 2. *Let μ be a capacity on Ω (or $\mathcal{A}$ to be precise) and $f : \Omega \to \mathbb{R}_+$ be a nonnegative real valued function on Ω. Let v_i be the value of f on E_i, $i = 1, 2, \ldots, n$, in decreasing order, where $\{E_1, \ldots, E_n\}$ is of course a partition of Ω. So, $v_1 \geq v_2 \geq \ldots \geq v_n$ and set $v_{n+1} = 0$. Then the Choquet Integral is defined as*

$$\begin{aligned}
\int_\Omega f \, d\mu &= \sum_{i=1}^{n} (v_i - v_{i+1}) \mu(\cup_{j=1}^{i} E_j) \\
&= \sum_{i=1}^{n} v_i (\mu(\cup_{j=1}^{i} E_j) - \mu(\cup_{j=1}^{i-1} E_j)),
\end{aligned}$$

where $\cup_{j=1}^{0} E_j = \emptyset$.

If μ is additive, then the definition is equivalent to the usual Riemann integral. Also, if f is any bounded, nonnegative function, then

$$\int_\Omega f \, d\mu = \int_0^\infty \mu(f \geq t) dt,$$

where the right hand side is the usual Riemann Integral. Notice that this is well defined because $\mu(f \geq t)$ is a nondecreasing function of t. For functions f that may be negative, let $c > 0$ be such that $f + c \geq 0$. Then,

$$\int_\Omega f\,d\mu = \int (f + c)d\mu - c.$$

Note that $\int_\Omega (af + b)d\mu = a\int_\Omega f\,d\mu + b$, however $\int_\Omega (f + g)d\mu \neq \int_\Omega f\,d\mu + \int_\Omega g\,d\mu$, in general. If we had $\int_\Omega (f + g)d\mu = \int_\Omega f\,d\mu + \int_\Omega g\,d\mu$ always, then using indicator functions, we can conclude that μ is additive, which is a contradiction.

Definition 3. *Two real valued functions f, g on Ω are comonotonic if there exists no $s, t \in \Omega$ with $f(s) > f(t)$ and $g(t) > g(s)$. Another way of saying this is that f, g are comonotonic if for all $s, t \in \Omega$, $(f(s) - f(t))(g(s) - g(t)) \geq 0$.*

Lemma 1. *If f and g are comonotonic, then $\int_\Omega (f + g)d\mu = \int_\Omega f\,d\mu + \int_\Omega g\,d\mu$.*

Proof. □

2.2 Preliminaries

Let X be a finite set of prizes. Ω is a finite set of states. An act is a function $h : \Omega \to \mathcal{L}(X)$. Let H be the set of all acts. For $f, g \in H,\ a \in [0, 1]$ define $af(1 - a)g$ by

$$(af + (1 - a)g)(s) = af(s) + (1 - a)g(s)\ \forall s \in \Omega.$$

where $af(s) + (1 - a)g(s)$ is the mixture space operation of vNM theory.

A word on the notation here. $f(s)$ will be used to denote both an element of $\mathcal{L}(X)$ that maps f maps into, in state s and also a constant act in H that gives the same lottery $f(s)$ in every state. The usage will be clear from the context.

Definition 4. *The acts f and g are comonotonic if there is no pair $s, s' \in \Omega$ with $f(s) \succ f(s')$ and $g(s') \succ g(s)$.*

Next, we state the axioms involved.

Axiom 1: $\succsim$ on H is a preference relation.

Axiom CI: For pairwise comonotonic acts $f, g, h \in H, f \succ g$ and $a \in (0, 1) \implies af + (1 - a)h \succ ag + (1 - a)h$.

Axiom 3: For all $f, g, h \in H,\ f \succ g \succ h \implies \exists a, b \in (0, 1)$ such that $af + (1 - a)h \succ g \succ bf + (1 - b)g$.

Axiom 4: There exists f and g in H such that $f \succ g$.

Axiom 5: For all $f, g \in H,\ f(s) \succsim g(s)\ \forall s \in \Omega \implies f \succsim g$,

CI stands for comonotonic independence. Axiom 3 is the usual continuity axiom, Axiom 4 is a non-degeneracy axiom and Axiom 5 is often called monotonicity.

2.3 Results

Theorem 1. *$\succsim$ satisfies Axioms 1, CI, 3-5 iff there exists a non-constant linear function $U : \mathcal{L}(X) \to \mathbb{R}$ and a capacity $\mu : \mathcal{A} \to [0,1]$ such that the function defined by $W(f) = \int (U \circ f) d\mu$ represents $\succsim$.*

Proof. $\square$

Definition 5. *Let $\succsim$ be binary relation on H. Then, $\succsim$ is said to be uncertainty averse if $f \succsim h, g \succsim h$ and $a \in [0,1] \implies af + (1-a)g \succsim h$. An equivalent definition is $f \succsim g$ and $a \in [0,1] \implies af + (1-a)g \succsim g$.*

Intuitively, uncertainty aversion means that "smoothing" or averaging utility distributions makes the decision maker better off. Another way is to say that substituting objective mixing to subjective mixing makes the decision maker better off. We now present a full mathematical characterization of the concept

Definition 6. *A capacity $\mu : \mathcal{A} \to \mathbb{R}$ is said to be convex if for all $A, B \in \mathcal{A}$,*

$$\mu(A) + \mu(B) \leq \mu(A \cup B) + \mu(A \cap B).$$

Let $P(\Omega, \mathcal{A})$ be the set of all probability measures on $(\Omega, \mathcal{A})$.

Definition 7. *The core of capacity μ, denoted by $C(\mu)$, is the set of probability measures that assign to each event at least a probability equal to the capacity,*

$$C(\mu) = \{p \in P(\Omega, \mathcal{A}) |\, p(A) \geq \mu(A)\ \forall A \in \mathcal{A}\}.$$

So, for example if $\Omega = \{1,2\}$, and $\mu(1) = \mu(2) = 0.4$, then, $C(\mu) = \{(p_1, p_2) |\ p_1 = 1 - p_2$ and $p_1 \in [0.4, 0.6]\}$.

Theorem 2. *Let $\succsim$ be a binary relation on H that satisfies Axioms 1, CI, 3-5, and $\mu : \mathcal{A} \to [0,1]$ be a capacity. Then, the following are equivalent:*

(i) *$\succsim$ is uncertainty averse.*

(ii) *μ is convex.*

(iii) *$\int_\Omega v d\mu = min_{p \in C(\mu)} \int_\Omega v d\mu$ for all $v : \Omega \to \mathbb{R}$.*

Proof. $(i) \implies (ii)$ Normalize the utility of the worst prize to be 0 and of the best prize to be 1. $\succsim$ induces a preference relation on $\mathcal{L}(X)$ which in turn induces a preference relation on X. Since X is finite, a normalization between 0 and 1 is possible. For every $v : \Omega \to \mathbb{R}$, define,

$$I(v) = \int_\Omega v d\mu.$$[1]

$\square$

[1]Note that for $v : \Omega \to [0,1]$, there is an act h_v such that $U(h_v) = v$, where U is the vNM utility function derived in theorem 1. Thus, I is basically the same as W in the theorem, taking into account duality between $v \leftrightarrow h_v$.

3 Gilboa-Schmeidler's Maxmin expected utility