Savage representation theorem

A weakness of the vNM theorem is that it assumes the probabilities in the lotteries $L \in \mathcal{L}(E)$ are known. One could imagine disputing that vNM is useful for modelling agency on the basis that agents often need to make decisions whose outcomes are uncertain in a non-quantified way.

The thrust of the Savage theorem is that "reasonable" agents in fact assign subjective probabilities to possible states of the world and make choices to maximize subjective expected utility.

To formalize this, let S be a set of states, and Θ a set of outcomes. Then the set A of acts is

$$A = \{\text{fns } S \longrightarrow \Theta\}.$$

We will show that reasonable agents assign probabilities $P(s)$ to states $s \in S$ and utilities $U(x)$ to outcomes $x \in \Theta$, then choose s to maximize $\sum_{s \in S} U(f(s))P(s)$.

Defn

(1) A probability measure on S is a map $P: S \longrightarrow [0,1]$ s.t. $\sum_{s \in S} P(s) = 1$.

(2) Given a probability measure P on S and a utility fn $U: \Theta \longrightarrow \mathbb{R}$, the subjective expected utility of an act $f: S \longrightarrow \Theta$ is $SEU_{P,U}(f) := \sum_{s \in S} U(f(s))P(s)$.

(3) If P is a prob. measure on S and $U: \Theta \longrightarrow \mathbb{R}$ is a utility fn, then we get an induced weak order $\leq_{P,U}$ on A via:

$$f \leq_{P,U} g \iff SEU_{P,U}(f) \leq SEU_{P,U}(g).$$

Thus, as with the vNM theorem, we have a map

$$\left\{\begin{matrix}\text{pairs } (P,U) \text{ of} \\ P: \text{prob. meas. on } S \\ U: \Theta \to \mathbb{R}\end{matrix}\right\} \longrightarrow \left\{\begin{matrix}\text{weak orders} \\ \text{on } A\end{matrix}\right\}.$$

As before, we can ask:

(1) When do two pairs (P,U) induce the same order?

(2) Which weak orders are induced by some (P,U)?

(3) Which weak orders are reasonable for an agent?

Thm (Savage representation theorem)

Let $\preceq$ be a weak order on A. Then $\preceq$ satisfies conditions (P1)–(P5) below iff $\preceq$ is induced by a (P,U). In this case P is unique, and U is unique up to pos. affine transformation.

Rmk This implies that a rational agent's beliefs about the world can be inferred from their preferences!

Example

Suppose we flip weighted coins C_1 and C_2. You choose one coin such that you win \$10 if the coin is heads. If you choose C_1 over C_2, that implies you think C_1 is more likely to land heads than C_2.

Rmk Note that our set $A = \{\text{fns } S \longrightarrow \Theta\}$ of acts includes some "impossible" acts, like the act which always results in the best outcome regardless of the state. Even if the agent doesn't have the option to choose any possible $f \in A$, the agent still has preferences over them. Given a set of possible acts $A' \subseteq A$, the agent selects the most preferred $f \in A'$.

Savage axioms

First we fix some notation.

Defn

(1) An event is a subset $E \subseteq S$.

(2) If $\alpha \in \Theta$ is an outcome, then the constant act $\underline{\alpha} \in A$ is defined by $\underline{\alpha}(s) = \alpha$ for all $s \in S$.

(3) If $E \subseteq S$ is an event and $f, g \in A$ are acts, then define the act $f_E g$ by

$$f_E g(s) = \begin{cases} f(s) & \text{if } s \in E \\ g(s) & \text{if } s \notin E. \end{cases}$$

(4) An event E is nonnull if there are outcomes $\gamma, \beta \in \Theta$ (γ = "good", β = "bad") s.t.

$$\gamma_E \beta \succ \beta. \quad E \text{ is null otherwise.}$$

Rmk The null events are the events that can be ignored when comparing preferences. They will be assigned probability 0.

(P1) (Sure thing principle) for all $\alpha, \beta \in \Theta$, $f, g \in A$

$$\alpha_E f \preceq \alpha_E g \iff \beta_E f \preceq \beta_E g.$$

(P2) (Monotonicity) If E is nonnull and $\gamma, \beta \in \Theta$, $f \in A$ then $\beta \preceq \gamma \iff \beta_E f \preceq \gamma_E f$

Rmk These axioms are Savage's version of independence.

(P3) (state neutrality) Let E, B be events, and $\beta, \gamma, \beta', \gamma' \in \Theta$ with $\beta \prec \gamma$, $\beta' \prec \gamma'$. Then

$$\gamma_E \beta \preceq \gamma_B \beta \iff \gamma'_E \beta' \preceq \gamma'_B \beta'.$$

Rmk This axiom implies that the agent has beliefs over the relative likelihoods of E and B. More precisely, given a weak order $\preceq$ on A satisfying (P3), we can define an order $\leq_P$ on the set of events by

$$E \leq_P B \iff \gamma_E \beta \preceq \gamma_B \beta \text{ for some } \beta \prec \gamma \in \Theta.$$

By (P3), this is independent of the particular choice of β, γ.

(P4) (continuity) For all $f, g \in A$, $\beta, \gamma \in \Theta$ and events $E \subseteq S$ with $\beta \prec \gamma$ and $\beta_E f \prec g \prec \gamma_E f$, there is some partition $E = E_1 \cup E_2$ s.t.

$$g \sim \beta_{E_1} \gamma_{E_2} f.$$

(P5) (nontriviality) There are $\beta, \gamma \in \Theta$ s.t. $\beta \prec \gamma$.