SAVAGE'S SUBJECTIVISM

1. The Theory

1.1. Decision matrix. The primary goal of Savage's theory is to provide a Bayesian subjective interpretation of probability that is widely employed in virtually all stages of statistical inferences. The method adopted by subjectivists was to ground the concept of probability in a normative theory of decision making.

Example (Omelet)**.** The husband volunteers to finish making the omelet started by his wife. Already, there are five eggs in the mixing bowl, the task is to break the sixth egg into to the bowl and start the cooking process. However, for some reason, the sixth egg is likely to be either good or rotten. The husband needs to make a decision as to how to proceed. ◁

Table 1. Decision Matrix of the Omelet Example.

Act	State	
	Good	Rotten
break into bowl	six-egg omelet	no omelet and all five eggs destroyed
break into saucer	six-egg omelet and a saucer to wash	five-egg omelet and a saucer to wash
throw away	five-egg omelet and one good egg wasted	five-egg omelet

Notations and definitions.

- X**:** a set $\{o_{1,1}, o_{1,2}, \ldots\}$ of consequences. "Consequences are conceived of as what the person experiences and what he has preferences for even where there is no uncertainty."
- S**:** an (infinite) set $\{s_1, s_2, \ldots\}$ of states of the world. "A state of the world is, again informally, a possible list of answers to all questions that might be pertinent to the decision situation at hand. The states need not be conceived of as absolutely atomic but only atomic for the context; and even this use of atoms is didactic rather than necessary."
- $\mathcal{A}$**:** the set $\{f_1, f_2, \ldots\}$ of all practically and theoretically possible actions. "An act is defined as a function, or schedule, associating a consequence for the person with each possible state of the world."

Date: October 15, 2015.

Seminar on Decision Theory, Michaelmas Term 2015.

TABLE 2. Savage's decision matrix.

	s_1	s_2	$\cdots$	s_n	$\cdots$
f_1	$o_{1,1}$	$o_{1,2}$	$\cdots$	$o_{1,n}$	$\cdots$
f_2	$o_{2,1}$	$o_{2,2}$	$\cdots$	$o_{2,n}$	$\cdots$
$\vdots$			$\ddots$		
f_m	$o_{m,1}$	$o_{m,2}$	$\cdots$	$o_{m,n}$	$\cdots$

Definition (Combined acts)**.** For any $f, g \in \mathcal{A}$, define the *combination* of f and g with respect to an event E (a set of states), written $f \oplus_E g$, to be such that:

$$(f \oplus_E g)(s) = \begin{cases} f(s) & \text{if } s \in E \\ g(s) & \text{if } s \in E^C, \end{cases} \tag{1}$$

where $E^C = S - E$ is the compliment of E.[1]

1.2. **Representation theorem.** As a primitive assumption, the agent is assumed to have *preferences* over acts, which are modeled by a preorder $\succsim$ on $\mathcal{A}$. Thus, for any acts $f, g \in \mathcal{A}$, $f \succsim g$ is taken to mean that act f is *weakly preferred to* act g (or that g is not preferred to f) by the agent.

$$f \succ g \ =_{\text{Df}} f \succsim g \text{ and } g \not\succsim f$$
$$f \sim g \ =_{\text{Df}} f \succsim g \text{ and } g \succsim f$$

The objective is *to represent the agent's preferences among acts by their expected utilities*: if the agent's preference relation $\succsim$ satisfies a set of postulated rationality and structural axioms the there exist a (subjective) probability measure μ on events and a real-valued utility function u on X such that, for any $f, g \in \mathcal{A}$,

$$f \succsim g \iff \int_S u\big[f(s)\big]\, d\mu \geq \int_S u\big[g(s)\big]\, d\mu. \tag{2}$$

1.3. **The sure-thing principle and postulate 2.** The cornerstone of Savage's decision model is a postulated rationality principle known as the "sure-thing principle". The following is the example used by Savage to motivate this principle.

Example (Businessman)**.** A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness the purchase. So, to clarify the matter for himself, he asks whether he would buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that the would buy in either event, he decides that the should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say. ◁

[1]Some writers use '(f, E, g)' or 'fEg' or '$f|E + g|E^C$' or '$[f$ on E, g on $E^C]$' for combined acts.

STP: If the decision maker prefers one act over another assuming either certain event obtains or that the compliment of the event obtains, then her preference over the two acts shall remain unchanged.

- The principle is sometimes referred to as the *dominance principle*.
- STP can be easily generalized for n-partitions.
- Savage takes STP to be fundamental to rational decision making: "I know of no other extra-logical principle governing decisions that finds such ready acceptance."

Conditional preferences. Note that the statement of the sure-thing principle above employs explicitly a concept of conditional preference, that is, one act being preferred to another *given* the occurrence of certain event. Since the current formal setup is based entirely on *unconditional* preferences over acts, the notion of conditional preference is not directly expressible. Some alternative arrangements hence need to be made.

Definition (Conditional preference)**.** Let E be some event, then, given acts $f, g \in \mathcal{A}$, f is said to be weakly preferred to g *given* E, written $f \succsim_E g$, if for *all* pairs of acts $f', g' \in \mathcal{A}$ the following condition is satisfied

$$\left.\begin{array}{ll} f(s) = f'(s),\, g(s) = g'(s) & \text{if } s \in E \\ f'(s) = g'(s) & \text{if } s \in E^C \end{array}\right\} \implies f' \succsim g'. \tag{3}$$

Now, given the definition of conditional preference, **STP** can be translated into

$$\left[f \succsim_E g, f \succsim_{E^C} g \right] \implies f \succsim g. \tag{STP}$$

Postulate 2. However, Savage was unwilling to incorporate (STP) directly into his system for the concern that the concept of conditional preference is based on knowledge of the possible occurrences of some events, the introduction of which may lead to, it is said, unsought philosophical complications. Savage (1972, p.22)

- "The sure-thing principle [i.e., **STP** above] cannot appropriately be accepted as a postulate in the sense that P1 is, because it would introduce new undefined technical terms referring to knowledge and possibility that would refer it mathematically useless without still more postulates governing these terms."
- "It will be preferable to regard the principle as a loose one that suggests certain formal postulates [i.e. **P2** below] well articulated with P1."

Instead, Savage formulated his second postulate as follows

$$f \oplus_E h \succsim g \oplus_E h \iff f \oplus_E h' \succsim g \oplus_E h', \tag{P2}$$

P2: If the consequences of two acts differ on the occurrence of some event E but otherwise agree with each other, then their preferential comparison between these two acts shall be decided on those states in E and their corresponding consequences.

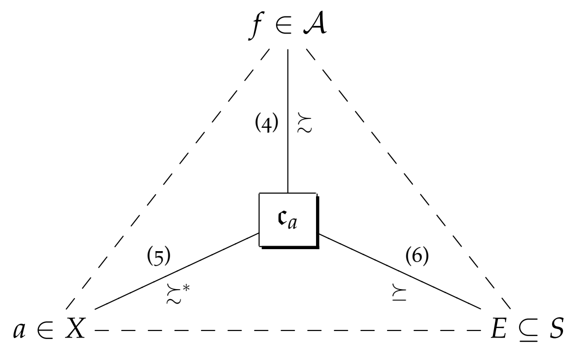

FIGURE 1. Constant act $\mathfrak{c}_a$ and other parameters in Savage's model.

1.4. **Deriving numerical probabilities and utilities.** Savage postulate a set of axioms on the preference relation $\succsim$. From the first five postulates a comparative notion of subjective probability is derived which reflects the agent's qualitative probabilistic judgments over possible circumstances under which these actions are taking place. With the sixth postulate, the derived qualitative probability is further represented by a numerical probability measure together with a personal utility function for simple acts (i.e., acts that may lead to finitely many potential consequences under different states). The last postulate is brought in so that the utility function for simple acts can be extended to all acts (cf. Table 3).

TABLE 3. Inferential order in Savage's system.

P1-5		**+ P6**		**+ P7**
Qualitative probability	$\Rightarrow$	Quantitative probability Utility for simple acts	$\Rightarrow$	Utility for all acts

Definition (Constant acts)**.** For any $a \in X$, an act is said to be *constant* with respect to consequence a, in symbols $\mathfrak{c}_a$, if

$$\mathfrak{c}_a(s) = a \quad \text{for all } s \in S. \tag{4}$$

- This concept is a key structural component of Savage's theory.
- Given a preference ordering $\succsim$ on $\mathcal{A}$, an ordering $\succsim^*$ over consequences can be defined using constant acts by

$$a \succsim^* b \quad \iff \quad \mathfrak{c}_a \succsim \mathfrak{c}_b \quad \text{for all } a, b \in X. \tag{5}$$

Qualitative probability. For any events $E, F \in \mathcal{F}$, say that E is *weakly more probable* than F, written $E \succeq F$ (or $F \preceq E$), if, for any $a, b \in X$ with $a \succsim b$,

$$\mathfrak{c}_a \oplus_E \mathfrak{c}_b \succsim \mathfrak{c}_a \oplus_F \mathfrak{c}_b \tag{6}$$

(or equivalently if $\mathfrak{c}_b \oplus_F \mathfrak{c}_a \succsim \mathfrak{c}_b \oplus_E \mathfrak{c}_a$). E and F are said to be *equally probable*, in symbols $E \simeq F$, if both $E \succeq F$ and $F \succeq E$ hold.

Separability. The following two axioms are usually referred to as the *independence axioms* of Savage's system.

P3: For any consequences $a, b \in X$ and for any non-null event E, $\mathfrak{c}_a \succsim_E \mathfrak{c}_b$ if and only if $a \succsim b$.

P4: For any $a, b, c, d \in X$ satisfying $a \succsim b$ and $c \succsim d$ and for any events $E, F \subseteq S$, $\mathfrak{c}_a \oplus_E \mathfrak{c}_b \succsim \mathfrak{c}_a \oplus_F \mathfrak{c}_b$ if and only if $\mathfrak{c}_c \oplus_E \mathfrak{c}_d \succsim \mathfrak{c}_c \oplus_F \mathfrak{c}_d$.

- To guarantee various formal concept defined above are well defined.
- The agent's value judgments over the consequences and her probabilistic estimations are mutually independent.

2. Some Evaluations

Goals of decision theory.

Ahmed. Decision theory works roughly like this. Suppose an agent has several options in some situation. The situation has many possible outcomes that the agent desires or dreads in varying degrees. For each option, if she takes it, some outcomes are more likely than others. At any rate she takes some outcomes to be more likely than others if she takes the option. Suppose we have some idea of how good each outcome is for her. And suppose we have some idea of how likely she considers each outcome, given each option. Decision theory takes all this as input – the options, the possible outcomes, and how good and how likely she considers them. Its output is a *recommendation* of some option or options. p.6

Savage. The preference theory could be studied as an axiom system without interpretation; but to mathematicians it is of only modest interest. It can be interpreted as a theory about how people, corporations, or other organisms, actually behave; but as psychology, it has very limited validity and use. The interpretation for which it was developed is a normative one by which a person can *police his own potential decisions for incoherency*. ...

To use the preference theory is to search for incoherence among potential decisions, of which you, the user of the theory, must then revise one or more. The theory itself does not say which way back to coherence is to be chosen, and presumably should not be expected to.

Behaviorism.

Ahmed. So it seems that we must build in the following specification from the start: an appropriate partition of S into events must be one that the agent believes to be (somehow: either causally or evidentially) independent of which option he chooses. But then we must already have a conception of what it is for the agent to have such beliefs. ...But now we see that that is wrong. What is right is this: that for the agent to think sunshine more probable than rain is for him to prefer the first bet to the second if he already believes that whether there is sunshine or rain tomorrow is (in some sense) independent of his choice between these bets.

So it turns out that the agents beliefs are after all not just facts about his actual or possible behaviour. They are facts about the actual and possible behaviour of an agent equipped with such-and-such other beliefs, in particular with other beliefs about the independence of events from acts. So we must give up on the aim of such purely operational explications of beliefs and desires as Savages personal probability and utilities were intended by him to be. p.33

References

Ahmed, A. (2014). *Evidence, Decision and Causality*. Cambridge University Press.

Savage, L. J. (1967). Difficulties in the theory of personal probability. *Philosophy of Science 34*(4), 305–310.

Savage, L. J. (1972). *The Foundations of Statistics* (Second Revised ed.). Dover Publications, Inc.