Notes sur l'axiomatisation de Savage (1954)

Mikael Cozic (Paris IV-Sorbonne - IHPST)

14 mai 2004

1 Le théorème de représentation

Définition 1
Soient S un ensemble d'états de la nature, C un ensemble de conséquences, $F = \{f : S \to C\}$ l'ensemble des actes et $>$ une relation sur F dite ***préférence sur les actes****.*

On note :
$f \sim g$ *si non* $f > g$ *et non* $g > f$,
$f \geq g$ *si* $f > g$ *ou* $f \sim g$,
$f =_E x$ *si* $f(s) = x$ *pour tout* $s \in E$,
$f =_E g$ *si* $f(s) = g(s)$ *pour tout* $s \in E$, *et*
$f = xEy$ *si* $f =_E x$ *et* $f =_{E^c} y$.

A partir de la relation de préférence sur les actes, on définit deux relations de préférence dérivées :

Définition 2

(D1) la relation de ***préférence sur les actes conditionnelle à*** E*, notée* $>_E$*, est définie ainsi :*

$f >_E g$ si $\forall f', g' \in F$, si $f' =_E f$ et $g' =_E g$, alors $f' > g'$

(D2) la relation de ***préférence sur les conséquences*** *est définie ainsi :*

$x > y$ *ssi* $f > f'$ où $\forall s \in S$, $f(s) = x$ *et* $f'(s) = y$

(D3) Pour tout $E \in \wp(S)$, E *est* **nul** *si* $f \sim g$ *quand* $f =_{E^c} y$ [1]

Postulats 1

(P1) $>$ *est un ordre asymétrique et négativement transitif*
(P2) si $f =_E f'$, $g =_E g'$, $f =_{E^c} g$ *et* $f' =_{E^c} g'$, alors $f > g$ *ssi* $f' > g'$
(P3) si E *n'est pas nul,* $f =_E x$ *et* $g =_E y$, alors $f >_E g$ *ssi* $x > y$

[1] [Sav72], [Fis94]. J'ai suivi la formulation de ce dernier, qui contient une légère amélioration par rapport au résultat original de Savage.

(P4) si $x > y$ et $x' > y'$, alors $xEy > xE'y$ ssi $x'Ey' > x'E'y'$

(P5) il existe certains $z, w \in C$ tels que $x > y$

(P6) si $f > g$, alors, étant donné x, il existe une partition finie de S telle que, pour tout membre S_i de la partition, si $f' =_{S_i} x$ et $f' =_{S_i^c} f$ alors $f' > g$, et si $g' =_{S_i} x$ et $g' =_{S_i^C} g$, alors $f > g'$.

(P7) si, pour tous $s \in E$, $f >_E g(s)$, alors $f \geq_E g$; et si, pour tous $s \in E$, $f(s) >_E g$, alors $f \geq_E g$

La notion de probabilité que l'on peut extraire à partir de ces postulats est plus faible que la notion classique de Kolmogorov [2] du point de vue de l'additivité ; voici celle qui est adaptée au théorème de représentation.

Définition 3
*Soit S un ensemble d'états ; un sous-ensemble $\mathcal{E}$ de $\wp(S)$ est une **algèbre de Boole** si*

(i) $S \in \mathcal{E}$

(ii) si $E \in \mathcal{E}$, alors $E^c \in \mathcal{E}$ (stabilité par passage au complémentaire)

(iii) si, pour tous $E, E' \in \mathcal{E}$, $E \cup E' \in \mathcal{E}$ (stabilité par réunion finie)

*On dit alors que $(S, \mathcal{E})$ est un **espace booléen**.*

Définition 4
*Soit un espace booléen $(S, \mathcal{E})$; la relation $>$ sur $\mathcal{E}$ est une **probabilité qualitative**[3] si pour tout $E, E', E'' \in \mathcal{E}$,*

(i) $>$ est un ordre faible,

(ii) $S > \emptyset$,

(iii) non $\emptyset > E$, et

(iv) si $(E \cup E') \cap E'' = \emptyset$, alors $E > E'$ ssi $E \cup E'' > E' \cup E''$.

Définition 5
*Soit un espace booléen $(S, \mathcal{E})$; P est une **distribution de probabilité finiment additive** sur $(S, \mathcal{E})$ si*

(i) P est une fonction de $\mathcal{E}$ dans $[0, 1]$,

(ii) si les ensembles E, E' sont disjoints, $P(E \cup E') = P(E) + P(E')$ (additivité finie), et

(iii) $P(S) = 1$.

Définition 6
*Soit un espace booléen $(S, \mathcal{E})$, une probabilité qualitative $>$ sur $\mathcal{E}$, une distribution de probabilité P sur $(S, \mathcal{E})$. P **représente** $>$ si*

$\forall E, E' \in \mathcal{E}$, $E > E'$ ssi $P(E) > P(E')$.

Le résultat principal de Savage peut alors s'énoncer de la manière suivante :

Théorème 1

Théorème de représentation

Si la relation $>$ sur F satisfait (P1)-(P7), alors

[2] Voir sur ce point [Sav72], pp. 40-43, [Fis94], p.1415, [Gil00], pp. 65 et sqq.
[3] Voir [Sav72], p. 32, [Fis94], p. 1414.

(i) il existe une distribution de probabilité finiment additive P sur $(S, \wp(S))$ telle que (a) $P(E) > P(E')$ si $xEy > xE'y$ quand $x > y$ et (b) $P(E) = 0$ quand E est un événement nul.

(ii) il existe une fonction bornée $u : C \to \mathbb{R}$, unique à une transformation affine croissante près, telle que

$\forall f, g \in F$, $f > g$ *ssi* $\int_S u(f(s))dP(s) > \int_S u(g(s))dP(s)$

2 Discussion du théorème de représentation

Ce que montre l'axiomatisation de Savage, c'est donc que si les préférences sur les actes d'un agent satisfont les postulats (P1)-(P7), alors on peut rationaliser ses préférences par le principe de maximisation de l'utilité espérée : il existe une distribution de probabilité et une fonction d'utilité sur les conséquences telles que l'agent préfère un acte à un autre si, et seulement si, l'espérance d'utilité du premier acte est supérieure à celle du second. Le principal antécédent historique du résultat de Savage est le théorème de représentation de Von Neumann et Morgenstern (1947) ; la différence essentielle tient à ce que, dans leur axiomatisation, Von Neumann et Morgenstern considèrent la distribution de probabilité comme donnée, au même titre que les préférences sur les actes. Les axiomes sont des conditions suffisantes pour qu'il existe une fonction d'utilité sur les conséquences telles que les préférences sur les actes soient représentées par l'espérance d'utilité définie à partir de cette fonction et de la distribution de probabilité. Dans le cadre de Von Neumann et Morgenstern, on infère l'utilité à partir des préférences sur les actes et de la probabilité, tandis que dans celui de Savage on infère l'utilité et la probabilité à partir des seules préférences sur les actes.

A quoi sert une axiomatisation comme celle de Savage ? Quelles sont les motivations d'un tel théorème de représentation ?[4]

Il est important de remarquer que les préférences sur les actes sont étroitement attachées aux comportements de l'agent : l'agent préfère f à g si, quand il a le choix entre f et g, il choisit f. C'est un point fondamental pour comprendre les différentes motivations du théorème de représentation.

2.1 Motivations fondationnelles

La principale motivation qu'affiche Savage est fondationnelle : il construit son axiomatisation comme un argument en faveur de son interprétation des probabilités.

Pour établir le théorème de représentation, on a défini une probabilité qualitative à partir de la relation de préférence sur les actes, et ensuite on l'a représentée numériquement par une probabilité quantitative. On a là ce qu'on pourrait appeler une caractérisation **pragmatique**[5] des croyances de l'agent

[4]Pour une discussion générale et approfondie, voir [Pic96].

[5][Fis94] dit "*preference-based*". [Pic96] : "[Savage] suggère, sur un mode que l'on pourrait dire pragmatiste, que le comportement d'une personne exprime mais surtout *donne sens* aux

puisqu'elles sont caractérisées par leur rôle dans le choix de l'agent. On retrouve une intuition fondamentale de Ramsey ([Ram26]) :

"...le degré d'une croyance est une propriété causale de la croyance, qu'on peut vaguement définir comme étant la propension à agir sur la base de cette croyance." [6]

En matière d'interprétation des probabilités, Savage considère (p.60) que cette caractérisation pragmatique des croyances le distingue, avec Ramsey, d'autres subjectivistes comme Koopman et Good qui ne font pas jouer un rôle fondamental aux décisions dans la façon dont ils caractérisent la "confiance" d'un individu dans la vérité de proposition. Pour Savage et Ramsey, cette confiance se mesure à la façon dont l'agent agit. Voici comment [Fis86] formule cette cartographie des subjectivistes :
"*The theory of subjective probability attempts to make precise the connection between coherent disposition toward uncertainty and quantitative probability as axiomatized by Kolmogorov (1933) and others. It accomodates the classical interpretations of probability in Bayes (1763) and Laplace (1812), the intuitive views of Koopman (1940a, 1940b) and Good (1950), and the decision-theoretic approach of Ramsey (1931), de Finetti (1931, 1937), and Savage (1954).*"

2.2 Motivations descriptives

Un second type de motivation pour le théorème de représentation porte sur l'utilisation du modèle d'utilité espérée comme théorie descriptive du comportement humain en situation d'incertitude. En effet, on peut alors considérer les postulats essentiellement comme les conséquences observables du principe de maximisation de l'utilité espérée. Ils fournissent donc un moyen de mettre à l'épreuve la validité descriptive de la théorie. L'axiomatisation est dans ce cas un outil d'évaluation descriptive. J'illustre les commentaires que je viens de faire sur l'axiomatique de Savage avec la déviation des postulats que l'article célèbre [Ell61] a mis en évidence.

Une urne contient 90 boules ; 30 sont rouges, et 60 sont noires ou jaunes, dans une proportion inconnue de l'agent. On tire une boule au hasard de l'urne, et on demande à l'agent de choisir entre des paires de paris qui portent sur l'issue du tirage.

<u>Choix 1</u> :
f_1 = parier sur $Rouge$ = 100 euros si la boule tirée est rouge, 0 sinon
f_2 = parier sur $Noir$ = 100 euros si la boule tirée est noire, 0 sinon

<u>Choix 2</u> :
f_3 = parier sur $Rouge \cup Jaune$ = 100 euros si la boule tirée est rouge ou jaune, 0 sinon

jugements d'une personne."(p.206) ; [?] : "*Probabilities and utilities are defined in term's of a person's preferences, in so far as theses preferences satisfy certain consistency assumptions. The definition is constructive ; that is, the probabilities and utilities can be calculated from observed preferences.*" (p.1999) ; [Sav67] : "The probability measure associated with events by prizes is intimely associated with the person's attitude and potential behavior toward uncertain event." (p.307)

[6] [Ram03], p. 164.

$f_4 =$ parier sur $Noir \cup Jaune =$ 100 euros si la boule tirée est noire ou jaune, 0 sinon

Posons $E = Rouge \cup Noir$; alors $f_1 =_E f_3$, $f_2 =_E f_4$, $f_1 =_{E^c} f_2$ et $f_3 =_{E^c} f_4$. Le postulat (P2) s'applique et donc $f_1 > f_3$ ssi $f_2 > f_3$. Or, on observe expérimentalement que la majorité des sujets préfèrent f_1 à f_2 mais f_4 à f_3. Le postulat (P2) est invalidé expérimentalement. Cela implique que les conditions de justification du MEU ne sont pas remplies : il n'est *pas certain* que l'on puisse (en l'occurence) *expliquer* ce comportement à l'aide du MEU. En fait, il n'est pas possible de définir une relation de probabilité qualitative sur la base de ces deux choix : si $f_1 > f_2$, alors *Rouge* $>$ *Noir* ; mais $f_4 > f_3$ implique $Noir \cup Jaune > Rouge \cup Jaune$ donc *Noir* $>$ *Rouge*. Il est donc, dans ce cas, *impossible* d'expliquer le comportement à l'aide du MEU, *pour autant que les probabilités sont définies pragmatiquement.*

2.3 Motivation normative

La motivation normative est analogue à la motivation desciptive : l'axiomatisation nous donne les moyens d'évaluer la validité normative du modèle d'espérance d'utilité ; ce sont d'ailleurs souvent les mêmes expérimentations qui servent de base de discussion aux évaluations normative et descriptive du modèle. l'axiomatisation est un moyen de répondre à la question de savoir si la maximisation de l'utilité est le bon guide à suivre en situation d'incertitude. Concrètement, pour mettre le modèle en difficulté du point de vue normatif, il suffit de construire un ensemble de situations de choix tel que les choix qui paraissent raisonnables aux agents enfreignent un des axiomes. C'était par exemple l'intention originale de M. Allais ([All53])dans ses célèbres choix entre paris. Un des avantages de l'axiomatisation, c'est qu'elle permet de guider la construction de situations de choix qui violent le principe de maximisation de l'utilité espérée. Elle fournit certains critères qui permettent de déterminer qu'un ensemble de choix n'est pas rationalisable.

2.4 Remarques conclusives

Les motivations épistémologiques et normatives sont communes aux théorèmes de représentation de Savage et de Von Neumann et Morgenstern ; en revanche, la motivation fondationnelle est propre aux théorèmes de représentation de Savage (et à ceux qui sont apparentés), parce que l'axiomatisation ne suppose pas donnée une distribution de probabilité mais permet de l'engendrer.

Références

[All53] M. Allais, *Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l'Ecole Américaine*, Econometrica **21** (1953), 503–546.

[Ell61] D. Ellsberg, *Risk, Ambiguity and the Savage Axioms*, Quaterly Journal of Economics **75** (1961), 643–669.

[Fis86] P.C. Fishburn, *The Axioms of Subjective Probability*, Statistical Science **1** (1986), no. 3, 335–358.

[Fis94] ———, *Utility and Subjective Probability*, Handbook of Game Theory (R.J. Aumann and S. Hart, eds.), Elsevier Science, 1994.

[Gil00] D. Gillies, *Philosophical Theories of Probability*, Routledge, Londres, 2000.

[Pic96] E. Picavet, *Choix Rationnel et Vie Publique*, Fondements de la politique, PUF, 1996.

[Ram26] F.P. Ramsey, *Truth and Probability*, The Foundations of Mathematics and other Logical Essays (R.B. Braithwaite, ed.), Kegan Paul, 1926.

[Ram03] ———, *Logique, philosophie et probabilités*, Vrin, Paris, 2003.

[Sav67] L.J. Savage, *Difficulties in the Theory of Personal Probability*, Philosophy of Science **34** (1967), 305–310.

[Sav72] L. Savage, *The foundations of statistics*, 2nd ed., Dover, New York, 1972.