RAIRO. RECHERCHE OPÉRATIONNELLE

JEAN-YVES JAFFRAY

Généralisation du critère de l'utilité espérée aux choix dans l'incertain régulier

RAIRO. Recherche opérationnelle, tome 23, n° 3 (1989), p. 237-267

<http://www.numdam.org/item?id=RO_1989__23_3_237_0>

Recherche opérationnelle/Operations Research
(vol. 23, n° 3, 1989, p. 237 à 267)

GÉNÉRALISATION DU CRITÈRE DE L'UTILITÉ ESPÉRÉE AUX CHOIX DANS L'INCERTAIN RÉGULIER (*)

par Jean-Yves JAFFRAY [1]

Résumé. — *On définit l'incertain régulier comme toute situation où l'information sur les événements est compatible avec un ensemble de lois de probabilités, caractérisable par son enveloppe inférieure, qui est une capacité convexe. On justifie alors axiomatiquement un critère de choix dans l'incertain régulier, qui généralise le critère de l'utilité espérée.*

Mots clés : Décision dans l'incertain; capacité convexe; fonction de croyance; utilité espérée; von Neumann-Morgenstern.

Abstract. — *Regular uncertainty denotes any situation where data concerning the events are consistent with a set of probability measures, characterizable by its lower envelope, which is a convex capacity. A criterion of choice under regular uncertainty is then justified axiomatically and consists in a generalization of the expected utility criterion.*

Keywords : Decision under uncertainty; convex capacity; belief function; expected utility; von Neumann-Morgenstern.

1. INTRODUCTION

La probabilité de défaillance, p, de tel ou tel composant d'un produit de haute technologie en constitue souvent une caractéristique essentielle. Un composant pouvant provenir de lots de qualités diverses, la valeur de p n'est en général qu'imparfaitement connue et l'information disponible permettra parfois tout au plus de situer p entre des limites p^- et p^+.

Pourtant le modèle standard de l'analyse de la décision, le modèle Bayésien, n'admet pas une telle représentation de l'incertitude et exige dans cette situation, comme plus généralement dans toute situation où l'information objective est du type « la loi de probabilité P sur les événements est l'une

(*) Reçu juin 1988, révisé en janvier 1989.
[1] Laboratoire d'Économétrie, Université de Paris-VI, 4, place Jussieu, 75005 Paris.

des lois de la famille $\mathscr{P}$ », que l'on se ramène à une représentation probabiliste de l'incertitude par l'introduction de « probabilités sur les probabilités » : loi de probabilité sur $[p^-, p^+]$ dans notre exemple, loi de probabilité sur Θ, espace de paramètres, dans le cas général où $\mathscr{P}=\{P_\theta, \theta\in\Theta\}$. Le fait que dans de telles situations le choix de cette loi, dite « subjective », apparaisse souvent comme purement arbitraire ne plaide pas en faveur de l'usage du modèle Bayésien.

Encore faut-il disposer d'une solution de rechange. L'objet de cet article est de proposer un modèle de décision acceptant une information objective brute du type « P est dans $\mathscr{P}$ » et conduisant à un critère généralisant le critère de l'utilité espérée.

On sait que dans le modèle Bayésien ce dernier critère se justifie ainsi : Du fait de l'existence de probabilités sur les événements, l'incertitude sur l'effet de chaque décision peut elle-même être décrite par une loi de probabilité sur l'espace des résultats; on admet que le Décideur exprime directement des préférences dans cet ensemble de lois et que ses préférences y satisfont aux axiomes de la théorie de l'utilité linéaire de von Neumann-Morgenstern; la validité du critère de l'utilité espérée découle alors immédiatement de l'existence d'une utilité linéaire.

Nous allons montrer qu'il est possible de suivre une démarche parallèle pour modéliser les choix dans l'incertain imparfaitement probabilisé, à condition toutefois de se limiter à l'*incertain régulier*, c'est-à-dire aux situations de choix où la famille de lois $\mathscr{P}$ décrivant l'information disponible sur les événements est représentable par son enveloppe inférieure f, capacité (fonction d'ensemble croissante) associant à tout événement A le nombre $f(A)=\inf_{P\in\mathscr{P}} P(A)$ et où, de plus, la capacité f est convexe (c'est-à-dire suradditive).

Sous ces hypothèses, en effet (*cf.* § 2.3), l'incertitude sur le résultat de chaque décision δ, décrite par l'ensemble de lois de probabilité $\bar{\mathscr{P}}_\delta$, image de $\mathscr{P}$ par δ, peut elle-même être résumée par l'enveloppe inférieure, $\bar{f}$, de $\bar{\mathscr{P}}_\delta$, qui est encore une capacité convexe (sur l'espace des résultats x).

Or, la théorie de l'utilité linéaire est applicable formellement à des préférences définies parmi des capacités. Quant à la justification des axiomes de rationalité sur lesquels repose cette théorie, elle exige la cohérence de l'opération de combinaison linéaire convexe de capacités et de l'opération correspondante sur les ensembles de lois dont elles sont les enveloppes, cohérence qui n'est obtenue que pour des capacités convexes (*cf.* § 3.2). La restriction de notre modèle à l'incertain régulier est donc tout à la fois techniquement commode et psychologiquement justifiée.

Appliquant la théorie de l'utilité linéaire aux choix dans l'incertain régulier, nous obtenons (section 4) un critère dont l'expression généralise celle du critère de l'utilité espérée du modèle Bayésien. On sait que ce dernier critère attribue une valeur

$$\sum_{x \in \mathscr{X}} \bar{P}(\{x\})\, u(x),$$

(où la fonction d'utilité u dépend du Décideur) à la décision donnant le résultat x avec probabilité $\bar{P}(\{x\})$. Dans notre modèle, la valeur d'une décision δ s'exprime encore comme une somme pondérée de termes

$$\sum_{B \in 2^{\mathscr{X}} \setminus \{\varnothing\}} \varphi(B)\, v(m_B, M_B),$$

où chaque terme est maintenant relatif à un ensemble de résultats B et où :

— les poids $\varphi(B)$ sont reliés (bi-univoquement) à la capacité convexe $\bar{f}$ sur x associée à δ par la formule

$$\bar{f}(A) = \sum_{B \subset A} \varphi(B),$$

qui généralise la relation

$$\bar{P}(A) = \sum_{\{x\} \subset A} \bar{P}(\{x\})$$

du cas probabiliste;

— le niveau d'utilité attribué à B ne dépend que du plus mauvais et du meilleur résultat, m_B et M_B, de B et est intermédiaire entre les utilités attribuées à ces derniers dans le modèle Bayésien :

$$u(m_B) \leqq v(m_B, M_B) \leqq u(M_B).$$

Nous allons maintenant procéder à l'exposé détaillé du modèle, que nous compléterons par une discussion.

2. L'INCERTAIN RÉGULIER ET LA FORMALISATION DES PROBLÈMES DE DÉCISION

2.1. L'information sur les événements et sa représentation

Soit $\mathscr{X} = x_1, \ldots, x_l, \ldots, x_L\}$ *l'ensemble* (fini) *des états de la nature*, $\mathscr{A} 2^{\mathscr{X}}$ l'*algèbre des événements* et $\mathscr{L}$ l'ensemble de toutes les lois de probabilité sur $(\mathscr{X}, \mathscr{A})$.

En situation d'incertitude probabilisée, encore appelée *risque*, à tout état d'information sur les événements correspond une loi P de $\mathscr{L}$. Nous considérons ici une situation d'incertitude plus générale où à tout état d'information est associé un sous-ensemble *non-vide* $\mathscr{P}$ de $\mathscr{L}$. Une interprétation possible d'une telle situation est qu'il existe une loi de probabilité que l'information disponible, trop partielle, permet seulement de localiser dans $\mathscr{P}$.

Tout sous-ensemble non-vide $\mathscr{P}$ de $\mathscr{L}$ a pour enveloppes sa *probabilité inférieure*, f, et sa *probabilité supérieure*, F, applications de $\mathscr{A}$ dans [0, 1] définies, respectivement, par

$$A \mapsto f(A) = \operatorname*{Inf}_{P \in \mathscr{P}} P(A) \qquad \text{et} \qquad A \mapsto F(A) = \operatorname*{Sup}_{P \in \mathscr{P}} P(A); \tag{1}$$

nous écrivons symboliquement $f = \inf\limits_{P \in \mathscr{P}} P$, $F = \operatorname*{Sup}\limits_{P \in \mathscr{P}} P$.

[*N.B.* : De même, nous écrivons : $P \geqq f$ lorsque $P(A) \geqq f(A)$ pour tout A de $\mathscr{A}$; $P = \sum\limits_{i \in I} \alpha_i P_i$ lorsque $P(A) = \sum\limits_{i \in I} \alpha_i P_i(A)$ pour tout A de $\mathscr{A}$; etc.].

De ce que pour tout P de $\mathscr{L}$, tout A de $\mathscr{A}$ et $A^c = \mathscr{X} \setminus A$,

$$P(A) + P(A^c) = 1,$$

on déduit immédiatement que

$$f(A) + F(A^c) = 1, \tag{2}$$

de sorte que F est déterminée par la donnée de f (et vice-versa) et que

$$F \geqq P \geqq f \text{ si et seulement si } P \geqq f. \tag{3}$$

Dans la suite, c'est f que nous privilégions par rapport à F. Cependant, f ne peut en général servir à caractériser $\mathscr{P}$, car *l'ensemble des lois de probabilité compatibles avec f*,

$$\mathscr{P}_f = \{ P \in \mathscr{L} : P \geqq f \} = \{ P \in \mathscr{L} : F \geqq P \geqq f \},$$

n'est lié à $\mathscr{P}$ que par l'inclusion évidente $\mathscr{P} \subset \mathscr{P}_f$, qui peut être stricte.

Exemple 1: $\mathscr{X}\{x_1, x_2, x_3\}$; $\mathscr{P}\{P_1, P_2\}$.

$\mathscr{A}$	$\varnothing$	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathscr{X}$
P_1.	0	1/4	1/2	1/4	3/4	1/2	3/4	1
P_2.	0	1/3	1/6	1/2	1/2	5/6	2/3	1
f.	0	1/4	1/6	1/4	1/2	1/2	2/3	1
P_3.	0	1/4	1/4	1/2	1/2	3/4	3/4	1

La loi P_3 appartient à $\mathscr{P}_f$ mais pas à $\mathscr{P}$, ni non plus à conv $\mathscr{P} = \{Q \in \mathscr{L} : Q = \lambda P_1 + (1-\lambda) P_2,\ \lambda \in [0, 1]\}$, fermeture convexe de $\mathscr{P}$, dont f est encore l'enveloppe inférieure. ■

L'égalité $\mathscr{P}\mathscr{P}_f$ n'est donc pas vérifiée en général, et la convexité de $\mathscr{P}$ (nécessaire puisque $\mathscr{P}_f$ est toujours convexe) ne suffit pas à l'entraîner.

Cependant l'égalité $\mathscr{P}\mathscr{P}_f$ est assurée dans le cas, fréquent, où l'ensemble $\mathscr{P}$ est décrit uniquement à l'aide d'informations du type « la probabilité de A est comprise entre telle et telle valeur » :

$\mathscr{P}\{P \in \mathscr{L} : a_i \leqq P(A_i) \leqq b_i,$ tout $i \in I\}$; ceci est, en effet, évident dès que l'on remarque que pour tout P de $\mathscr{P}_f$ et tout i de I,

$$a_i \leqq f(A_i) \leqq P(A_i) \leqq F(A_i) \leqq b_i.$$

Une première limitation du champ d'application de notre modèle sera que les situations d'incertitude rencontrées devront pouvoir être décrites par un ensemble de lois de probabilité $\mathscr{P}$ telles que $\mathscr{P}\mathscr{P}_f$, c'est-à-dire satisfaisant l'hypothèse :

H. 1. *$\mathscr{P}$ est non vide et vérifie* $\mathscr{P}\{P' \in \mathscr{L} : P' \geqq \operatorname*{Inf}_{P \in \mathscr{P}} P\}$.

Nous pourrons alors caractériser $\mathscr{P}$ par la probabilité inférieure correspondante, f.

2.2. Probabilité inférieures et capacités convexes

Il est facile de vérifier qu'une probabilité inférieure, f, est nécessairement une *capacité* sur $\mathscr{A}$, c'est-à-dire, par définition, que :

(i) $f(\varnothing) = 0$;

(ii) $f(\mathscr{X}) = 1$;

(iii) $[A, B \in \mathscr{A}, A \subset B] \Rightarrow f(A) \leqq f(B)$. (4)

Une capacité est dite *convexe* (ou monotone d'ordre deux) lorsque pour tous A, $B \in \mathscr{A}$,

$$f(A \cup B) + f(A \cap B) \geqq f(A) + f(B). \tag{5}$$

En particulier, on appelle *fonction de croyance* (Shafer [20]) une capacité monotone d'ordre ∞, c'est-à-dire satisfaisant pour tout $K \geqq 2$, tous $A_k \in \mathscr{A}$,

$$f\left(\bigcup_{k=1}^{K} A_k\right) \geqq \sum_{\substack{I \subset \{1, \ldots, K\} \\ I \neq \varnothing}} (-1)^{|I|+1} f\left(\bigcap_{k \in I} A_k\right).$$

Comme le montrent des exemples (Chateauneuf et Jaffray [2]), une probabilité inférieure peut ne pas être une fonction de croyance, ni même être convexe [1]; mais les capacités convexes ont des propriétés particulières qui, comme nous le verrons plus loin (§§ 2.3 et 3.2.), feront de cette classe de capacités le champ d'extension naturel du modèle. Avant d'énoncer la plus importante de ces propriétés, nous avons besoin de quelques définitions.

Étant donné la capacité f et une permutation $S = (x_{i_1}, \ldots, x_{i_l}, \ldots, x_{i_L})$ des éléments de $\mathscr{X}$, on peut définir une loi de probabilité P_S sur $\mathscr{A}$ par

$$P_S(\{x_{i_l}\}) = f(S_l) - f(S_{l-1}), \qquad 1 \leqq l \leqq L, \tag{6}$$

où $S_l = \{x_{i_1}, \ldots, x_{i_l}\}$, $1 \leqq l \leqq L$ et $S_0 = \varnothing$.

Remarquons que $P_S(S_l) = f(S_l)$, $1 \leqq l \leqq$ L, mais que l'on n'a pas nécessairement $P_S \geqq f$. Appelant $\mathscr{S}$ l'ensemble des permutations S des éléments de $\mathscr{X}$, posons $\mathscr{P}_{\mathscr{S}} = \{P_S,\ S \in \mathscr{S}\}$ et introduisons sa fermeture convexe

$$\operatorname{conv} \mathscr{P}_{\mathscr{S}} = \{P \in \mathscr{L} : P = \sum_{S \in \mathscr{S}} \alpha_S P_S \text{ où } \alpha_S \geqq 0, S \in \mathscr{S}, \text{ et } \sum_{S \in \mathscr{S}} \alpha_S = 1\}.$$

---

(1) Les propriétés de convexité d'une capacité et de convexité d'un ensemble de lois de probabilité n'ont aucun lien entre elles : $\mathscr{P}_f$ est convexe, que f soit ou ne soit pas convexe !

On a alors la propriété caractéristique suivante, que nous ne démontrerons pas [2] :

PROPOSITION 1 (Shapley [22], Ichiishi [13]) : *Étant donné une capacité f sur $\mathcal{A}$, les propriétés suivantes sont équivalentes :*

(i) *f est convexe;*

(ii) $\mathcal{P}_f \supset \mathcal{P}_{\mathcal{S}}$;

(iii) $\mathcal{P}_f = \operatorname{conv} \mathcal{P}_{\mathcal{S}}$.

Remarquons que (ii) affirme simplement que pour tout S de $\mathcal{S}$, $P_S \geqq f$ et que, si l'on identifie les lois P de $\mathcal{L}$ aux vecteurs $(P(\{x_1\}), \ldots, P(\{x_l\}), \ldots, P(\{x_L\}))$ du simplexe de $\mathbb{R}^L$, (iii) exprime que les lois P_S de $\mathcal{P}_{\mathcal{S}}$, qui sont des points extrêmes du polyèdre convexe $\mathcal{P}_f$, en constituent en fait tout le profil.

Notons aussi que l'on a alors $f = \operatorname*{Min}_{P \in \mathcal{P}_{\mathcal{S}}} P = \operatorname*{Min}_{P \in \mathcal{P}_f} P$.

2.3. L'information sur les résultats possibles d'une décision et sa représentation

En situation d'incertitude, une *décision* peut être définie comme une application δ qui à tout état de la nature x de $\mathcal{X}$ fait correspondre un élément $\bar{x} = \delta(x)$ d'un ensemble de résultats $\bar{\mathcal{X}}$. Nous supposerons $\bar{\mathcal{X}}$ fini, poserons $\bar{\mathcal{A}} = 2^{\bar{\mathcal{X}}}$ et noterons $\bar{\mathcal{L}}$ l'ensemble des lois de probabilité sur $\bar{\mathcal{A}}$.

Dans le cas particulier de l'incertain probabilisé, l'incertitude sur les résultats possibles de δ est elle-même probabilisée. La loi de probabilité P sur $\mathcal{A}$ a pour image par δ la loi $\bar{P}$ définie en tout $\bar{A}$ de $\bar{\mathcal{A}}$ par $\bar{P}(\bar{A}) = P(\delta^{-1}(\bar{A}))$, que l'on note encore $\bar{P} = P \circ \delta^{-1}$.

Lorsque, plus généralement, l'état de l'information sur les événements de $\mathcal{A}$ est caractérisable par une probabilité inférieure f, c'est-à-dire est compatible avec un ensemble de lois de probabilité $\mathcal{P}_f$, une décision δ offre encore, pour chaque $\bar{A}$ de $\bar{\mathcal{A}}$, la garantie d'obtenir l'un des résultats de $\bar{A}$ avec une probabilité au moins égale à $\inf_{P \in \mathcal{P}_f} P(\delta^{-1}(\bar{A})) = f(\delta^{-1}(\bar{A}))$.

Ceci suggère que l'incertitude sur le résultat de δ soit décrite à l'aide de l'application, $\bar{f} = f \circ 5^{-1}$, de $\bar{\mathcal{A}}$ dans [0, 1] définie par

$$\bar{A} \mapsto \bar{f}(\bar{A}) = f(\delta^{-1}(\bar{A}))$$

---

(2) Établie par les auteurs cités dans le cadre de la théorie des jeux coopératifs, f étant alors la fonction caractéristique du jeu. Pour la démonstration, on peut aussi se reporter à Chateauneuf et Jaffray [2].

qui est une capacité. Mais la difficulté qui, au paragraphe 2.1., nous a conduits à faire l'hypothèse H 1, apparaît à nouveau ici : la donnée de la capacité $\bar{f}$ ne permet de remonter qu'à $\bar{\mathscr{P}}_{\bar{f}}=\{\bar{P}\in\bar{\mathscr{L}}:\bar{P}\geqq\bar{f}\}$, ensemble des lois sur $\bar{\mathscr{A}}$ qui sont compatibles avec $\bar{f}$; or la décision δ ne pouvant induire sur $\bar{\mathscr{A}}$ que les images des lois de $\mathscr{P}$, l'ensemble des lois compatibles avec l'information donnée est en fait

$$\bar{\mathscr{P}}=\{\bar{P}\in\bar{\mathscr{L}}:\bar{P}=P\circ\delta^{-1},\text{ où } P\in\mathscr{P}_f\}$$

$$=\{\bar{P}\in\bar{\mathscr{L}}:\bar{P}=P\circ\delta^{-1},\text{ où } P\in\mathscr{L}\text{ et } P\geqq f\};$$

comme, par définition, $\bar{f}=\operatorname*{Inf}_{\bar{P}\in\bar{\mathscr{P}}}\bar{P}$, l'inclusion $\bar{\mathscr{P}}\subset\bar{\mathscr{P}}_{\bar{f}}$ est évidente, mais il faut l'identité $\bar{\mathscr{P}}=\bar{\mathscr{P}}_{\bar{f}}$ pour que $\bar{f}$ caractérise effectivement l'information sur le résultat de δ. D'où l'intérêt du résultat suivant :

PROPOSITION 2 : *Si la capacité f caractérisant l'information sur les événements de $\mathscr{A}$ est convexe, alors, pour toute décision δ, la capacité $\bar{f}=f\circ\delta^{-1}$ est elle-même convexe et $\bar{\mathscr{P}}=\bar{\mathscr{P}}_{\bar{f}}$.*

Démonstration. — Soit δ une décision; $\bar{f}=f\circ\delta^{-1}$ est convexe puisque, pour tous $\bar{A}$, $\bar{B}\in\bar{\mathscr{A}}$,

$$\bar{f}(\bar{A}\cup\bar{B})+\bar{f}(\bar{A}\cap\bar{B})=f(\delta^{-1}(\bar{A}\cup\bar{B}))+f(\delta^{-1}(\bar{A}\cap\bar{B}))$$

$$=f(\delta^{-1}(\bar{A})\cup\delta^{-1}(\bar{B}))+f(\delta^{-1}(\bar{A})\cap\delta^{-1}(\bar{B}))$$

$$\geqq f(\delta^{-1}(\bar{A}))+f(\delta^{-1}(\bar{B}))=\bar{f}(\bar{A})+\bar{f}(\bar{B}).$$

Soit alors $\bar{P}\in\bar{\mathscr{P}}_{\bar{f}}$; d'après la proposition 1, $\bar{f}$ étant convexe, $\bar{P}$ est de la forme $\bar{P}=\sum_{T\in\mathscr{T}}\beta_T\bar{P}_T$, avec $\beta_T\geqq 0$, $T\in\mathscr{T}$ et $\sum_{T\in\mathscr{T}}\beta_T=1$, où $\mathscr{T}$ est l'ensemble des permutations $T=(\bar{x}_{j_1},\ldots,\bar{x}_{j_k},\ldots,\bar{x}_{j_K})$ des résultats de $\bar{\mathscr{X}}$ et la loi $\bar{P}_T$ est définie par $\bar{P}_T(\{\bar{x}_{j_k}\})=\bar{f}(\bar{T}_k)-\bar{f}(\bar{T}_{k-1})$, $1\leqq k\leqq K$, avec $\bar{T}_k=\{\bar{x}_{j_1},\ldots,\bar{x}_{j_k}\}$, $1\leqq k\leqq K$ et $\bar{T}_0=0$.

La suite d'événements $A_k=\delta^{-1}(\bar{T}_k)$, $0\leqq k\leqq K$, étant croissante, il existe au moins une permutation S des éléments de $\mathscr{X}$ qui énumère successivement les états de la nature de A_1, $A_2\setminus A_1,\ldots,A_k\setminus A_{k-1},\ldots,A_K\setminus A_{K-1}$. Soit P_S la loi de $\mathscr{L}$ associée à S par (6). Les A_k étant des S_l, elle vérifie $P_S(A_k)=f(A_k)=\bar{f}(\bar{T}_k)$, $0\leqq k\leqq K$.

Montrons que $\bar{P}_T = P_S \circ \delta^{-1}$: Tout $\bar{A}$ de $\bar{\mathscr{A}}$ pouvant s'écrire sous la forme $\bar{A} = \bigcup_{k \in K(\bar{A})} \{\bar{x}_{j_k}\}$, il vient

$$\bar{P}_T(\bar{A}) = \sum_{k \in K(\bar{A})} [\bar{f}(\bar{T}_k) - \bar{f}(\bar{T}_{k-1})] = \sum_{k \in K(\bar{A})} [P_S(A_k) - P_S(A_{k-1})]$$
$$= \sum_{k \in K(\bar{A})} P_S(A_k \setminus A_{k-1}) = P_S(\bigcup_{k \in K(\bar{A})} A_k \setminus A_{k-1}),$$

et comme $\delta^{-1}(\bar{A}) = \bigcup_{k \in K(\bar{A})} \delta^{-1}(\{\bar{x}_k\}) = \bigcup_{k \in K(\bar{A})} A_k \setminus A_{k-1}$, nous obtenons bien $\bar{P}_T(\bar{A}) = P_S(\delta^{-1}(\bar{A}))$.

Mais de ce que $\bar{P}_T = P_S \circ \delta^{-1}$ pour une permutation $S = S(T)$ de $\mathscr{S}$, il résulte immédiatement que :

$$\bar{P} = \sum_{T \in \mathscr{T}} \beta_T \bar{P}_T = (\sum_{T \in \mathscr{T}} \beta_T P_{S(T)}) \circ \delta^{-1} = (\sum_{S \in \mathscr{S}} \alpha_S P_S) \circ \delta^{-1},$$

où $\alpha_S = \sum_{S(T) = S} \beta_T \geqq 0$ pour tout $S \in \mathscr{S}$ et $\sum_{S \in \mathscr{S}} \alpha_S = 1$.

La proposition 1 permet d'affirmer que, f étant convexe, $P = \sum_{S \in \mathscr{S}} \alpha_S P_S$ appartient à $\mathscr{P}\mathscr{P}_f$. Comme $\bar{P} = P \circ \delta^{-1}$, nous avons établi l'appartenance à $\bar{\mathscr{P}}$ de tout $\bar{P}$ de $\bar{\mathscr{P}}_{\bar{f}}$; d'où $\bar{\mathscr{P}} \supset \bar{\mathscr{P}}_{\bar{f}}$ et, puisque l'inclusion inverse est évidente, $\bar{\mathscr{P}} = \bar{\mathscr{P}}_{\bar{f}}$. ■

2.4. L'incertain régulier : définition et exemple

Nous appellerons *incertain régulier* toute situation d'incertitude où l'information sur les événements peut être décrite par un ensemble de lois de probabilité $\mathscr{P}$ et où $\mathscr{P}$ et son enveloppe inférieure $f = \inf_{P \in \mathscr{P}} P$ satisfont aux deux hypothèses suivantes :

H 1. *$\mathscr{P}$ est non vide et vérifie $\mathscr{P}\{P' \in \mathscr{L} : P' \geqq f\}$*

et

H 2. *La capacité f est convexe.*

L'hypothèse H 1 permet de caractériser l'information disponible sur les événements par la capacité f; l'hypothèse H 2 rend de même, en vertu de la proposition 2, l'information disponible sur les résultats de toute décision caractérisable par une capacité convexe $\bar{f}$ sur $\bar{\mathscr{A}}$.

Ces propriétés de l'incertain régulier sont illustrées par l'exemple suivant :

Exemple 2 : Tombé en panne, le conducteur d'un camion a fait remplacer la pièce défectueuse de son moteur par une pièce semblable envoyée par le fabricant, qui l'a prise au hasard dans son stock, et qui peut être de qualité variable : b(onne), a(cceptable) ou i(nsuffisante); tout ce qu'il sait est qu'au plus la moitié des pièces sont d'une qualité donnée, information qui peut être résumée par une probabilité inférieure, f, sur les parties de $\{b, a, i\}$:

$\mathscr{A}$	$\varnothing$	$\{b\}$	$\{a\}$	$\{i\}$	$\{b, a\}$	$\{b, i\}$	$\{a, i\}$	$\{b, a, i\}=\mathscr{X}$
f.	0	0	0	0	1/2	1/2	1/2	1

La symétrie de f permet de vérifier facilement que c'est une capacité convexe.

Il peut alors prendre soit la décision δ', d'effectuer le reste du trajet à pleine vitesse avec pour résultats possibles $\delta'(b)=2$ (gain normal) et $\delta'(a)=\delta'(i)=0$ (nouvelle panne), soit la décision δ'', de rouler au ralenti avec pour résultats possibles $\delta''(b)=\delta''(a)=1$ (pénalisation du retard) ou $\delta''(i)=0$ (nouvelle panne). L'incertitude sur les résultats de δ' et de δ'' se traduit, respectivement, par les probabilités inférieures $\bar{f}'$ et $\bar{f}''$ suivantes :

$\bar{\mathscr{A}}$	$\varnothing$	$\{0\}$	$\{1\}$	$\{2\}$	$\{0, 1\}$	$\{0, 2\}$	$\{1, 2\}$	$\{0, 1, 2\}=\bar{\mathscr{X}}$
$\bar{f}'=f\circ\delta'^{-1}$.	0	1/2	0	0	1/2	1	0	1
$\bar{f}''=f\circ\delta''^{-1}$.	0	0	1/2	0	1	0	1/2	1

On vérifie facilement que $\bar{f}'=f\circ\delta'^{-1}$, $\bar{f}''=f\circ\delta''^{-1}$ et que $\bar{f}'$ et $\bar{f}''$ sont bien des capacités convexes; on voit de plus que : (i) comme $\bar{P}\geqq\bar{f}'$ équivaut à $\bar{P}(\{0\})\geqq 1/2$ et $\bar{P}(\{2\})=1-\bar{P}(\{0\})$, toute loi P telle que $P(\{b\})=\bar{P}(\{2\})$, $P(\{a\})\leqq 1/2$ et $P(\{i\})\leqq 1/2$ satisfait $P\geqq f$ et engendre $\bar{P}$ par δ'; (ii) comme $\bar{P}\geqq\bar{f}''$ équivaut à $\bar{P}(\{1\})\geqq 1/2$ et $\bar{P}(\{0\})=1-\bar{P}(\{1\})$, toute loi P telle que $P(\{b\})\leqq 1/2$, $P(\{a\})\leqq 1/2$ et $P(\{i\})=\bar{P}(\{0\})$ satisfait $P\geqq f$ et engendre $\bar{P}$ par δ''. ∎

3. APPLICATION DE LA THÉORIE DE L'UTILITÉ LINÉAIRE AUX DÉCISIONS DANS L'INCERTAIN RÉGULIER

Procédant à l'identification des décisions aux capacités qu'elles engendrent, nous allons ramener tout problème de décision dans l'incertain régulier à un problème de choix dans un ensemble de capacités convexes. La théorie de l'utilité linéaire pouvant formellement s'appliquer dans ce cadre, nous allons rappeler les traits essentiels de cette théorie avant de discuter le bien-fondé de son application aux décisions dans l'incertain régulier.

3.1. La théorie de l'utilité linéaire de von Neumann et Morgenstern

Cette théorie, introduite à l'origine par von Neumann et Morgenstern [18], peut se présenter ainsi :

Appelons *ensemble mixable* (dénomination due à Herstein et Milnor [11]) un ensemble muni d'une opération de composition externe sur [0, 1], associant à tous f, g de $\mathscr{F}$ et tout λ de [0, 1] un élément, noté $\lambda f+(1-\lambda)g$, de $\mathscr{F}$, qui possède les propriétés suivantes :

(i) $1 f+0g=f$;

(ii) $\lambda f+(1-\lambda)g=(1-\lambda)g+\lambda f$;

(iii) $\lambda[\mu f+(1-\mu)g]+(1-\lambda)g=\lambda\mu f+(1-\lambda\mu)g$.

Considérons alors une relation binaire, notée $\succsim$, dans $\mathscr{F}$ et les axiomes suivants (dus à Jensen [16]), où $\succ$ et $\sim$ sont, respectivement, la partie asymétrique et la partie symétrique de $\succsim$.

A1 (*Transitivité et Totalité*). — *$\succsim$ est une relation transitive et complète (préordre complet) dans $\mathscr{F}$.*

A2 (*Indépendance*). — *Pour tous f', f'', g de $\mathscr{F}$ et tout λ de*]0, 1[,

$$f' \succ f'' \text{ entraîne } \lambda f' + (1-\lambda)g \succ \lambda f'' + (1-\lambda)g.$$

A3 (*Continuité*). — *Pour tous f, g, k de $\mathscr{F}$ satisfaisant $f \succ k \succ g$, il existe λ et μ dans*]0, 1[*tels que*

$$\lambda f+(1-\lambda)g \succ k \succ \mu f+(1-\mu)g.$$

On peut alors énoncer :

THÉORÈME 1 (von Neumann et Morgenstern) : *La relation binaire $\succsim$ dans $\mathscr{F}$ vérifie les axiomes* A1, A2 *et* A3 *si et seulement si il existe une application V de $\mathscr{F}$ dans $\mathbb{R}$ telle que :*

(i) *pour tous f, g de $\mathscr{F}$, $f \succsim g$ si et seulement si $V(f) \geqq V(g)$;*

(ii) *pour tous f, g de $\mathscr{F}$ et tout λ de* [0, 1],

$$V(\lambda f+(1-\lambda)g)=\lambda V(f)+(1-\lambda) V(g).$$

De plus, les propriétés (i) *et* (ii) *caractérisent V à une transformation affine strictement croissante près.*

Pour la démonstration de ce théorème, on pourra se reporter, par exemple, à Fishburn ([8] ou [9]).

Lorsque une relation $\succsim$ est interprétée comme une relation de préférence, une fonction V satisfaisant (i) est appelée *fonction d'utilité*; elle est appelée *fonction d'utilité linéaire* lorsqu'elle satisfait de plus (ii).

3.2. Discussion de l'application de la théorie aux problèmes de décision dans l'incertain régulier

[N. B. : *Afin d'alléger les notations, nous désignerons les seuls objets avec lesquels nous aurons à travailler désormais, c'est-à-dire l'ensemble des résultats, l'ensemble de ses parties et les capacités caractérisant les décisions par* $\mathscr{X}$, $\mathscr{A}$, f, *et non* $\bar{\mathscr{X}}$, $\bar{\mathscr{A}}$, $\bar{f}$].

Remarquons tout d'abord que si $\mathscr{F}$ est l'ensemble de toutes les capacités convexes sur $\mathscr{A}$, ensemble des parties d'un ensemble $\mathscr{X}$, l'opération qui à f, g de $\mathscr{F}$ et λ de $[0, 1]$ associe l'application $h=\lambda f+(1-\lambda)g$ de $\mathscr{A}$ dans $[0, 1]$ définie par $A \mapsto h(A)=\lambda f(A)+(1-\lambda)g(A)$, fait de $\mathscr{F}$ un ensemble mixable, car on vérifie facilement que h est bien une capacité convexe et que les propriétés (i), (ii) et (iii) sont satisfaites.

Considérons alors un Décideur en *situation de choix dans l'incertain régulier* avec $\mathscr{X}$ pour ensemble des résultats.

Du point de vue normatif, il est naturel d'imposer aux préférences du Décideur parmi des décisions de dépendre exclusivement des capacités sur $\mathscr{A}$ qui les caractérisent. Il existe alors une relation binaire $\succsim$ dans $\mathscr{F}$ traduisant ces préférences, $f \succ g$ signifiant que s'il avait à choisir entre deux décisions, dont l'une, δ_f, offrirait une incertitude sur les résultats caractérisable par la probabilité inférieure f et l'autre, δ_g, une incertitude caractérisable par g, il préférerait δ_f à δ_g; de même, $f \sim g$ signifiera que δ_f est indifférent à δ_g.

Quant aux propriétés que l'on peut imposer à la relation $\succsim$, remarquons d'abord que l'hypothèse de transitivité, étant naturelle pour des choix, le reste pour $\succsim$ et que les choix évoqués pouvant n'être qu'hypothétiques, il n'est pas absurde de supposer $\succsim$ complète. L'axiome A1 ne paraît donc pas moins fondé que dans toute autre situation de choix.

En ce qui concerne la justification de l'axiome d'indépendance A2, toujours du point de vue normatif, il faut regarder si l'on peut étendre le raisonnement généralement invoqué dans le cas de l'incertain probabilisé, et qui est le suivant : le tirage d'un résultat selon une loi de probabilité $R=\lambda P+(1-\lambda)Q$, mixage des lois P et Q, peut être réalisé par un processus en deux temps, où le résultat est finalement tiré suivant la loi P ou suivant la loi Q selon qu'auparavant l'événement E, de probabilité λ, ou son complémentaire E^c, aura été observé. D'où, si l'on doit choisir entre des mixages

$R'=\lambda P'+(1-\lambda)Q$ et $R''=\lambda P''+(1-\lambda)Q$, on peut imaginer qu'ils soient tirés en deux temps et, remarquant que le tirage final est le même lorsque E^c se produit, mais que, lorsque c'est E qui se produit, il se fait suivant la loi P' si l'on a choisi R', la loi P'' si l'on a choisi R'', on en conclut que si l'on préfère P' à P'', on doit aussi préférer R' à R''.

Pour que ce raisonnement puisse être transposé aux situations d'incertitude plus générales que nous étudions, il faut, et suffit, que l'état d'information caractérisé par une capacité $h=\lambda f+(1-\lambda)g$ soit identique à l'état d'information précédant l'observation d'un événement E de probabilité λ, ou de son complémentaire E^c, qui amèneraient, respectivement, à des états d'information caractérisés par f et g. Comme les lois de probabilité compatibles avec un tel état d'information initial forment l'ensemble

$$\mathcal{R}=\{R=\lambda P+(1-\lambda)Q : P\in\mathcal{P}_f,\ Q\in\mathcal{P}_g\},$$

cette condition n'est autre que l'identité de $\mathcal{P}_h$ et de $\mathcal{R}$. Or,

PROPOSITION 3 : *Étant donné des capacités convexes f et g et leur mixage $h=\lambda f+(1-\lambda)g$, où $\lambda\in[0, 1]$, les ensembles $\mathcal{P}_h=\{R\in\mathcal{L} : R\geqq h\}$ et $\mathcal{R}=\{R\in\mathcal{L} : R=\lambda P+(1-\lambda)Q,\ P\geqq f$ et $Q\geqq g\}$ sont identiques.*

Démonstration : L'inclusion $\mathcal{R}\subset\mathcal{P}_h$ est évidente. Pour établir l'inclusion inverse, introduisons, pour toute permutation S de $\mathcal{X}$ les lois P_S, Q_S et R_S de $\mathcal{L}$ associées respectivement aux capacités f, g et h par (6) ; de ce que, pour $1\leqq l\leqq L$,

$$\begin{aligned}R_S(\{x_{i_l}\})=h(S_l)-h(S_{l-1})&=\lambda[f(S_l)-f(S_{l-1})]\\&\quad+(1-\lambda)[g(S_l)-g(S_{l-1})]\\&=\lambda P_S(\{x_{i_l}\})+(1-\lambda)Q_S(\{x_{i_l}\}),\end{aligned}$$

on déduit que, plus généralement, $R_S=\lambda P_S+(1-\lambda)Q_S$.

Par la proposition 1, nous savons que tout R de $\mathcal{P}_h$ peut s'écrire sous la forme $R=\sum_{S\in\mathcal{S}}\alpha_S R_S$, où $\alpha_S\geqq 0$, $S\in\mathcal{S}$ et $\sum_{S\in\mathcal{S}}\alpha_S=1$; d'où ici

$$R=\sum_{S\in\mathcal{S}}\alpha_S[\lambda P_S+(1-\lambda)Q_S]=\lambda P+(1-\lambda)Q,$$

où l'on a posé

$$P=\sum_{S\in\mathcal{S}}\alpha_S P_S \qquad \text{et} \qquad Q=\sum_{S\in\mathcal{S}}\alpha_S Q_S\,;$$

mais f et g étant des capacités convexes, nous savons, toujours par la proposition 1, que $P \in \mathscr{P}_f$ et $Q \in \mathscr{Q}_g$; donc $R \in \mathscr{R}$. ■

L'exemple suivant montre qu'en revanche, dès que l'une des capacités f et g n'est pas convexe, l'identité de $\mathscr{P}_h$ et $\mathscr{R}$ n'est plus garantie.

Exemple 3 : $\mathscr{X} = \{x_1, x_2, x_3, x_4\}$; f, g et h, capacités données par le tableau :

$\mathscr{A}$	$\varnothing$	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_4\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_1, x_4\}$	$\{x_2, x_3\}$	$\{x_2, x_4\}$	$\{x_3, x_4\}$
f	0	0	1/4	1/4	0	5/12	5/12	0	7/12	1/2	1/2
g	0	0	0	0	0	1/12	1/12	0	1/12	0	0
h	0	0	1/8	1/8	0	1/4	1/4	0	1/3	1/4	1/4

$\mathscr{A}$	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$	$\mathscr{X}$
f	3/4	1/2	1/2	5/6	1
g	5/12	1/4	1/4	1/3	1
h	7/12	3/8	3/8	7/12	1

$h = 1/2f + 1/2g$; la loi R définie par : $R(\{x_1\}) = 1/8$, $R(\{x_2\}) = 5/8$ et $R(\{x_3\}) = R(\{x_4\}) = 1/8$ satisfait $R \geqq h$, mais il n'existe pas $P \geqq f$ et $Q \geqq g$ tels que $R = 1/2\,P + 1/2\,Q$; en effet, on montre successivement que P devrait satisfaire

$$P(\{x_3\}) = 1/4,\ P(\{x_4\}) = 1/4 \qquad \text{et} \qquad P(\{x_1\}) = 0,$$

d'où

$$P(\{x_1, x_3\}) = 1/4 < 5/12 = f(\{x_1, x_3\}).$$

On peut vérifier que g et h sont convexes (et même, en utilisant la proposition 4 du paragraphe 4.1 ci-après, que ce sont des fonctions de croyance), mais que f ne l'est pas, puisque

$$f(\{x_1, x_2, x_3\}) + f(\{x_1\}) < f(\{x_1, x_2\}) + f(\{x_1, x_3\});$$

g et h, comme toutes capacités convexes, sont enveloppes inférieures de familles de lois de probabilité; on peut montrer que f en est également une (*cf.* Chateauneuf et Jaffray, exemple 6 [2]). ■

Il est donc possible de reprendre, dans le cas de l'incertain régulier — et seulement dans ce cas — en l'adaptant, le raisonnement qui servait à justifier l'axiome d'indépendance dans le cas probabilisé : supposons l'incertitude sur le résultat d'une première décision, δ', caractérisée par l'ensemble de loi $\mathscr{P}_{h'}$ où $h' = \lambda f' + (1-\lambda) g$, l'incertitude sur le résultat d'une deuxième décision, δ'', par $\mathscr{P}_{h''}$, où $h'' = \lambda f'' + (1-\lambda) g$. Devant choisir entre δ' et δ'', le Décideur

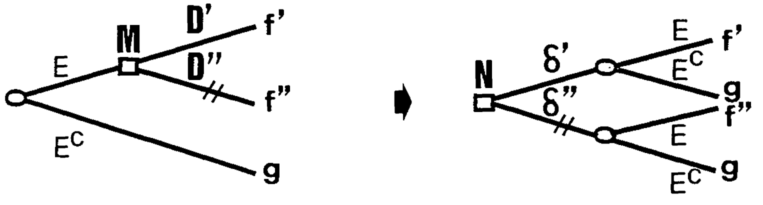

Figure — L'axiome d'indépendance A2 affirme que, préférant en M la décision D' à la décision D'', le Décideur doit en N préférer la décision δ' à la décision δ''.

peut supposer, en vertu de la proposition 3, que cette incertitude est résolue en deux temps et que l'observation de l'événement E, de probabilité λ, transforme l'incertitude sur les résultats en incertitude descriptible par f' (compatibilité avec les lois de $\mathscr{P}_{f'}$) pour δ' et par f'' (compatibilité avec les lois de $\mathscr{P}_{f''}$) pour δ'', alors que l'observation de E^c conduit, pour δ' comme δ'', à l'état d'incertitude g (compatibilité avec les lois de $\mathscr{P}_g$). La décision δ' doit donc être préférée à la décision δ'' dès que l'on préfère l'état d'incertitude décrit par f' à celui décrit par f''.

L'axiome d'indépendance A2 ne paraît donc ni plus ni moins justifiable dans l'ensemble mixable des capacités convexes interprétées comme des probabilités inférieures que dans celui des lois de probabilité.

Examinons enfin l'axiome de continuité A3. Dans le cas de l'incertain probabilisé, on justifie sa première exigence (argument symétrique pour la deuxième) par le fait que, les mixages $R_\lambda = \lambda P + (1-\lambda) Q$ tendant en loi vers la loi P lorsque λ tend vers 1, un Décideur qui préfère que le résultat soit tiré selon la loi P que selon une autre loi K devrait encore préférer au tirage selon la loi K un tirage selon la loi R_λ pour λ suffisamment proche de 1 ; A3 repose donc sur l'idée que des modifications minimes des probabilités qu'offre une décision d'obtenir un résultat appartenant à tel ou tel ensemble A ne doivent pas influer sur son classement par rapport aux autres décisions.

Cet argument s'applique encore dans notre cas, où un état d'incertitude caractérisé par une capacité $\lambda f + (1-\lambda) g$ correspond à la possibilité d'un tirage selon chacune des lois de l'ensemble

$$\mathscr{R}_\lambda = \{ R_\lambda = \lambda P + (1-\lambda) Q : P \in \mathscr{P}_f,\ Q \in \mathscr{P}_g \}$$

qui tendent uniformément en loi vers les lois de $\mathscr{P}_f$ lorsque λ tend vers 1.

Les axiomes A1, A2 et A3 paraissent donc autant justifiés, d'un point de vue normatif, dans le cas de l'incertain régulier, où l'incertitude sur le résultat de chaque décision est caractérisée par une probabilité inférieure convexe, qu'ils l'étaient dans le cas probabiliste. En revanche, les arguments présentés pour justifier A2 et A3 ne pourraient s'étendre à des situations d'incertitude où ces probabilités inférieures ne seraient plus convexes.

En ce sens, l'incertain régulier apparaît bien comme le champ d'extension naturel du modèle de maximisation d'une fonction d'utilité linéaire.

Remarque : En ce qui concerne le point de vue descriptif, il est clair que l'ensemble mixable des capacités convexes contenant celui des probabilités, le Paradoxe d'Allais [1], comme toutes les autres violations des prédictions de la théorie de l'utilité linéaire qui ont été mises en évidence expérimentalement (Kahneman et Tversky [17]) dans le cas du risque fournissent autant d'exemples de comportement réel dans l'incertain régulier ne satisfaisant pas aux exigences axiomatiques précédentes.

4. GÉNÉRALISATION DU CRITÈRE DE L'UTILITÉ ESPÉRÉE

On sait que dans le cas du risque des préférences représentables par une fonction d'utilité linéaire U s'expliquent par un critère d'utilité espérée, en ce sens que, pour toute loi P de $\mathscr{L}$,

$$U(P)=\sum_{x\in S_P} P(\{x\})\,u(x), \tag{7}$$

où S_P est le support de la loi P, et la fonction u, appelée *utilité de von Neumann-Morgenstern* du Décideur, est liée à U par

$$\text{pour tout } x\in\mathscr{X},\quad u(x)=U(\delta_x), \tag{8}$$

δ_x désignant la loi certaine en x.

Dans l'incertain régulier, il existe une expression de l'utilité linéaire qui généralise (7). Avant de donner cette expression qui fait intervenir l'inverse de Möbius de la capacité considérée, il nous faut définir cette inverse et énoncer ses principales propriétés.

4.1. L'inversion de Möbius et ses propriétés

A toute application $f:\mathscr{A}\to\mathbb{R}$ on peut associer son *inverse de Möbius* application $\varphi:\mathscr{A}\to\mathbb{R}$ définie par

$$\varphi(A)=\sum_{B\subset A}(-1)^{|A\setminus B|}f(B); \tag{9}$$

f est caractérisable par son inverse, car

$$f(A)=\sum_{B \subset A} \varphi(B). \tag{10}$$

Cette propriété est démontrée dans Shafer [19], ainsi que dans Chateauneuf et Jaffray [2], où l'on trouvera également la preuve des propriétés suivantes :

PROPOSITION 4 : (1) *Une application f de $\mathscr{A}$ dans $\mathbb{R}$ est une capacité si et seulement si son inverse φ vérifie :*

(i) $\varphi(\varnothing)=0$;

(ii) $\sum_{B \in \mathscr{A}} \varphi(B)=1$; (11)

(iii) *pour tous $A \in \mathscr{A}$ et $x \in A$,* $\sum_{\{x\} \subset B \subset A} \varphi(B) \geqq 0$.

(2) *Une capacité f est convexe si et seulement si pour tous $A \in \mathscr{A}$ et x_1, $x_2 \in A$, $x_1 \neq x_2$,*

$$\sum_{\{x_1, x_2\} \subset B \subset A} \varphi(B) \geqq 0. \tag{12}$$

(3) *Une capacité f est une fonction de croyance si et seulement si φ est non négative.*

Nous appellerons *ensemble focal* d'une capacité f l'ensemble $\mathscr{E}=\{B \in \mathscr{A} : \varphi(B) \neq 0\}$. Notons que $\varnothing \notin \mathscr{E}$.

Une loi de probabilité P est une fonction de croyance dont l'ensemble focal ne contient que des singletons : S_P étant le support de P,

$$\mathscr{E}=\{\{x\} : x \in S_P\} \quad \text{et} \quad \varphi(\{x\})\}=P(\{x\}), \quad x \in S_P.$$

Pour tout $B \in \mathscr{A} \setminus \{\varnothing\}$, on peut définir la *fonction de croyance élémentaire* en B, e_B, comme la capacité définie par :

$$e_B(A)=1 \quad \text{pour } A \supset B, \qquad e_B(A)=0 \text{ sinon};$$

elle a pour ensemble focal $\mathscr{E}_B=\{B\}$, son inverse, φ_B, satisfaisant $\varphi_B(B)=1$ et $\varphi_B(A)=0$ pour $A \neq B$ (puisque $\varphi_B \geqq 0$, c'est bien une fonction de croyance).

Toute capacité f s'exprime comme combinaison linéaire de fonctions de croyance élémentaires puisque, pour tout $A \in \mathscr{A}$,

$$f(A)=\sum_{B \subset A} \varphi(B)=\sum_{B \subset \mathscr{A} \setminus \{\varnothing\}} \varphi(B) e_B(A)=\sum_{B \in \mathscr{E}} \varphi(B) e_B(A),$$

et donc

$$f=\sum_{B\in\mathscr{E}}\varphi(B)\,e_B. \tag{13}$$

Toute fonction de croyance f est alors combinaison linéaire convexe des e_B, puisque $\varphi(B)>0$ pour tout $B\in\mathscr{E}$ et que $\sum_{B\in\mathscr{E}}\varphi(B)=1$.

4.2. Expression de l'utilité linéaire

Dans l'ensemble $\mathscr{F}$ des capacités convexes, l'opération de combinaison linéaire convexe (c. l. c.) de n capacités avec des coefficients $\lambda_i\geqq 0$, $i=1,\ldots,n$, $\sum_{i=1}^{n}\lambda_i=1$, où $h=\sum_{i=1}^{n}\lambda_i f_i$ est définie, pour tout A de $\mathscr{A}$, par $h(A)=\sum_{i=1}^{n}\lambda_i f_i(A)$, coïncide, pour $n=2$, avec l'opération de mixage.

La linéarité de la fonction d'utilité V [théorème 1 (ii)] implique alors la propriété plus générale

$$V\left(\sum_{i=1}^{n}\lambda_i f_i\right)=\sum_{i=1}^{n}\lambda_i V(f_i). \tag{14}$$

L'expression (13), appliquée à une *fonction de croyance* f [afin que les $\varphi(B)$ soient positifs], et la formule (14) conduisent alors à l'expression

$$V(f)=\sum_{B\in\mathscr{E}}\varphi(B)\,V(e_B). \tag{15}$$

Appliquée à une loi P de $\mathscr{P}$ cette expression devient

$$V(P)=\sum_{x\in S_P}P(\{x\})\,V(e_{\{x\}}),$$

où, en définissant la fonction $u:\mathscr{X}\to\mathbb{R}$ par

$$u(x)=V(e_{\{x\}}), \tag{16}$$

on reconnaît la formule de l'utilité espérée

$$V(P)=E_P u=\sum_{x\in S_P}P(\{x\})\,u(x), \tag{17}$$

dont (15) constitue bien la généralisation à l'ensemble des fonctions de croyance.

Dans Jaffray [15] nous avions déjà obtenu l'expression (15) dans le cas où f était une fonction de croyance.

Par une propriété générale des fonctions linéaires [au sens de (ii) du théorème 1] sur un ensemble convexe, le champ de validité de cette expression s'étend en fait, comme nous allons le redémontrer, à l'ensemble des capacités convexes.

PROPOSITION 5 : *Soit V une fonction d'utilité linéaire sur $\mathcal{F}$, ensemble des capacités convexes sur $(\mathcal{X}, \mathcal{A})$, muni du préodre complet $\succsim$. Alors, pour tout f de $\mathcal{F}$,*

$$V(f)=\sum_{B\in\mathcal{E}}\varphi(B)\,V(e_B). \tag{15}$$

Démonstration : Il suffit d'établir la validité de (15) lorsque f n'est pas une fonction de croyance; il existe alors $B_0\in\mathcal{E}$ tel que $\varphi(B_0)<0$, φ désignant l'inverse de f.

Associons à φ les fonctions $\varphi^+=\sup(\varphi, 0)$ et $\varphi^-=\sup(-\varphi, 0)$, puis, remarquant que

$$\alpha=\sum_{B\in\mathcal{E}}\varphi^+(B)>1 \qquad \text{et} \qquad \sum_{B\in\mathcal{E}}\varphi^-(B)=\alpha-1>0,$$

les fonctions

$$\psi^+=\frac{1}{\alpha}\varphi^+ \qquad \text{et} \qquad \psi^-=\frac{1}{\alpha-1}\varphi^-$$

et, enfin, les fonctions de croyance

$$g^+=\sum_{B\in\mathcal{E}^+}\psi^+(B)\,e_B \qquad \text{et} \qquad g^-=\sum_{B\in\mathcal{E}^-}\psi^-(B)\,e_B,$$

où

$$\mathcal{E}^+=\{B\in\mathcal{E}:\varphi(B)>0\} \qquad \text{et} \qquad \mathcal{E}^-=B\in\mathcal{E}:\varphi(B)<0\}.$$

De $f=\sum_{B\in\mathscr{E}}\varphi(B)e_B$, on déduit alors facilement que $g^+=(1/\alpha)f+(1-(1/\alpha))g^-$ et la linéarité de V dans $\mathscr{F}$ entraîne que

$$V(g^+)=\frac{1}{\alpha}V(f)+\left(1-\frac{1}{\alpha}\right)V(g^-).$$

La validité de (15) pour les fonctions de croyance ayant déjà été établie,

$$V(g^+)=\sum_{B\in\mathscr{E}^+}\psi^+(B)V(e_B)\qquad\text{et}\qquad V(g^-)=\sum_{B\in\mathscr{E}^-}\psi^-(B)V(e_B);$$

d'où l'on déduit que

$$\begin{aligned}V(f)&=\alpha V(g^+)-(\alpha-1)V(g^-)\\&=\sum_{B\in\mathscr{E}}[\alpha\psi^+(B)-(\alpha-1)\psi^-(B)]V(e_B)\\&=\sum_{B\in\mathscr{E}}\varphi(B)V(e_B).\quad\blacksquare\end{aligned}$$

4.3. L'axiome de dominance et ses implications

L'intérêt de l'expression (17) de l'utilité linéaire dans le cas du risque est de montrer que les préférences sont complètement déterminées par la donnée de l'application u de $\mathscr{X}$ dans $\mathbb{R}$.

Dans l'incertain, l'expression (15) ne réduit l'information nécessaire à la détermination des préférences qu'à la connaissance de la fonction w,

$$B\mapsto w(B)=V(e_B),$$

qui est une application de $\mathscr{A}$ dans $\mathbb{R}$. D'un point de vue pratique, la différence est énorme : si $|\mathscr{X}|=N$, il y a N nombres $u(x)$, mais 2^N nombres $w(B)$.

Un axiome supplémentaire, qui est un axiome de dominance, va permettre de transformer (15) en une expression plus simple.

Dans son énoncé, nous utiliserons la relation de préférence dans le certain cohérente avec $\succsim$, et notée par le même symbole, relation définie dans $\mathscr{X}$ par

$$x'\succsim x''\quad\Leftrightarrow\quad e_{\{x'\}}\succsim e_{\{x''\}};\tag{18}$$

cette définition ne requiert aucune propriété particulière de $\succsim$ dans $\mathscr{F}$, mais dans le cadre de la théorie de l'utilité linéaire (16) entraînera que

$$x'\succsim x''\quad\Leftrightarrow\quad u(x')\geq u(x'')\tag{19}$$

ce qui permettra de caractériser $\succsim$ dans $\mathscr{X}$ par sa fonction d'utilité u.

Dès que $\succsim$ est un préordre total (axiome A1), tout ensemble B de $\mathscr{A} \setminus \{\varnothing\}$, étant fini, contient un plus mauvais et un meilleur résultat, éléments m_B et M_B (non uniques) satisfaisant

$$m_B, M_B \in B \qquad \text{et} \qquad m_B \precsim x \precsim M_B \quad \text{pour tout } x \text{ de } B \tag{20}$$

ou encore, lorsque $\succsim$ est représentable par u,

$$m_B, M_B \in B \qquad \text{et} \qquad u(m_B) = \min_{x \in B} u(x), \qquad u(M_B) = \max_{x \in B} u(x). \tag{21}$$

L'axiome suivant est donc bien défini, en présence de A1, grâce à (18) et (20), ou, en présence de A1, A2 et A3, grâce à (18) et (21).

A4 (*Dominance*) : *Pour tous B', B'' de $\mathscr{A} \setminus \{\varnothing\}$, si $m_{B'} \succsim m_{B''}$ et $M_{B'} \succsim M_{B''}$ alors $e_{B'} \succsim e_{B''}$.*

Cet axiome peut se justifier ainsi : Rappelons d'abord que, comme nous l'avons remarqué dans la discussion de l'axiome A1, l'existence de la relation $\succsim$ dans $\mathscr{F}$ suppose que *les préférences parmi les décisions dépendent uniquement des capacités qu'elles engendrent dans l'espace des résultats.* Or, lorsque $m_{B'} \succsim m_{B''}$ et $M_{B'} \succsim M_{B''}$, on peut toujours construire (comme l'illustre l'exemple 4 ci-dessous) un espace d'états de la nature $\mathscr{Y}$ et deux décisions δ' et δ'' (applications de $\mathscr{Y}$ dans $\mathscr{X}$), d'images $\delta'(\mathscr{Y}) = B'$ et $\delta''(\mathscr{Y}) = B''$ satisfaisant, pour tout y de $\mathscr{Y}$, $\delta'(y) \succsim \delta''(y)$ (on dit que δ' *domine* δ''). Il est alors naturel d'exiger que δ' soit préférée ou indifférente à δ'', *quelle que soit l'information dont on dispose* sur $(\mathscr{Y}, 2^{\mathscr{Y}})$.

En particulier, il peut y avoir absence complète d'information, cas où l'information, étant compatible avec toute loi de probabilité sur $2^{\mathscr{Y}}$, est caractérisée par la fonction de croyance $e_{\mathscr{Y}}$ (incertain total) [3].

Dans cette situation δ' et δ'' engendrent respectivement les fonctions de croyance $e_{B'}$ et $e_{B''}$ sur $\mathscr{A}$. Le respect du principe de dominance, appliqué aux décisions δ' et δ'', conduit donc ainsi à l'exigence : $e_{B'} \succsim e_{B''}$.

Exemple 4. — $\mathscr{X}\{1, 2, 3, 4\}$; $B' = \{1, 2, 4\}$; $B'' = \{1, 3, 4\}$; $u(x) = x$ (donc $\succsim$, dans $\mathscr{X}$, n'est autre que $\geqq$); d'où $m_{B'} = m_{B''} = 1$; $M_{B'} = M_{B''} = 4$.

([3]) Dans Cohen et Jaffray [4], [5] et Jaffray [14], la notion d'incertain total a été introduite avec un sens voisin, mais dans un contexte différent, les préférences étant définies dans l'ensemble des décisions; c'est alors un théorème qui affirme que, dans l'incertain total, les préférences ne doivent dépendre que des résultats extrêmes de chaque décision et croître (au sens large) avec ces résultats.

Introduisons $\mathscr{Y}\{y_1, y_2, y_3, y_4\}$ et les décisions δ' et δ'' définies par :

$\delta'(y_1)=\delta''(y_1)=1$; $\delta'(y_2)=2$; $\delta''(y_2)=1$;

$\delta'(y_3)=4$, $\delta''(y_3)=3$; $\delta'(y_4)=\delta''(y_4)=4$. δ' domine δ'' (et la domine même strictement); de plus, $\delta'(\mathscr{Y})=B'$ et $\delta''(\mathscr{Y})=B''$.

Remarquons que l'on peut aussi construire ici δ'_1 et δ''_1 telles que δ''_1 domine δ'_1, $\delta'_1(\mathscr{Y})=B'$ et $\delta''_1(\mathscr{Y})=B''$ (ce ne serait pas le cas si l'on avait $m_{B'}>m_{B''}$ ou $M_{B'}>M_{B''}$), si bien que l'on peut justifier à la fois les exigences $e_{B'}\succsim e_{B''}$ et $e_{B''}\succsim e_{B'}$, donc $e_{B'}\sim e_{B''}$. ■

La propriété de l'exemple 4 est générale; de l'axiome A4, il résulte immédiatement que :

PROPOSITION 6 : *Pour tous B', B'' de $\mathscr{A}\setminus\{\varnothing\}$, si $m_{B'}\sim m_{B''}$ et $M_{B'}\sim M_{B''}$ alors $e_{B'}\sim e_{B''}$ et $V(e_{B'})=V(e_{B''})$.*

Il est donc possible de définir une application v de

$$\mathscr{M}\{(m, M)\in\mathscr{X}\times\mathscr{X}: m\precsim M\}$$

dans $\mathbb{R}$ par

$$(m, M)\mapsto v(m, M)=V(e_B),$$

où B est n'importe quel événement de $\mathscr{A}$ tel que $m_B\sim m$ et $M_B\sim M$.

En présence de A4, l'expression (15) s'écrit donc encore

$$V(f)=\sum_{B\in\mathscr{E}}\varphi(B)\,v(m_B, M_B), \tag{22}$$

où les résultats m_B et M_B de B satisfont (21), ce qui ramène la détermination de V à la seule connaissance d'une fonction v définie sur $\mathscr{M}$, donc caractérisée par environ $N^2/2$ nombres.

Le rapprochement de (16) et (22) montre que v et u sont liées par pour tout x de $\mathscr{X}$,

$$u(x)=V(e_{\{x\}})=v(x, x). \tag{23}$$

L'axiome A4 entraîne immédiatement que, compte tenu de (19) et (23), v possède la propriété de croissance suivante :

$[(m', \mathrm{M}'), (m'', \mathrm{M}''), (m', m''), (\mathrm{M}', \mathrm{M}'')\in\mathscr{M}]$

$$\text{entraîne}\quad v(m', M')\leqq v(m'', M''), \tag{24}$$

dont on déduit, en particulier, que

$$[v(m', m') = v(m'', m'') \leqq v(M', M') = v(M'', M'')] \quad \text{entraîne } v(m', M') = v(m'', M'').$$

propriété affirmant que v ne dépend que des classes d'indifférence de ses arguments, ainsi que

$$\text{pour tout } (m, M) \text{ de } \mathscr{M},\ u(m) \leqq v(m, M) \leqq u(M). \tag{25}$$

L'essentiel de nos résultats se trouve résumé par l'énoncé du théorème 2 ci-dessous; l'implication (i) entraîne (ii) vient d'être démontrée; la validité de la réciproque est immédiate si l'on remarque que l'utilité V définie par (22) est automatiquement linéaire, d'où A1, A2 et A3, et que (24) entraîne A4 [(24) assure aussi que $v(m_B, M_B)$ soit bien défini].

THÉORÈME 2 : *Soit $\mathscr{F}$ l'ensemble des capacités convexes sur $(\mathscr{X}, \mathscr{A})$, où $\mathscr{X}$ est fini et $\mathscr{A}2^{\mathscr{X}}$, et soit $\succsim$ une relation binaire dans $\mathscr{F}$.*

Les deux propositions suivantes sont équivalentes :

(1) $\succsim$ *satisfait aux axiomes* A1, A2 *et* A3 *de la théorie de l'utilité linéaire et à l'axiome de dominance* A4;

(ii) $\succsim$ *est représentable par une fonction d'utilité V de la forme*

$$\mathrm{V}(f) = \sum_{B \in \mathscr{E}} \varphi(B)\, v(m_B, \mathscr{M}_B), \tag{22}$$

où : $-v$, définie sur $\{(x, x) : x \in \mathscr{X}\}$, l'est sur tout

$$\mathscr{M} = \{(m, M) \in \mathscr{X} \times \mathscr{X} : v(m, m) \leqq (M, M)\},$$

et vérifie

$$[(m', M'), (m'', M''), (m', m''), (M', M'') \in \mathscr{M}] \quad \text{entraîne } v(m', M') \leqq v(m'', M'') \tag{24}$$

— *pour tout B de $\mathscr{E}$, m_B et M_B satisfont*

$$v(m_B, m_B) = \operatorname*{Min}_{x \in B} v(x, x),\ v(M_B, M_B) = \operatorname*{Max}_{x \in B} v(x, x), \tag{26}$$

— *φ est l'inverse de Möbius de f et $\mathscr{E}$ son ensemble focal.*

De plus, lorsque $\succsim$ *vérifie* (ii), *V et v satisfaisant* (22) *sont uniques modulo une même transformation affine strictement croissante.*

Pour la suite, remarquons que le fait que v ne dépende que des classes d'indifférence de m et M, se traduit par l'existence d'une application F de

$$\mathscr{D}=\{(z_1, z_2)\in\mathbb{R}^2 : z_1, z_2\in u(\mathscr{X}) \text{ et } z_1\leqq z_2\}$$

dans $\mathbb{R}$ telle que $v=F\circ(u, u)$:

$$v(m, M)=F(u(m), u(M)); \tag{27}$$

F est non décroissante et pour tout z de $u(\mathscr{X})$,

$$F(z, z)=z; \tag{28}$$

d'où l'on déduit que pour $z_1, z_2\in\mathscr{D}$,

$$z_1\leqq F(z_1, z_2)\leqq z_2. \tag{29}$$

On peut alors remplacer l'expression (22) par

$$V(f)=\sum_{B\in\mathscr{E}}\varphi(B)F(u_B^-, u_B^+), \tag{30}$$

où

$$u_B^-=\operatorname*{Min}_{x\in B} u(x) \quad \text{et} \quad u_B^+=\operatorname*{Max}_{x\in B} u(x); \tag{31}$$

4.4. Cohérence entre les préférences dans le risque et les préférences dans l'incertain régulier

Étant donné qu'une capacité convexe f sert à caractériser l'état d'incertitude où l'information est compatible avec l'ensemble de lois $\mathscr{P}\{P\in\mathscr{L} : P\geqq f\}$, la cohérence entre les préférences dans le risque du Décideur et ses préférences dans l'incertain régulier exige que $V(f)$ soit compris entre $\operatorname*{Inf}_{P\in\mathscr{P}} V(P)$ et $\operatorname*{Sup}_{P\in\mathscr{P}} V(P)$, c'est-à-dire, d'après (17), entre $\operatorname*{Inf}_{P\in\mathscr{P}} E_P u$ et $\operatorname*{Sup}_{P\in\mathscr{P}} E_P u$.

Or l'intégrale de Choquet [3] de la fonction u par rapport à la capacité f a pour expression, $\mathscr{X}$ étant fini et ses éléments x_l, $l=1, \ldots, L$, rangés par utilité croissante,

$$\int_{\mathrm{Ch}} u\,df=u(x_1)+\sum_{l=2}^{L}(u(x_l)-u(x_{l-1}))\,f(\{x_l, x_{l+1}, \ldots, x_L\})$$

et possède, dès que f est convexe, la propriété

$$\int_{\mathrm{Ch}} u\,df = \operatorname*{Inf}_{P \in \mathscr{P}} E_P u = \sum_{B \in \mathscr{E}} \varphi(B)\, u_B^-, \tag{32}$$

où $u_B^- = u(m_B) = \operatorname*{Min}_{x \in B} u(x)$.

Ce résultat est démontré dans Chateauneuf et Jaffray [2], prop. 10 et cor. 4, et il suffit de l'appliquer à la fonction $-u$ pour montrer que

$$\operatorname*{Sup}_{P \in \mathscr{P}} E_P u = \sum_{B \in \mathscr{E}} \varphi(B)\, u_B^+, \tag{33}$$

où $u_B^+ = u(m_B) = \operatorname*{Max}_{x \in B} u(x)$.

Étant donné l'expression (30) de $V(f)$, il nous suffit pour établir la cohérence des préférences, de démontrer la proposition générale suivante :

Proposition 7 : *Soit f une capacité convexe sur $(\mathscr{X}, \mathscr{A})$, d'inverse φ et d'ensemble focal $\mathscr{E}$. Soit u une application de $\mathscr{X}$ dans $\mathbb{R}$ et F une application de $\mathscr{D}\{(z_1, z_2) \in \mathbb{R}^2 : z_1, z_2 \in u(\mathscr{X})$ et $z_1 \leqq z_2\}$ dans $\mathbb{R}$, non décroissante et telle que, pour tout z de $u(\mathscr{X})$, $F(z, z) = z$.*

Alors, pour $u_B^- = \operatorname*{Min}_{x \in B} u(x)$ et $u_B^+ = \operatorname*{Max}_{x \in B} u(x)$,

$$\sum_{B \in \mathscr{E}} \varphi(B)\, u_B^- \leqq \sum_{B \in \mathscr{E}} \varphi(B)\, F(u_B^-, u_B^+) \leqq \sum_{B \in \mathscr{E}} \varphi(B)\, u_B^+. \tag{34}$$

Démonstration : On peut toujours indicer $\mathscr{X}\{x_1, \ldots, x_l, \ldots, x_L\}$ de manière que $i < l$ entraîne $u(x_i) \leqq u(x_l)$ et donc $u(x_i) \leqq F(u(x_i), u(x_l)) \leqq u(x_l)$.

Introduisons alors pour tout (i, l) de $\mathscr{I}\{(i, l) : 1 \leqq i < l \leqq L\}$ l'ensemble $\mathscr{B}_{i,l} = \{B \in \mathscr{E} : \{x_i, x_l\} \subset B \subset \{x_i, x_{i+1}, \ldots, x_{l-1}, x_l\}\}$. Pour tout B de $\mathscr{B}_{i,l}$, $u_B^- = u(x_i)$ et $u_B^+ = u(x_l)$; d'autre part, il résulte de la propriété (12) de la proposition 4 que $\sum_{B \in \mathscr{B}_{i,l}} \varphi(B) \geqq 0$.

Comme les ensembles $\mathscr{B}_{i,l}$, $(i, l) \in \mathscr{I}$ forment une partition de

$$\mathscr{E}^*\{B \in \mathscr{E} : |B| \geqq 2\},$$

$$\begin{aligned}\sum_{B \in \mathscr{E}^*} \varphi(B)\, u_B^- &= \sum_{(i,l) \in \mathscr{I}} \Big(\sum_{B \in \mathscr{B}_{i,l}} \varphi(B)\Big) u(x_i) \\ &\leqq \sum_{(i,l) \in \mathscr{I}} \Big(\sum_{B \in \mathscr{B}_{i,l}} \varphi(B)\Big) F(u(x_i), u(x_l))\end{aligned}$$

$$\leqq \sum_{(i,\ l) \in \mathscr{I}} \Big(\sum_{B \in \mathscr{B}_{i,\ l}} \varphi(B) \Big) u(x_l) = \sum_{B \in \mathscr{E}^*} \varphi(B) u_B^+ .$$

Comme $F(u(x_i), u(x_l)) = F(u_B^-, u_B^+)$ et comme pour $B \in \mathscr{E} \setminus \mathscr{E}^*$, c'est-à-dire pour $B = \{x\}$, $u_B^- = u_B^+ = F(u_B^-, u_B^+) = u(x)$, on en déduit immédiatement que :

$$\sum_{B \in \mathscr{E}} \varphi(B) u_B^- \leqq \sum_{B \in \mathscr{E}} \varphi(B) F(u_B^-, u_B^+) \leqq \sum_{B \in \mathscr{E}} \varphi(B) u_B^+ . \quad \blacksquare$$

Nous pouvons donc énoncer :

COROLLAIRE : 1. — *Lorsque les préférences $\succsim$ dans $\mathscr{F}$ sont représentables par une fonction d'utilité V de la forme* (22) *du théorème* 2, *l'utilité $V(f)$ de la capacité $f = \operatorname{Inf}_{P \in \mathscr{P}} P$* satisfait

$$\operatorname{Inf}_{P \in \mathscr{P}} E_P u \leqq V(f) \leqq \operatorname{Sup}_{P \in \mathscr{P}} E_P u, \tag{35}$$

où u, utilité de von Neumann-Morgenstern dans le risque, est liée à v de (22) *par*

$$\textit{pour tout } x \textit{ de } \mathscr{X},\ u(x) = v(x, x) \tag{23}$$

et vérifie

$$\textit{pour tout } (m, M) \textit{ de } \mathscr{M}, \quad u(m) \leqq v(m, M) \leqq u(M). \tag{25}$$

Remarques : 1. L'expression (34) de la proposition 7 s'écrit, de manière équivalente :

$$\sum_{B \in \mathscr{E}} \varphi(B) u(m_B) \leqq \sum_{B \in \mathscr{E}} \varphi(B) v(m_B, M_B) \leqq \sum_{B \in \mathscr{E}} \varphi(B) u(M_B). \tag{36}$$

2. Considérons l'axiome de dominance suivant :

A4*. *Pour tout $\bar{x}$ de $\mathscr{X}$ et tout B de $\mathscr{A} \setminus \{\varnothing\}$:*

(i) *Si, pour tout x de B, $x \succsim \bar{x}$, alors $e_B \succsim e_{\{\bar{x}\}}$;*

(ii) *Si, pour tout x de B, $\bar{x} \succsim x$, alors $e_{\{\bar{x}\}} \succsim e_B$.*

Cet axiome, qui est d'interprétation plus immédiate que A4, mais plus faible, n'assure plus l'existence de la représentation (22), mais suffit à entraîner que

$$\text{pour tout } B \text{ de } \mathscr{A} \setminus \{\varnothing\}, \quad u_B^- \leqq V(e_B) \leqq u_B^+,$$

d'où, pour toute fonction de croyance f, par non-négativité de φ,

$$\sum_{B \in \mathscr{E}} \varphi(B) u_B^- \leqq \sum_{B \in \mathscr{E}} \varphi(B) V(e_B) \leqq \sum_{B \in \mathscr{E}} \varphi(B) u_B^+,$$

qui n'est autre que (35).

Cependant cette propriété ne s'étend pas aux capacités convexes, comme le montre l'exemple suivant :

Exemple 5 : $\mathscr{X}\{x_1, x_2, x_3, x_4\}$; f, capacité symétrique par rapport aux x_i, d'inverse φ, donnée par le tableau suivant :

$\mathscr{A}$	$\varnothing$	$\{x_i\}$	$\{x_i, x_j\}$	$\{x_i, x_j, x_k\}$	$\mathscr{X}$
f.	0	0	1/6	1/3	1
φ.	0	0	1/6	$-1/6$	2/3

f est convexe, car (12) de la proposition 4 est vérifiée.

Pour u et V satisfaisant $u(x_i) = i$, $i = 1, 2, 3, 4$
et

$$V(e_B) = u_B^- \quad \text{pour} \quad |B| \neq 3,$$
$$V(e_B) = u_B^+ \quad \text{pour} \quad |B| = 3,$$

on obtient

$$V(f) = \sum_{B \in \mathscr{E}} \varphi(B) V(e_B) = -1/6 < 3/2 = \sum_{B \in \mathscr{E}} \varphi(B) u_B^- = \operatorname*{Inf}_{P \in \mathscr{S}} E_P u. \quad \blacksquare$$

4.5. Détermination des préférences

La formule (23) montre que les préférences d'un Décideur dans l'incertain régulier sont déterminées dès que l'on connaît v.

Nous allons voir que l'on peut construire v, en construisant séparément deux fonctions : La première, u, est l'utilité de von Neumann-Morgenstern du Décideur, qui caractérise, comme on le sait, son attitude vis-à-vis du risque. La seconde, α, caractérise son attitude vis-à-vis de l'*ambiguïté*, au sens d'Ellsberg [7] et est appelée *indice de pessimisme local* de Hurwicz [12]. α est une application de $\mathscr{M}$ dans [0,1] et sa valeur $\alpha(m, M)$ en chaque (m, M) de $\mathscr{M}$ est caractérisée par l'indifférence du Décideur entre :

(i) recevoir un résultat au moins aussi bon que m et au plus aussi bon que M, sans autre information;
et

(ii) recevoir soit m, avec probabilité $\alpha(m, M)$, soit M, avec probabilité $(1-\alpha(m, M))$.

Comme l'incertitude sur le résultat de la première éventualité correspond à la fonction de croyance élémentaire e_B, où $B=\{x \in \mathscr{X} : m \lesssim x \lesssim M\}$ et celle sur le résultat de la deuxième à une loi de probabilité de la forme

$$P_\alpha = \alpha e_{\{m\}} + (1-\alpha) e_{\{M\}} \qquad \text{avec} \quad \alpha = \alpha(m, M),$$

nécessairement $V(e_B) = V(P_{\alpha(m, M)})$ ce qui, en vertu de (23) et (17) équivaut à :

$$v(m, M) = \alpha(m, M) u(m) + (1-\alpha(m, M)) u(M). \tag{37}$$

Il existe des méthodes bien connues de construction de u (Hershey, Kunreuther et Schoemaker [10]); pour construire α, on peut, par exemple, proposer au Décideur des séries de comparaisons du type (i), (ii) ci-dessus. v est alors déterminée par u et α selon la formule (37). Cette formule met en évidence la simplicité avec laquelle l'attitude vis-à-vis du risque et l'attitude vis-à-vis de l'ambiguïté se combinent pour former les préférences dans l'incertain régulier.

4.6. Exemple

Reprenons l'exemple 2 avec ses notations (espace des états de la nature $(\mathscr{X}, \mathscr{A})$, espace des résultats $(\bar{\mathscr{X}}, \bar{\mathscr{A}})$, etc.)

Exemple 2 (suite) : Calculons d'abord, à l'aide de (9), les inverses φ de f, ψ' de $\bar{f}'$ et ψ'' de $\bar{f}''$:

$\mathscr{A}$	$\varnothing$	$\{b\}$	$\{a\}$	$\{i\}$	$\{b, a\}$	$\{b, i\}$	$\{a, i\}$	$\{b, a, i\}=\mathscr{X}$
$\varnothing$	0	0	0	0	1/2	1/2	1/2	$-1/2$

$\bar{\mathscr{A}}$	$\varnothing$	$\{0\}$	$\{1\}$	$\{2\}$	$\{0,1\}$	$\{0,2\}$	$\{1,2\}$	$\{0,1,2\}=\bar{\mathscr{X}}$
ψ'	0	1/2	0	0	0	1/2	0	0
ψ''	0	0	1/2	0	1/2	0	0	0

On peut vérifier que (10) redonne bien f, $\bar{f}'$ et $\bar{f}''$.

φ satisfaisant (11), mais prenant une valeur négative en $\mathscr{X}$, f est une capacité convexe, mais n'est pas une fonction de croyance. ψ' et ψ'' étant non négatives, $\bar{f}'$ et $\bar{f}''$ sont des fonctions de croyance.

Étant donné l'utilité u, vérifiant $u(0) < u(1) < u(2)$, on peut calculer

$$\sum_{\bar{B} \in \bar{\mathscr{E}}'} \psi'(\bar{B}) u_{\bar{B}}^- = u(0) \qquad \text{et} \qquad \sum_{\bar{B} \in \bar{\mathscr{E}}'} \psi'(\bar{B}) u_{\bar{B}}^+ = 1/2 [u(0) + u(2)]$$

et vérifier que ces valeurs sont respectivement égales à $\inf_{\bar{P} \geq \bar{f}'} E_{\bar{P}} u$ et $\sup_{\bar{P} \geqq \bar{f}'} E_{\bar{P}} u$. Introduisant ensuite v et α liées à u par (37), on calcule

$$V(\bar{f}') = 1/2\, u(0) + 1/2\, [\alpha(0,2)\, u(0) + (1-\alpha(0,2))\, u(2)]$$

soit encore

$$V(\bar{f}') = \alpha(0,2)\, u(0) + (1-\alpha(0,2)) \frac{u(0)+u(2)}{2},$$

qui, puisque $\alpha(0,2) \in [0,1]$, montre que (35) est satisfait.

Pour $\bar{f}''$ les calculs donnent, semblablement,

$$\sum_{\bar{B} \in \bar{\mathscr{E}}''} \psi''(\bar{B})\, u_{\bar{B}}^{-} = 1/2\, [u(0) + u(1)], \qquad \sum_{\bar{B} \in \bar{\mathscr{E}}''} \psi''(\bar{B})\, u_{\bar{B}}^{+} = u(1)$$

et

$$\begin{aligned} V(\bar{f}'') &= 1/2\, u(1) + 1/2\, [\alpha(0,1)\, u(0) + (1-\alpha(0,1))\, u(1)] \\ &= \alpha(0,1) \frac{u(0)+u(1)}{2} + (1-\alpha(0,1))\, u(1). \end{aligned}$$

Le Décideur préfère donc δ' à δ'' si et seulement si $V(\bar{f}') > V(\bar{f}'')$, c'est-à-dire lorsque

$$(1-\alpha(0,2))\,(u(2)-u(1)) > (1+\alpha(0,2)-\alpha(0,1))\,(u(1)-u(0)).$$

On peut en déduire, par exemple, qu'aucun Décideur adversaire du risque (attitude caractérisée ici par : $u(2)-u(1) \leqq u(1)-u(0)$) et de coefficient de pessimisme constant (d'où $\alpha(0,2) = \alpha(0,1)$) ne préfére δ' à δ''. ■

5. DISCUSSION ET CONCLUSION

Nous venons de présenter un modèle de décision dans l'incertain régulier qui propose un critère de choix généralisant de façon naturelle le critère de l'utilité espérée. L'intérêt de ce modèle dépend évidemment de son adéquation à la représentation des problèmes de décisions réels et de l'existence de techniques de résolution efficaces.

En ce qui concerne le premier point, l'existence d'applications potentielles découle de la propriété suivante, établie par Dempster [6] : Chaque fois que l'information prend la forme d'un message aléatoire imprécis, elle engendre sur les événements une fonction de croyance, C'est là une situation courante puisque toute enquête statistique (échantillonnage, sondage) dont les données sont ambiguës ou incomplètes fournit une information de ce type. Cependant, on rencontre aussi des situations où les événements ont une probabilité inférieure qui n'est pas monotone d'ordre infini, mais seulement d'ordre deux, comme le montre l'exemple 2 du paragraphe 2.4, simple prototype d'une large classe de problèmes, dont la résolution nécessitait l'extension à l'incertain régulier du modèle présenté dans Jaffray [15].

Il ne devrait pas y avoir de difficulté conceptuelle à lever l'hypothèse, très limitative, que l'ensemble des états de la nature $\mathscr{X}$ est fini; en effet, si l'inverse de Möbius n'existe que dans le cas fini, il existe, dans le cas général, une mesure (sur une sous-algèbre des parties de parties de $\mathscr{X}$) aux propriétés analogues (Choquet [3], Revuz [18], Shafer [21]). Cependant, aussi bien la détermination de cette mesure que son maniement risquent de se révéler assez complexes.

Cette remarque nous conduit au second point : lorsque l'on passe du risque à l'incertain régulier, les problèmes pratiques posés par l'usage du modèle de l'utilité linéaire augmentent très sensiblement, non seulement parce que les capacités sont plus complexes que les lois de probabilité, mais aussi parce que la technique des arbres de décision ne se généralise pas : au mixage de deux décisions ne correspond pas le mixage des capacités qu'elles engendrent et, de toute façon, prendre une décision conditionnellement à un événement, une autre conditionnellement à un événement complémentaire, ne constitue pas une décision mixte, les événements n'ayant pas de probabilités ! Il reste donc à développer des méthodes efficaces pour caractériser une capacité, évaluer son utilité, calculer l'effet du conditionnement par un événement, etc.; l'avenir du modèle en dépend de façon cruciale.

Dans le domaine de l'aide à la décision dans l'incertain, où le manque d'outils valables est flagrant, nous pensons que ce modèle offre quelque espoir.

REMERCIEMENTS

Je remercie Alain Chateauneuf, Peter Wakker et deux rapporteurs anonymes, dont Denys Bouyssou, pour leurs critiques et conseils, que j'ai essayé de mettre à profit dans cette nouvelle rédaction.

BIBLIOGRAPHIE

1. M. ALLAIS, The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulate and Axioms of the American School (1952, traduit du Français), dans M. ALLAIS et O. HAGEN Ed., *Expected Utility Hypotheses and the Allais Paradox*, D. Reidel, Dordrecht, 1979, p. 27-145.
2. A. CHATEAUNEUF et J. Y. JAFFRAY, Some Characterizations of Lower Probabilities and Other Monotone Capacities through the Use of Möbius Inversion, *Mathematical Social Sciences*, vol. 17, 1989 (à paraître).
3. G. CHOQUET, Théorie des capacités, *Ann. Inst. Fourier* (*Grenoble*), V, 1953, p. 131-295.
4. M. COHEN et J. Y. JAFFRAY, Rational Behavior Under Complete Ignorance, *Econometrica*, vol. 48, p. 1281-1299.
5 M. COHEN et J. Y. JAFFRAY, Approximations of Rational Criteria Under Complete Ignorance and the Independence Axiom. *Theory and Decision*, vol. 15, 1983, p. 121-150.
6. A. P. DEMPSTER, Upper and Lower Probabilities Induced by a Multivalued Mapping, *Annals of Mathematical Statistics*, vol. 38, 1967, p. 325-339.
7. D. ELLSBERG, Risk, Ambiguity and the Savage Axioms, *Quaterly J. of Economics*, vol. 75, 1961, p. 643-669.
8. P. C. FISHBURN, *Utility theory for decision making*, Wiley, New York, 1970.
9. P. C. FISHBURN, *The foundations of Expected Utility*, D. Reidel Publishing Company, Dordrecht, 1982.
10. J. C. HERSHEY, H. C. KUNREUTHER et P. J. SCHOEMAKER, Sources of bias in assessment procedures for utility functions, *Management Sci.*, vol. 28, 1986, p. 936-954.
11. I. N. HERSTEIN et J. MILNOR, An Axiomatic Approach to Measurable Utility, *Econometria*, vol. 21, 1953, p. 291-297.
12. L. HURWICZ, Optimality Criteria for Decision Making Under Ignorance, Cowles Commission discussion paper, Statistics #370, 1951, mimeographed.
13. T. ICHIISHI, Super-Modularity: Applications to Convex Games and to the Greedy Algorithm for LP, *J. Econom. Thory*, vol. 25, 1981, p. 283-286.
14. J. Y. JAFFRAY, Jeux contre la Nature, *Economies et Sociétés, XIV*, 1980, *p.* 1345-1367.
15. J. Y. JAFFRAY, Linear Utility Theory for Belief Functions, *Operations Research Letters*, vol. 8, 1989, p. 107-112.
16. N. E. JENSEN, An Introduction to Bernoullian Utility Theory. I: Utility Functions, *Swedish Journal of Economics*, vol. 69, 1967, p. 163-183.
17. D. KAHNEMAN et A. TVERSKY, Prospect Theory: an Analysis of Decision Under Risk, *Econometrica*, vol. 47, 1979, p. 263-291.
18. J. VON NEUMANN et O. MORGENSTERN, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, 1944.
19. A. REVUZ, Fonctions croissantes et mesures sur les espaces topologiques ordonnés, *Ann. Instit. Fourier* (Grenoble), VI, 1955, p. 187-269.
20. G. SHAFER, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, 1976.
21. G. SHAFER, allocations of probability, *Ann. Prob.*, vol. 7, 1979, p. 827-839.
22. L. S. SHAPLEY, Cores of Convex Games, *Internat. J. Game Theory* 1, 1971, p. 11-26.