

M1 IAD UE Décision et Jeux

Notes de cours (2)

Jean-Yves Jaffray Patrice Perny

28 février 2007

DECISION DANS LE RISQUE

1 Introduction

1.1 Décision dans l'incertitude et décision dans risque

Ce chapitre traite de la *décision dans l'incertitude*, c.-à-d. de situations de choix où les résultats des actions ne peuvent être prévus avec certitude. Nous supposons que cette incertitude est probabilisée, c.-à-d. que le résultat obtenu ne dépend que de la réalisation d'événements de probabilités connues. On utilise alors le terme de *décision dans le risque*.

Dans quelle mesure cette hypothèse est-elle restrictive? Il existe une école de pensée, l'école bayésienne (du nom de BAYES), fondée par DE FINETTI et SAVAGE, qui soutient que tout décideur rationnel doit se comporter comme si tous les événements avaient des probabilités, celles-ci pouvant varier d'une personne à l'autre, d'où leur dénomination de "*probabilités subjectives*". Cette conviction est loin de faire l'unanimité (en particulier chez les statisticiens). Les non-bayésiens considèrent les situations de risque comme un cas particulier des situations d'incertitude: c'est pour eux le comportement du décideur qui permet de reconnaître s'il attribue des probabilités aux événements et, si oui, quelles sont leurs valeurs.

1.2 Un bref historique

La proposition de PASCAL (1654) pour un partage équitable des sommes mises dans une partie interrompue marque à la fois la naissance du calcul des probabilités et de la théorie de la décision dans le risque. Son évaluation s'appuyait sur la comparaison des *espérances mathématiques de gain* (EG). La justification de ce critère est claire dans le cas de situations de décision répétitives et à aléas indépendants d'une situation à l'autre: la loi des grands

nombres nous assure qu'en répétant n fois une décision offrant à chaque coup une EG positive, nous aurons une probabilité tendant vers 1 avec n d'obtenir un gain moyen positif et un gain cumulé arbitrairement grand.

Que vaut le critère EG pour des situations décisionnelles qui ne se rencontrent qu'une seule fois? Prenons le cas d'un jeu à *Pile* ou *Face* avec une pièce de monnaie non biaisée ($Pr(Pile) = Pr(Face) = \frac{1}{2}$). Miser 1€ sur *Pile* pour un gain (brut) de 2€ donne une EG de $\frac{1}{2}(2 - 1) + \frac{1}{2}(-1) = 0$ comme le *statu quo* (ne pas jouer); le critère EG prescrit l'indifférence entre jouer et ne pas jouer, ce qui ne paraît pas déraisonnable. Augmentons l'importance du pari: mise de 1000€ sur *Pile* pour un gain (brut) de 2000€ ou encore: mise de 1 M€ sur *Pile* pour un gain (brut) de 2 M€; les jeux restent encore équitables (au sens de l'EG); pourtant la plupart des gens refuseraient de miser à ce jeu et même à certains jeux favorables comme celui où la mise serait de 1M € sur *Pile* pour un gain de 2.5 M €.

Quelle est l'explication?

D.BERNOULLI, réfléchissant, en 1738, au *paradoxe de SAINT-PETERSBOURG* (personne n'est prêt à miser plus de quelques € à un certain jeu, malgré une EG infinie), parvint à la conclusion que c'était dû au fait que l'*utilité marginale* des gains est rapidement décroissante: un € de plus augmente plus votre utilité si vous êtes pauvre que si vous êtes riche. Il proposa donc de remplacer le critère de l'espérance de gain, EG, par un *critère de l'espérance d'utilité du gain*, EU, avec pour fonction d'utilité dans le certain une fonction *concave*, la fonction $u(x) = \log_2(c_0 + x)$, où c_0 est la fortune initiale du joueur.

Ce critère est compatible avec le refus de jouer à des jeux équitables et même favorables; par exemple, pour une fortune initiale $c_0 = 2M€$ l'utilité espérée du jeu favorable où l'on mise 1 M€ sur *Pile* pour un gain de 2.5 M€ est $EU = \frac{1}{2} \log_2[c_0 + (2.5 - 1) \times 10^6] + \frac{1}{2} \log_2[c_0 - 1 \times 10^6] = \frac{1}{2} \log_2[3.5 \times 10^6] + \frac{1}{2} \log_2[1 \times 10^6] = \frac{1}{2} \log_2[3.5 \times 10^{12}]$, inférieure à celle du *statu quo*, $EU = \log_2[2 \times 10^6]$, (car $3.5 < 4$).

L'explication de D.BERNOULLI suppose l'existence d'une *utilité cardinale dans le certain*, en ce sens que le classement des décisions dans le risque par un critère EU n'est pas modifié quand on remplace la fonction \log_2 par une fonction qui en est une transformée affine strictement croissante, par exemple par \ln ou $2\ln - 3$, mais peut changer si l'on opère une transformation strictement croissante qui n'est pas affine.

La théorie économique moderne a renoncé à introduire une utilité cardinale dans le certain, bien que son existence soit confirmée par l'introspection (tout le monde semble d'accord sur le fait que devenir deux fois plus riche ne rend pas deux fois plus heureux!); la raison en est qu'il est impossible de donner une confirmation valable de son existence; en effet, la psychologie

expérimentale considère qu'une observation n'a de valeur scientifique que si les alternatives de choix envisagées sont effectivement proposées aux sujets ; comme un sujet est à un moment donné riche ou pauvre, mais n'est pas les deux simultanément, on ne peut pas lui proposer d'alternative de choix permettant de vérifier s'il serait, ou non, plus content, étant riche, de recevoir 100 € que de recevoir, étant pauvre, 10 € .

Pourtant le critère usuel en décision dans le risque est celui de l'espérance d'utilité, EU ; mais on a renoncé à fournir une interprétation à la fonction d'utilité, u , impliquée. C'est simplement un paramètre, caractéristique du comportement du décideur, que l'on sait déterminer, ce qui est l'essentiel.

Dans le modèle de VON NEUMANN et MORGENSTERN que nous allons étudier maintenant, c'est du respect par le décideur d'un certain nombre d'axiomes de comportement que va découler nécessairement la validité du critère EU.

2 La théorie de l'utilité linéaire de VON NEUMANN et MORGENSTERN

2.1 Lois de probabilité engendrées par les décisions

Dans le risque, c.-à-d. dans l'incertitude probabilisée, toute action/décision, comme toute stratégie, suite d'actions éventuellement conditionnelles à des événements, apporte un résultat aléatoire caractérisé par une loi de probabilité P sur l'ensemble \mathcal{C} des résultats. On dit que cette action *offre* la *perspective aléatoire* P sur \mathcal{C} , ou encore qu'elle *engendre* la loi de probabilité P sur \mathcal{C} .

On appellera *loteries* les lois de probabilité sur \mathcal{C} dont le *support* \mathcal{S} est fini, c.-à-d. telles qu'il existe $\mathcal{S} = \{c_i, i \in I\}$, I fini, tel que $P(\mathcal{S}) = 1$. Une loterie P est caractérisable par la donnée des probabilités élémentaires $p_i = P(\{c_i\})$, $i \in I$; notons que puisque $p_i \geq 0, \forall i \in I$ et que $\sum_{i \in I} p_i = 1$, le vecteur $(p_i, i \in I)$ appartient au simplexe de $\mathbb{R}^{|I|}$.

La définition des lois à support infini exige que l'on munisse \mathcal{C} d'une σ -algèbre de parties, \mathcal{G} ; on supposera toujours que les singletons $\{c\}$ appartiennent à \mathcal{G} , afin que $P(\{c\})$ existe. Le plus souvent une loi P sera donnée par sa densité p par rapport à une mesure μ sur $(\mathcal{C}, \mathcal{G})$, de sorte que la probabilité que le résultat appartienne à un ensemble G sera $P(G) = \int_G p(c) d\mu$.

On notera δ_c la loi de support le singleton $\{c\}$, donc donnant le résultat c avec probabilité 1 ; on l'appelle *loi certaine*, ou loi de DIRAC, en c .

Exemple 1. Miser 1 M€ sur *Pile* pour un gain de 2 M€ est une action engendrant sur $\mathcal{C} = \mathbb{R}$, espace des gains nets (algébriques) la loterie P , de support $\{+1, -1\}$, avec $P(\{+1\}) = Pr(Pile)$ et $P(\{-1\}) = Pr(Face)$; refuser de miser est aussi une action, le *statu quo*, donnant la loi certaine δ_0 . Jouer une première fois au jeu précédent, puis y rejouer si *Pile* est sorti et seulement dans ce cas, offre pour perspective aléatoire \mathcal{C} la loterie Q de support $\{+2, 0, -1\}$, avec $Q(\{+2\}) = Pr(Pile, Pile)$, $Q(\{0\}) = Pr(Pile, Face)$ et $Q(\{-1\}) = Pr(Face)$.

On prend parfois pour ensemble de résultats \mathcal{C}' les *états de fortune* finaux; pour une fortune initiale c_0 , jouer engendre sur \mathcal{C}' la loi P' telle que $P'(\{c_0+1\}) = Pr(Pile)$ et $P'(\{c_0-1\}) = Pr(Face)$ et ne pas jouer la loi δ_{c_0} . De même, rejouer si *Pile* est sorti offre pour perspective aléatoire sur \mathcal{C}' la loterie Q' telle que: $Q'(\{c_0+2\}) = Pr(Pile, Pile)$, $Q'(\{c_0\}) = Pr(Pile, Face)$ et $Q'(\{c_0-1\}) = Pr(Face)$.

Les deux formalisations sont évidemment équivalentes.

2.2 Mélanges de lois de probabilité

Définition 1. Etant donné deux lois de probabilité quelconques, P, Q sur $(\mathcal{C}, \mathcal{G})$ et un nombre quelconque $\lambda \in [0, 1]$ on appelle *mélange des lois P et Q avec poids respectifs λ et $(1 - \lambda)$* la loi de probabilité, R , définie par :

$$\forall G \in \mathcal{G}, R(G) = \lambda P(G) + (1 - \lambda)Q(G)$$

on vérifie facilement que c'est bien une loi de probabilité; par exemple, pour $G \cap H = \emptyset$, $R(G \cup H) = \lambda P(G \cup H) + (1 - \lambda)Q(G \cup H) = \lambda[P(G) + P(H)] + (1 - \lambda)[Q(G) + Q(H)] = [\lambda P(G) + (1 - \lambda)Q(G)] + [\lambda P(H) + (1 - \lambda)Q(H)] = R(G) + R(H)$.

On note symboliquement $R = \lambda P + (1 - \lambda)Q$ ce mélange.

Définition 2. Plus généralement, on peut définir le mélange “d'ordre n ” $R = \sum_{i=1}^n \lambda_i P_i$ de n lois P_i ($i = 1, \dots, n$) avec poids λ_i ($i = 1, \dots, n$), tels que $\lambda_i \geq 0$, $\sum_{i=1}^n \lambda_i = 1$ par

$$\forall G \in \mathcal{G}, R(G) = \sum_{i=1}^n \lambda_i P_i(G);$$

Remarque 1 on emploie aussi le terme de “mixage” pour “mélange” (anglais: “mixture”).

Remarque 2 L'ensemble des lois de probabilité sur $(\mathcal{C}, \mathcal{G})$ peut être considéré comme un ensemble convexe d'un espace vectoriel (celui des mesures signées), l'opération de mélange n'étant autre que la combinaison linéaire convexe dans cet espace.

Exemple 2. On vous propose de miser 1 M€ sur *Pile* ($Pr(Pile) = Pr(Face) = \frac{1}{2}$) pour un gain de 2 M€ ; vous décidez de choisir aléatoirement si vous allez accepter ou non de jouer en lançant un dé (non pipé) et de dire *oui* si le 5 ou le 6 sort et *non* sinon.

Vous êtes alors devant la perspective aléatoire $R = \frac{1}{3}P + \frac{2}{3}\delta_0$, loi de support $\{+1, 0, -1\}$ avec : $R(\{+1\}) = \frac{1}{3}P(\{+1\}) + \frac{2}{3}\delta_0(\{+1\}) = \frac{1}{6}$; $R(\{0\}) = \frac{1}{3}P(\{0\}) + \frac{2}{3}\delta_0(\{0\}) = \frac{2}{3}$; $R(\{-1\}) = \frac{1}{3}P(\{-1\}) + \frac{2}{3}\delta_0(\{-1\}) = \frac{1}{6}$.

Exemple 3. En statistique, on utilise comme approximations d'autres lois des *mélanges de gaussiennes*, $R = \sum_{i=1}^n \lambda_i \mathcal{N}(m_i, \sigma_i)$, où $\mathcal{N}(m_i, \sigma_i)$ est la loi normale (dite aussi de LAPLACE-GAUSS) de moyenne m_i et d'écart-type σ_i .

La propriété suivante nous sera utile.

Proposition 1. *Toute loterie, c.-à-d. toute loi de probabilité discrète P à support $\mathcal{S} = \{c_i, i \in I\}$, I fini peut s'écrire comme mélange des lois certaines δ_{c_i} (c_i résultats du support de P), avec comme poids les probabilités $P(\{c_i\})$ d'obtenir ces résultats avec la loi P :*

$$P = \sum_{i=1}^n P(\{c_i\}) \delta_{c_i}$$

Démonstration

$$\begin{aligned} \forall c \in \mathcal{C}, \delta_{c_i}(\{c\}) &= 1, \text{ si } c = c_i \\ &= 0, \text{ si } c \neq c_i \end{aligned}$$

D'où,

$$\begin{aligned} \sum_{i=1}^n P(\{c_i\}) \delta_{c_i}(\{c\}) &= P(\{c_{i_0}\}), \text{ si } c = c_{i_0} \text{ pour un } i_0 \in I \\ &= 0, \text{ sinon} \end{aligned}$$

Pour mieux comprendre le sens et la portée des axiomes de comportement dans le risque que nous énoncerons plus loin, il est utile de remarquer que l'on est placé, à une date $t = 0$, devant la perspective aléatoire offerte par le mélange $R = \lambda P + (1 - \lambda)Q$ lorsque l'on sait que :

à $t = 1$, on apprendra si un événement E , de probabilité λ , est réalisé ou si, au contraire, c'est son complémentaire \bar{E} , de probabilité $1 - \lambda$, qui l'est ;
à $t = 2$, on obtiendra un résultat tiré selon la loi P ou la loi Q , selon que c'est E ou \bar{E} qui aura été réalisé précédemment.

En effet, la probabilité d'obtenir un résultat faisant partie de l'ensemble $G \in \mathcal{G}$ est alors, d'après la formule des probabilités composées : $Pr(G) = Pr(E)Pr(G/E) + Pr(\bar{E})Pr(G/\bar{E}) = \lambda P(G) + (1 - \lambda)Q(G) = R(G)$.

Conformément à la symbolique des arbres de décision dans le risque, qui ne sera introduite que plus loin, on représente le processus en deux temps réalisant un tirage selon la loi R comme dans la fig.1.

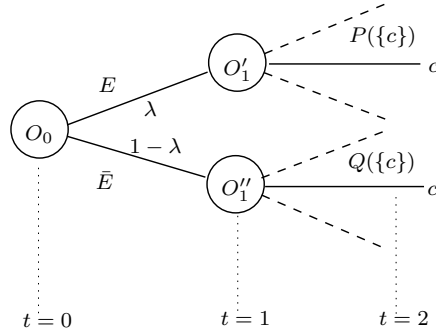


FIG. 1 – tirage aléatoire en deux temps d'un mélange

2.2.1 Propriétés des mélanges

Les principales propriétés des mélanges deviennent évidentes si l'on se souvient de ce que $R = \lambda P + (1 - \lambda)Q$ n'est qu'une écriture symbolique pour $\forall G \in \mathcal{G}, R(G) = \lambda P(G) + (1 - \lambda)Q(G)$, formule qui ne contient que des nombres réels, auxquels s'appliquent donc les règles de l'algèbre ; il en résulte en particulier les propriétés suivantes :

Proposition 2. $\forall P, Q$, lois de probabilités et $\forall \alpha, \lambda \in [0, 1]$,

- (i) $1P + 0Q = P$; $0P + 1Q = Q$;
- (ii) $\alpha P + (1 - \alpha)Q = (1 - \alpha)Q + \alpha P$;
- (iii) $\alpha[\lambda P + (1 - \lambda)Q] + (1 - \alpha)Q = \alpha\lambda P + [1 - \alpha\lambda]Q$.

2.3 Axiomes de la théorie de l'utilité linéaire

L'axiomatique que nous présentons ici est due à N.E.JENSEN (1967).

Soit un décideur en situation de risque, c.-à-d. associant à toute décision ou stratégie de décision une loi de probabilité (dite "engendrée" par elle) sur un espace de résultats $(\mathcal{C}, \mathcal{G})$, où la σ -algèbre \mathcal{G} contient les singletons. On note \mathcal{P} l'ensemble de toutes les lois de probabilité sur $(\mathcal{C}, \mathcal{G})$.

Axiome 1. *Existence de préférences complètes sur les lois*

Le décideur préfère une décision d à une décision d' si et seulement si les lois de probabilité P et P' qu'elles engendrent respectivement satisfont $P \succsim P'$, où la relation \succsim est un préordre total sur \mathcal{P} .

Cet axiome exige implicitement que seules les probabilités des événements déterminant les résultats comptent, mais pas les événements en eux-mêmes: si vous pensez qu'une pièce n'est pas biaisée, donc que $Pr(Pile) = Pr(Face) = \frac{1}{2}$, vous devez être indifférent, à mise et gain donnés, entre parier sur *Pile* et parier sur *Face*.

Vous devez aussi être insensible à la manière, qui peut être plus ou moins progressive, dont est déterminé le gain que vous obtenez finalement : si vous devez miser pour gagner si et seulement *Pile* sort deux fois, vous devez être indifférent entre la procédure où deux pièces sont lancées simultanément ou succesivement si vous pensez que votre probabilité de gagner est la même, $\frac{1}{4}$, dans les deux cas.

Cette hypothèse, connue sous le nom de *réduction des loteries composées* est rejetée par certains modèles car elle exclut certains aspects psychologiques courants, comme le *plaisir à jouer* éprouvé par les joueurs de cartes ou de roulette pendant le déroulement du jeu, qui leur fait préférer un processus lent de détermination du gain final à une résolution instantanée du jeu.

Axiome 2. Indépendance

$\forall P, P', Q \in \mathcal{P}, \forall \lambda \in]0, 1[$,

$$P \succ P' \Rightarrow \lambda P + (1 - \lambda)Q \succ \lambda P' + (1 - \lambda)Q$$

La justification classique de cet axiome fait appel au schéma de tirage aléatoire en deux temps des mélanges $R = \lambda P + (1 - \lambda)Q$ et $R' = \lambda P' + (1 - \lambda)Q$, avec les mêmes événements E et \bar{E} dans les deux cas (cf. fig.2).

L'argument est le suivant : si c'est \bar{E} qui est réalisé à $t = 1$, peu importe que vous ayez choisi à $t = 0$ la décision d engendrant le mélange R ou celle, d' engendrant le mélange R' ; c'est donc seulement les perspectives offertes conditionnellement à la réalisation de E qui peuvent vous faire pencher vers d ou vers d' ; or vous préférez la perspective P à la perspective P' ; vous devez donc choisir d .

Pourtant de nombreux décideurs peuvent ne pas respecter l'axiome d'indépendance dans certaines situations, comme l'a montré, entre autres, l'expérience suivante :

Paradoxe d'ALLAIS

Soit A, B, C, et D quatre loteries.

A: gagner 10 000 € = 10 k€ avec certitude ;

B: gagner 50 k€ avec probabilité $\frac{10}{11}$; ne rien gagner sinon ;

C: gagner 10 k€ avec probabilité $\frac{11}{100}$; ne rien gagner sinon ;

D: gagner 50 k€ avec probabilité $\frac{10}{100}$; ne rien gagner sinon.

L'expérience consiste à offrir aux sujets le choix entre A et B d'une part, C et D d'autre part.

Les perspectives aléatoires offertes par les diverses propositions sont les suivantes :

$$A \mapsto P = \delta_{10} ;$$

$$B \mapsto P' = \frac{10}{11}\delta_{50} + \frac{1}{11}\delta_0 = \frac{10}{11}\delta_{50} + \frac{1}{11}Q ; \text{ (on pose } Q = \delta_0 \text{)}$$

$$C \mapsto R = \frac{11}{100}\delta_{10} + \frac{89}{100}\delta_0 = \frac{11}{100}P + \frac{89}{100}Q ;$$

$$D \mapsto R' = \frac{10}{100}\delta_{50} + \frac{90}{100}\delta_0 = \frac{11}{100}[\frac{10}{11}\delta_{50} + \frac{1}{11}\delta_0] + \frac{89}{100}\delta_0 = \frac{11}{100}P' + \frac{89}{100}Q.$$

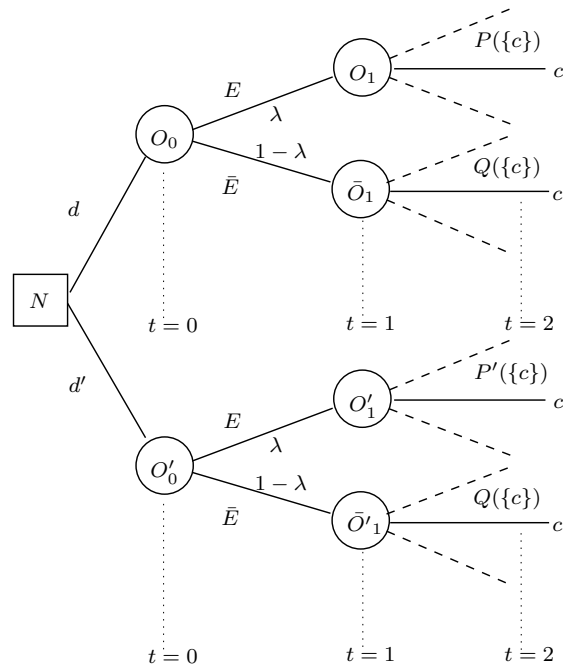


FIG. 2 – Axiome d'indépendance et résolution en deux temps

L'axiome d'indépendance exige que :

$$P \succ P' \Rightarrow \frac{11}{100}P + \frac{89}{100}Q \succ \frac{11}{100}P' + \frac{89}{100}Q$$

donc que C soit choisi si A l'est, mais que D soit choisi si c'est B qui l'est.

La moitié environ des sujets font les choix A et D et ne respectent donc pas l'axiome d'indépendance. Ils justifient souvent ces choix par un *effet de certitude* (KAHNEMAN et TVERSKY) : ils sont sensibles au fait que le gain minimum assuré par A est supérieur à celui de B , avantage que ne possède pas C par rapport à D , puisqu'avec chacun des deux il est possible ne rien gagner. On peut dire aussi que le critère MAXIMIN combine chez eux ses effets à ceux du critère EU.

On peut soutenir que si les décideurs ne respectent pas spontanément l'axiome d'indépendance, ils se montrent en cela déraisonnables. La défense moderne de l'axiome d'indépendance consiste à montrer que toute personne violant parfois cet axiome est susceptible de rencontrer des situations de choix où elle aura un comportement irrationnel, en gaspillant de l'argent ; on dit qu'elle peut être victime d'une *pompe monétaire*.

Pompe monétaire

L'idée est la suivante. Supposons qu'il existe $P, P', Q \in \mathcal{P}$ et $\lambda \in [0, 1]$, tels que $P \succ P'$ et que pourtant $R' = \lambda P' + (1 - \lambda)Q \succ R = \lambda P + (1 - \lambda)Q$ (violation forte de l'axiome; $R' \sim R$ en constituerait une violation faible). Considérons alors la situation où à $t = 0$ le décideur sait qu'à $t = 1$, si E est réalisé, il aura à choisir entre deux décisions, d et d' , lui offrant respectivement les perspectives P et P' , alors que, si c'est \bar{E} qui se produit, il se trouvera (sans avoir aucun choix) devant la perspective Q . A $t = 0$, il pense que choisir d' à $t = 1$ est préférable, car cela le mettrait alors devant la perspective R' qu'il préfère à R ; mais il prévoit qu'à $t = 1$ il choisira en fait d , puisqu'il préfère P à P' . Etant donné cette anticipation, il accepterait de payer une somme d'argent M pour se voir offrir directement, à $t = 0$, la perspective R' . Or cette perspective lui est déjà offerte par la situation initiale; ce n'est que son incohérence dynamique qui l'empêche d'en profiter et l'amène à gaspiller M sans nécessité (cf. fig.3).

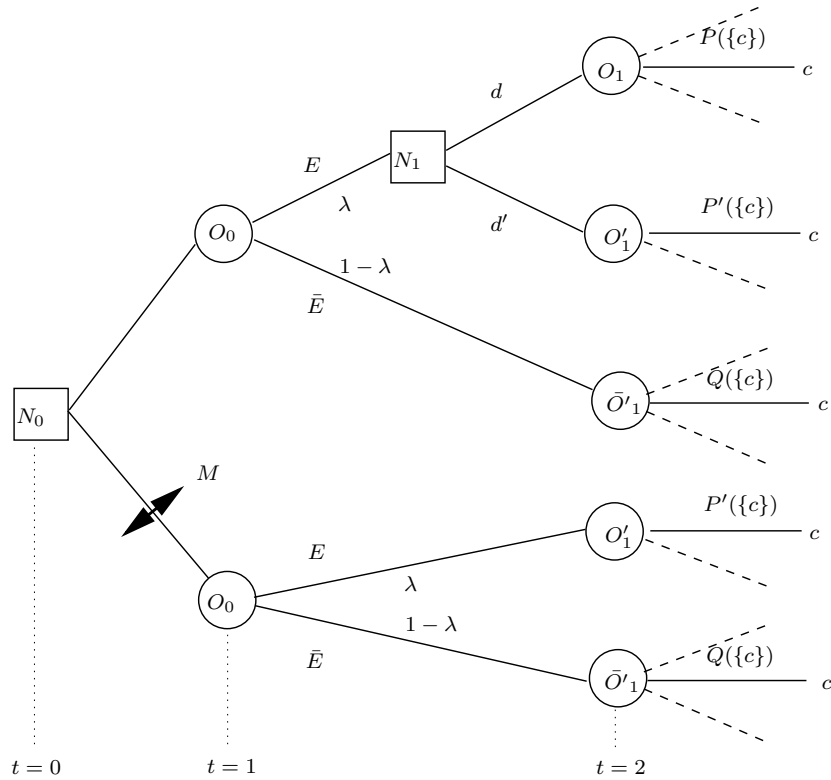


FIG. 3 – Pompe monétaire

L'argument précédent suppose que le décideur est *sophistiqué*, c.-à-d. capable d'anticiper sur ses actions futures; il admet de plus que ses choix en

tout sommet de décision sont guidés par la seule considération des perspectives offertes par les actions dans le sous-arbre issu de ce sommet (*consequentialisme*) et que ses préférences parmi les lois sont les mêmes partout (*invariance*). Rien de ceci ne va de soi et la discussion reste ouverte!

Axiome 3. Continuité

$\forall P, P', Q \in \mathcal{P}$ tels que $P \succ Q \succ P'$, $\exists \lambda, \mu \in]0, 1[$ tels que $\lambda P + (1 - \lambda)P' \succ Q \succ \mu P + (1 - \mu)P'$.

N.B. *Nous utilisons ci-dessous la notation: $R_\alpha = \alpha P + (1 - \alpha)P'$.*

Il faut comprendre que c'est pour λ proche de 1 et pour β proche de 0 que ceci doit être réalisé ; comme R_λ tend vers P en loi lorsque λ tend vers 1, et que de même R_μ tend vers P' lorsque μ tend vers 0, il s'agit bien d'une propriété de continuité des préférences.

On peut concevoir que cet axiome ne soit pas respecté dans certains cas. Par exemple supposons qu'il existe un résultat c_0 extrêmement déplaisant : ruine, souffrance, mort violente ou prématurée, etc. ; supposons que c_0 n'appartient pas au support de P ni à celui de Q et prenons $P' = \delta_{c_0}$; pour tout $\lambda \in]0, 1[$, le support de R_λ contiendra donc c_0 . Si le décideur veut avant tout éviter de courir le risque d'obtenir ce résultat, même avec une probabilité très faible, ce qui est possible avec Q mais pas avec R_λ , il préférera toujours le premier au second.

Les défenseurs de l'axiome avancent que dans le monde réel, aucune décision ne permet d'être sûr d'éviter c_0 et qu'on peut au plus chercher à réduire sa probabilité. C'est sans doute vrai si $c_0 = \textit{mort accidentelle}$ mais pas s'il s'agit, par exemple, de $c_0 = \textit{mort par noyade}$: on peut éviter de s'approcher de l'eau!

Un autre argument en faveur de l'axiome de continuité est que si la probabilité $(1 - \lambda)$ de c_0 est très très faible, de l'ordre de 10^{-6} ou 10^{-9} , la possibilité de ce résultat ne doit plus avoir aucun impact sur les préférences parce que l'on est persuadé qu'un événement ayant une telle probabilité ne se produira pas. Cet argument est fallacieux, car cela veut alors dire que cet événement a une probabilité subjective nulle et que, psychologiquement on a remplacé λ par 1, ce qui ne prouve donc plus rien.

Il existe des modèles n'exigeant que des versions affaiblies de l'axiome de continuité.

Nous ajouterons plus loin un quatrième axiome ; les trois premiers nous suffisent pour démontrer le théorème de représentation qui suit.

2.4 Le théorème de l'utilité linéaire et son corollaire

Définition 3. On appelle *utilité linéaire* sur l'ensemble \mathcal{P} , muni de l'opération de mélange et de la relation de préférence \succsim , une application $U : \mathcal{P} \rightarrow \mathbb{R}$ qui :

i) est une fonction d'utilité représentant \succsim , c.-à-d. satisfait

$$\forall P, Q \in \mathcal{P}, P \succsim Q \Leftrightarrow U(P) \geq U(Q);$$

et

ii) est linéaire pour les mélanges, c.-à-d. telle que

$$\forall P, Q \in \mathcal{P}, \lambda \in [0, 1], U(\lambda P + (1 - \lambda)Q) = \lambda U(P) + (1 - \lambda)U(Q).$$

Remarque le terme *linéaire* est la transcription de l'anglais "linear" ; il s'agit en fait d'une fonction *affine*.

Le lemme suivant nous sera utile :

Proposition 3. Si U est une fonction linéaire sur \mathcal{P} , alors pour tout mélange d'ordre n , $R = \sum_{i=1}^n \lambda_i P_i$ de n lois $(P_i, i = 1, \dots, n)$ avec poids avec poids $(\lambda_i, i = 1, \dots, n), \lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1$

$$U(R) = \sum_{i=1}^n \lambda_i U(P_i).$$

Démonstration Par récurrence. Si la propriété est vraie jusqu'à $n - 1$, de $R = \sum_{i=1}^n \lambda_i P_i = (1 - \lambda_n) [\sum_{i=1}^{n-1} \frac{\lambda_i}{1 - \lambda_n} P_i] + \lambda_n P_n$ on déduit que

$$\begin{aligned} U(R) &= (1 - \lambda_n) U(\sum_{i=1}^{n-1} \frac{\lambda_i}{1 - \lambda_n} P_i) + \lambda_n U(P_n) = \\ &= (1 - \lambda_n) \sum_{i=1}^{n-1} \frac{\lambda_i}{1 - \lambda_n} U(P_i) + \lambda_n U(P_n) = \sum_{i=1}^n \lambda_i U(P_i). \end{aligned}$$

Le résultat principal de la théorie s'énonce ainsi :

Proposition 4. Sous les axiomes 1, 2 et 3, il existe une fonction d'utilité linéaire sur \mathcal{P} . Elle est unique à une transformation affine strictement croissante près.

Avant de donner la démonstration de ce théorème, nous allons en donner un corollaire qui montre le lien existant entre utilité linéaire et utilité espérée.

Proposition 5. Soit U une fonction d'utilité linéaire U sur \mathcal{P} . Soit u l'application de \mathcal{C} dans \mathbb{R} définie par $\forall c \in \mathcal{C}, u(c) = U(\delta_c)$ où δ_c est la loi certaine en c .

Alors, l'utilité linéaire d'une loterie P , de support $\mathcal{S} = \{c_i, i \in I\}$, I fini, est l'espérance mathématique de la fonction u par rapport à la loi P ,

$$U(P) = \sum_{i=1}^n P(\{c_i\}) u(c_i)$$

Démonstration

D'après deux résultats précédents, P s'écrit

$$P = \sum_{i=1}^n P(\{c_i\})\delta_{c_i}$$

et par linéarité de U par rapport aux mélanges,

$$U(P) = \sum_{i=1}^n P(\{c_i\})U(\delta_{c_i}).$$

Introduisant alors u , application de \mathcal{C} dans \mathbb{R} définie par

$$\forall c \in \mathcal{C}, u(c) = U(\delta_c),$$

nous obtenons pour $U(P)$ l'expression annoncée.

La fonction u est appelée *utilité de von NEUMANN-MORGENSTERN* (vNM) du décideur. Comme U peut être remplacée par $V = a.U + b$, ($a > 0$), $v = a.u + b$ est aussi une utilité de vNM ; nous verrons que u , comme U , est unique modulo une telle transformation.

On peut définir les préférences dans le certain du décideur, $\succsim_{\mathcal{C}}$, à partir de ses préférences dans le "presque certain" par

$$c \succsim_{\mathcal{C}} c' \Leftrightarrow \delta_c \succsim \delta_{c'}.$$

On peut alors dire que u est une utilité représentant les préférences dans le certain du décideur.

Remarque Toute transformée strictement croissante de u est encore une utilité représentant $\succsim_{\mathcal{C}}$, mais seules celles qui en sont des transformées affines sont encore des utilités de vNM pour le décideur.

2.4.1 Démonstration du théorème de l'utilité linéaire

Cette démonstration s'articule en plusieurs points.

1. Comparaisons de mélanges de lois

$$\text{i) } \forall P, P' \in \mathcal{P}, \forall \lambda \in]0, 1], P \succ P' \Rightarrow P \succ \lambda P' + (1 - \lambda)P.$$

Preuve : Il suffit de remarquer que $P = \lambda P + (1 - \lambda)P$ et d'appliquer l'axiome 2 (ax. d'indépendance).

ii) $\forall P, P' \in \mathcal{P}, \forall \lambda, \mu \in [0, 1],$
 $[P \succ P' \text{ ET } \lambda > \mu] \Rightarrow \lambda P + (1 - \lambda)P' \succ \mu P + (1 - \mu)P'.$

Preuve : Soit $\alpha = \frac{\mu}{\lambda} \in [0, 1[$; d'après (i), $P \succ (1 - \alpha)P' + \alpha P = \alpha P + (1 - \alpha)P'$. L'axiome 2 appliqué à $P, \alpha P + (1 - \alpha)P'$ et P' entraîne alors, puisque $\lambda > 0$, que : $\lambda P + (1 - \lambda)P' \succ \lambda[\alpha P + (1 - \alpha)P'] + (1 - \lambda)P' = \lambda\alpha P + (1 - \alpha\lambda)P' = \mu P + (1 - \mu)P'.$

Ces deux propriétés se comprennent facilement si l'on pense à une résolution en deux temps des mélanges : le décideur préfère avoir une probabilité plus élevée d'être à la deuxième étape devant la perspective la plus favorable ; de plus, à probabilité donnée d'avoir une perspective, plus celle-ci est intéressante, plus il est satisfait ; etc.

2. Existence d'un mélange indifférent à une loi donnée

$\forall P, P', Q \in \mathcal{P}$ t.q. $P \succ Q \succ P', \exists \hat{\lambda} \in]0, 1[$ t.q. $Q \sim \hat{\lambda}P + (1 - \hat{\lambda})P'$;
de plus, $\hat{\lambda}$ est unique.

Preuve : Soit $E = \{\lambda \in]0, 1[: \lambda P + (1 - \lambda)P' \succ Q\}$; $E \neq \emptyset$, d'après l'axiome 3 (ax. de continuité) ; d'après (1.ii), si $\mu \in E$ et que $\lambda > \mu$, alors $\lambda \in E$ également ; E est donc un intervalle de la forme $]\hat{\lambda}, 1[$ ou $[\hat{\lambda}, 1[$; toujours d'après l'axiome 3, il est impossible que $E =]0, 1[$ et donc $\hat{\lambda} > 0$.

Comparons $\hat{\lambda}P + (1 - \hat{\lambda})P'$ et Q .

Si $\hat{\lambda}P + (1 - \hat{\lambda})P' \succ Q \succ P'$, par l'axiome 3, $\exists \alpha \in]0, 1[$ t.q.
 $\alpha[\hat{\lambda}P + (1 - \hat{\lambda})P'] + (1 - \alpha)P' \succ Q$, soit encore $\alpha\hat{\lambda}P + (1 - \alpha\hat{\lambda})P' \succ Q$
ce qui est impossible, puisque $\alpha\hat{\lambda} < \hat{\lambda}$ et donc $\alpha\hat{\lambda} \notin E$.

Si $P \succ Q \succ \hat{\lambda}P + (1 - \hat{\lambda})P'$, toujours par l'axiome 3, $\exists \beta \in]0, 1[$ t.q.
 $Q \succ \beta P + (1 - \beta)[\hat{\lambda}P + (1 - \hat{\lambda})P'] = [\beta + (1 - \beta)\hat{\lambda}]P + (1 - \beta)(1 - \hat{\lambda})P'$, ce
qui est impossible puisque $\beta + (1 - \beta)\hat{\lambda} > \hat{\lambda}$.

Nécessairement donc $\hat{\lambda}P + (1 - \hat{\lambda})P' \sim Q$; l'unicité de $\hat{\lambda}$ est alors une conséquence immédiate de (1.ii).

3. Comparaisons de mélanges (suite)

i) $\forall P, P' \in \mathcal{P}, \forall \lambda \in [0, 1], P \sim P' \Rightarrow \lambda P + (1 - \lambda)P' \sim P.$

Preuve : Par l'absurde. Supposons que $P \sim P'$ et $\lambda P + (1 - \lambda)P' \succ P$;
d'où $\lambda P + (1 - \lambda)P' \succ P'$ également. En employant deux fois l'axiome 2, il vient

$\lambda[\lambda P + (1 - \lambda)P'] + (1 - \lambda)[\lambda P + (1 - \lambda)P'] \succ \lambda P + (1 - \lambda)P'$, soit encore $\lambda P + (1 - \lambda)P' \succ \lambda P + (1 - \lambda)P'$; contradiction. Raisonnement analogue si $P \succ \lambda P + (1 - \lambda)P'$.

ii) $\forall P, P', Q \in \mathcal{P}, \forall \lambda \in [0, 1], P \sim P' \Rightarrow \lambda P + (1 - \lambda)Q \sim \lambda P' + (1 - \lambda)Q$.

Preuve : Le cas $\lambda = 0$ est trivial. Nous supposons donc que $\lambda > 0$.

Si $Q \sim P \sim P'$, on applique (i) deux fois.

Sinon, supposons par exemple que $Q \succ P$ et donc que $Q \succ P'$.

Par (1.ii), $Q \succ \lambda P + (1 - \lambda)Q$ et $Q \succ \lambda P' + (1 - \lambda)Q$ également. Raisonnant par l'absurde, supposons que $Q \succ \lambda P + (1 - \lambda)Q \succ \lambda P' + (1 - \lambda)Q$.

Par (2) ci-dessus, il existe un unique $\mu \in]0, 1[$ tel que $\lambda P + (1 - \lambda)Q \sim \mu Q + (1 - \mu)[\lambda P' + (1 - \lambda)Q] = (1 - \mu)\lambda P' + [1 - (1 - \mu)\lambda]Q$.

Mais, par (1.i), $Q \succ P' \Rightarrow (1 - \mu)P' + \mu Q \succ P' \sim P$; d'où par (1.ii),

$\lambda[(1 - \mu)P' + \mu Q] + (1 - \lambda)Q \succ \lambda P + (1 - \lambda)Q$; d'où une contradiction, puisque $\lambda[(1 - \mu)P' + \mu Q] + (1 - \lambda)Q = (1 - \mu)\lambda P' + [1 - (1 - \mu)\lambda]Q$.

Raisonnement analogue dans les autres cas.

4. Existence d'une utilité linéaire sur un intervalle d'ordre

Appelons *intervalle d'ordre de \mathcal{P}* tout ensemble de lois de la forme $\mathcal{M} = \{P \in \mathcal{P} : P^+ \succsim P \succsim P^-\}$, où $P^+ \succ P^-$.

i) Pour tout $P \in \mathcal{M}$ il existe un nombre λ_P tel que

$\lambda_P P^+ + (1 - \lambda_P)P^- \sim P$: cela résulte de (2) ci-dessus si $P^+ \succ P \succ P^-$;

c'est vrai avec $\lambda_P = 1$ si $P \sim P^+$, parce que $P^+ = 1P^+ + 0P^-$ et avec

$\lambda_P = 0$ si $P \sim P^-$, parce que $P^- = 0P^+ + 1P^-$.

Dans tous les cas il résulte de (1.ii) que λ_P est unique et que

$P \succ Q \Leftrightarrow \lambda_P P^+ + (1 - \lambda_P)P^- \succ \lambda_Q P^+ + (1 - \lambda_Q)P^- \Rightarrow \lambda_P > \lambda_Q$.

L'application $U : \mathcal{M} \rightarrow [0, 1]$ donnée par: $P \mapsto U(P) = \lambda_P$

est donc bien définie et satisfait, d'après (3.ii) et (1.ii) :

$P \sim Q \Rightarrow U(P) = U(Q)$ ainsi que $P \succ Q \Rightarrow U(P) > U(Q)$;

c'est donc une utilité représentant la restriction de \succsim à \mathcal{M} .

ii) Montrons que U est linéaire sur \mathcal{M} .

Soit $P, Q \in \mathcal{M}$, $\alpha \in [0, 1]$ et $R = \alpha P + (1 - \alpha)Q$. Comme

$[P^+ \succsim P \text{ ET } P^+ \succsim Q] \Rightarrow P^+ \succsim \alpha P + (1 - \alpha)Q$, et

$[P \succsim P^- \text{ ET } Q \succsim P^-] \Rightarrow \alpha P + (1 - \alpha)Q \succsim P^-$, on a bien $R \in \mathcal{M}$.

$U(R) = \lambda_R$, t.q. $R \sim \lambda_R P^+ + (1 - \lambda_R)P^-$; mais d'autre part, de

$R = \alpha P + (1 - \alpha)Q$, $P \sim \lambda_P P^+ + (1 - \lambda_P)P^-$ et $Q \sim \lambda_Q P^+ + (1 - \lambda_Q)P^-$,

il vient, en appliquant deux fois (3.ii), $R \sim \alpha[\lambda_P P^+ + (1 - \lambda_P)P^-] +$

$(1 - \alpha)[\lambda_Q P^+ + (1 - \lambda_Q)P^-] = [\alpha\lambda_P + (1 - \alpha)\lambda_Q]P^+ + [1 - \alpha\lambda_P - (1 - \alpha)\lambda_Q]P^-$.

Par unicité, $\lambda_R = \alpha\lambda_P + (1 - \alpha)\lambda_Q$, soit $U(R) = \alpha U(P) + (1 - \alpha)U(Q)$.

5. Cardinalité de l'utilité linéaire sur un intervalle d'ordre

Montrons que l'utilité linéaire U sur \mathcal{M} du (4) est *cardinale*, c-à-d unique à une transformation affine strictement croissante près.

Il est clair que toute fonction $V = aU + b$, $a > 0$, est encore une utilité sur \mathcal{M} ; elle est aussi linéaire, puisque: $V(\alpha P + (1 - \alpha)Q) = aU(\alpha P + (1 - \alpha)Q) + b = a[\alpha U(P) + (1 - \alpha)U(Q)] + b = \alpha[aU(P) + b] + (1 - \alpha)[aU(Q) + b] = \alpha V(P) + (1 - \alpha)V(Q)$.

Réciproquement, si V est une utilité linéaire,

$$\forall P \in \mathcal{M}, V(P) = V(\lambda_P P^+ + (1 - \lambda_P)P^-) = \lambda_P V(P^+) + (1 - \lambda_P)V(P^-) = [V(P^+) - V(P^-)]\lambda_P + V(P^-) = [V(P^+) - V(P^-)]U(P) + V(P^-).$$

V est bien de la forme cherchée avec

$$a = V(P^+) - V(P^-) \quad (a > 0, \text{ car } P^+ \succ P^-) \text{ et } b = V(P^-).$$

Notons que l'on peut toujours imposer des valeurs arbitraires aux utilités de deux lois: par exemple, $V(P') = c'$ et $V(P'') = c''$, avec $c' > c''$ si $P' \succ P''$; il existe une et une seule utilité linéaire satisfaisant ces deux conditions: $V = aU + b$ avec

$$a = \frac{V(P') - V(P'')}{U(P') - U(P'')}; \quad b = \frac{V(P')U(P'') - V(P'')U(P')}{U(P') - U(P'')}.$$

6. Extension de l'utilité linéaire à tout \mathcal{P}

Si \succsim est trivial, c-à-d si toutes les lois sont indifférentes les unes aux autres, $U = cte$ est une utilité linéaire.

Sinon, il existe P^0, P^1 t.q. $P^1 \succ P^0$. Soit $P \in \mathcal{P}$. et soit \mathcal{M} un intervalle d'ordre contenant à la fois P, P^0 et P^1 ; il en existe; par exemple, on peut prendre $\mathcal{M} = \{Q \in \mathcal{P} : \max_{\succsim}\{P, P^1\} \succsim Q \succsim \min_{\succsim}\{P, P^0\}\}$.

D'après (4) et (5), il existe sur \mathcal{M} une utilité linéaire, V , satisfaisant $V(P^0) = 0$ et $V(P^1) = 1$. Posons $U(P) = V(P)$.

Vérifions que $U(P)$ ne dépend pas de l'intervalle d'ordre choisi. Soit \mathcal{M}' un autre intervalle d'ordre contenant P, P^0 et P^1 et V' une utilité linéaire sur \mathcal{M}' satisfaisant $V'(P^0) = 0$ et $V'(P^1) = 1$. L'ensemble $\mathcal{M} \cap \mathcal{M}'$ est lui-même un intervalle d'ordre contenant P, P^0 et P^1 ; les restrictions de V et V' à cet ensemble sont des utilités linéaires et, coïncidant en P^0 et en P^1 , y sont partout égales; en particulier: $V(P) = V'(P)$.

L'application $P \mapsto U(P)$ est donc bien définie en tout $P \in \mathcal{P}$.

Pour vérifier que c'est bien une utilité linéaire sur \mathcal{P} , c-à-d que $P \succsim P' \Leftrightarrow U(P) \geq U(P')$ et $U(\lambda P + (1 - \lambda)Q) = \lambda U(P) + (1 - \lambda)U(Q)$, il suffit de remarquer qu'il existe sur l'intervalle d'ordre $\mathcal{M} = \{R \in \mathcal{P} : \max_{\succsim} \{P, Q, P^1\} \succsim R \succsim \min_{\succsim} \{P', Q, P^0\}\}$ une utilité linéaire qui y est partout égale à U .

2.5 De l'utilité linéaire à l'utilité espérée

Nous avons vu, avec le corollaire du théorème de l'utilité linéaire, que l'utilité d'une loterie, $U(P)$ pouvait s'exprimer comme l'espérance mathématique de l'utilité de vNM des résultats, u :

$$U(P) = \sum_{i=1}^n P(\{c_i\})u(c_i).$$

Pour que ce résultat s'étende à des lois de probabilités quelconques, il est en particulier nécessaire que pour une loi a.c. de densité $p(c)$ l'intégrale $\int_{\mathcal{P}} u(c)p(c)d\mu$ existe, ce qui exige tout d'abord que u soit mesurable; de plus, si u n'est pas bornée, cette intégrale divergera pour certaines lois. Nous allons donc devoir ajouter un nouvel axiome qui assurera que u est bornée.

Axiome 4. Mesurabilité et Dominance

i) $\forall c \in \mathcal{C}, \{c' \in \mathcal{C} : c' \succsim_{\mathcal{C}} c\}, \{c' \in \mathcal{C} : c \succsim_{\mathcal{C}} c'\} \in \mathcal{G}$;

ii) $\forall P \in \mathcal{P}, \forall c \in \mathcal{C}$,

$$P(\{c' \in \mathcal{C} : c' \succsim_{\mathcal{C}} c\}) = 1 \Rightarrow P \succsim \delta_c; P(\{c' \in \mathcal{C} : c \succsim_{\mathcal{C}} c'\}) = 1 \Rightarrow \delta_c \succsim P.$$

N.B. Puisque, $c \succsim_{\mathcal{C}} c' \Leftrightarrow \delta_c \succsim \delta_{c'}$, cet axiome pourrait être réécrit en ne faisant intervenir que \succsim .

Démonstration

i) *Mesurabilité de u .*

Montrons d'abord qu'il existe un ensemble dénombrable $S \subseteq u(\mathcal{C})$ tel que $[r, r' \in u(\mathcal{C}) \text{ ET } r > r'] \Rightarrow \exists s \in S, r \geq s \geq r'$.

Formons S en sélectionnant un élément s dans tout ensemble $]q, q'[\cap u(\mathcal{C}) \neq \emptyset$ où $q, q' \in \mathbb{Q}$ et $q' > q$ et en y adjoignant soit r soit r' lorsque $r, r' \in u(\mathcal{C}), r' > r$ et $]r, r'[\cap u(\mathcal{C}) = \emptyset$. Comme deux intervalles de ce type sont nécessairements disjoints, il y en a au plus une infinité dénombrable; comme par ailleurs $\mathbb{Q} \times \mathbb{Q}$ est aussi dénombrable, S l'est également.

S a bien la propriété annoncée, puisque, si $r, r' \in u(\mathcal{C})$ et $r > r'$,

- soit $]r, r'[\cap u(\mathcal{C}) = \emptyset$ et $s = r$ (ou $s = r'$) vérifie $r \geq s \geq r'$;

- soit $]r, r'[\cap u(\mathcal{C}) \neq \emptyset$ et il existe $\bar{r} \in u(\mathcal{C})$ et $q, q' \in \mathbb{Q}$ t.q. $r > q > \bar{r} > q' > r'$;

il existe donc aussi $s \in S$ tel que $r > q > s > q' > r'$ et a fortiori $r \geq s \geq r'$.

Soit $r \in \mathbb{R}$. Si, $r \in u(\mathcal{C})$, $r = u(c)$ pour un certain c ; donc $\{c' \in \mathcal{C} : u(c') \geq r\} = \{c' \in \mathcal{C} : c' \succ_{\mathcal{C}} c\} \in \mathcal{G}$ et $\{c' \in \mathcal{C} : r \geq u(c')\} = \{c' \in \mathcal{C} : c \succ_{\mathcal{C}} c'\} \in \mathcal{G}$ également. Si $r \notin u(\mathcal{C})$, $\{c' \in \mathcal{C} : u(c') \geq r\}$ est égal à $\bigcap_{s \in S, s < r} \{c' \in \mathcal{C} : u(c') > s\}$; comme $s \in S \Rightarrow s = u(c)$ pour un certain c , tous ces ensembles appartiennent à \mathcal{G} et, \mathcal{G} étant une σ -algèbre, leurs unions et intersections dénombrables aussi. On démontre de même que $\{c' \in \mathcal{C} : r \geq u(c')\} \in \mathcal{G}$.

ii) u est bornée .

Nous allons raisonner par l'absurde et utiliser une situation analogue à celle du jeu de SAINT-PETERSBOURG pour mettre en évidence une contradiction. Si u n'est pas bornée supérieurement, on peut trouver pour tout entier $n \geq 1$ un résultat c_n tel que $u(c_n) \geq \max\{2^n, u(c_{n-1})\}$ (d'où $c_n \succ_{\mathcal{C}} c_{n-1}$).

Introduisons les lois discrètes suivantes :

R , à support dénombrable, telle que $R(\{c_n\}) = \frac{1}{2^n}, \forall n \geq 1$;

pour chaque entier $m \geq 1$,

P_m , à support fini, telle que $P_m(\{c_n\}) = \frac{2^{m-1}}{2^{m-1}-1} \times \frac{1}{2^n}, \forall n \leq m-1$;

Q_m , à support dénombrable, telle que $Q_m(\{c_n\}) = 2^{m-1} \times \frac{1}{2^n}, \forall n \geq m$.

Pour tout m , R est un mélange de P_m et Q_m :

$$R = \frac{2^{m-1}-1}{2^{m-1}} P_m + \frac{1}{2^{m-1}} Q_m$$

Par l'axiome de dominance (ax. 4), $Q_m \succ \delta_{c_m}$; d'où par indépendance (ax. 2),

$$\forall m \geq 1, R \succ \frac{2^{m-1}-1}{2^{m-1}} P_m + \frac{1}{2^{m-1}} \delta_{c_m}$$

et donc

$$U(R) \geq \frac{2^{m-1}-1}{2^{m-1}} U(P_m) + \frac{1}{2^{m-1}} U(\delta_{c_m}), \text{ soit encore}$$

$$\forall m \geq 1, U(R) \geq [\frac{1}{2} u(c_1) + \frac{1}{2^2} u(c_2) + \dots + \frac{1}{2^{m-1}} u(c_{m-1})] + \frac{1}{2^{m-1}} u(c_m),$$

$$\text{d'où } \forall m \geq 1, U(R) \geq (m-1) + 2 = m+1,$$

ce qui est incompatible avec l'existence de $U(R) \in \mathbb{R}$. Contradiction.

On démontrerait de même que u doit être bornée inférieurement.

Proposition 6. *Sous les axiomes 1, 2, 3 et 4, il existe une fonction d'utilité linéaire, U , sur \mathcal{P} , unique à une transformation affine strictement croissante près.*

Soit u l'utilité de vNM associée, c-à-d l'application de \mathcal{C} dans \mathbb{R} définie par $\forall c \in \mathcal{C}, u(c) = U(\delta_c)$ où δ_c est la loi certaine en c .

L'utilité linéaire, $U(P)$, d'une loi P est l'espérance mathématique de la fonction u par rapport à la loi P ,

$$U(P) = \int_{\mathcal{C}} u(c) dP$$

Démonstration

i) Comme u est bornée, pour tout $G \in \mathcal{G}$, $G \neq \emptyset$, $u_G^- = \inf_{c \in G} u(c)$ et $u_G^+ = \sup_{c \in G} u(c)$ existent.

Nous allons d'abord montrer que, pour toute loi P ,

$$P(G) = 1 \Rightarrow u_G^- \leq U(P) \leq u_G^+$$

Soit donc P telle que $P(G) = 1$. Considérons $u(G) = \{u(c) : c \in G\}$.

Si sa borne inférieure est atteinte, i.e., $u_G^- \in u(G)$, il résulte de l'axiome 4, que $u_G^- \leq U(P)$; de même, si $u_G^+ \in u(G)$, $U(P) \leq u_G^+$.

Le résultat est donc vrai lorsque les deux bornes sont atteintes.

a) Supposons maintenant que seule la borne inférieure de $u(G)$ soit atteinte. $U(P)$ est donc bornée inférieurement par u_G^- . Elle est aussi bornée supérieurement, parce que $U(Q)$ l'est dans $\mathcal{Q} = \{Q \in \mathcal{P} : Q(G) = 1\}$, sinon il existerait une suite $Q_n \in \mathcal{Q}$ telle que $U(Q_n) \geq 2^n$ et pour $Q = \sum 2^{-n} Q_n$, $U(Q)$ ne pourrait exister (démonstration analogue à celle de u bornée dans la proposition précédente). Il existe donc M tel que $u_G^- \leq U(P) \leq M$.

Puisque u_G^+ n'est pas atteinte, les ensembles croissants $G_\epsilon = \{c \in G : u(c) < u_G^+ - \epsilon\}$ sont tels que $\sup_{\epsilon > 0} P(G_\epsilon) = P(G) = 1$.

S'il existe ϵ_0 tel que $P(G_{\epsilon_0}) = 1$, alors, par l'axiome 4, $U(P) \leq u_G^+$.

Dans le cas contraire, $P(G_\epsilon) < 1$ pour tout $\epsilon > 0$. Introduisons, pour chaque $\epsilon > 0$ la décomposition de P en mélange de ses lois conditionnelles aux événements G_ϵ et $H_\epsilon = G \setminus G_\epsilon$, que l'on note P^{G_ϵ} et P^{H_ϵ} .

$P = P(G_\epsilon)P^{G_\epsilon} + P(H_\epsilon)P^{H_\epsilon}$ et donc $U(P) = P(G_\epsilon)U(P^{G_\epsilon}) + P(H_\epsilon)U(P^{H_\epsilon})$. Comme $P^{G_\epsilon}(G_\epsilon) = 1$, il résulte de l'axiome 4 que $U(P^{G_\epsilon}) \leq u_G^+$, d'où, $U(P) \leq u_G^+ + P(H_\epsilon)M$ pour tout $\epsilon > 0$ et donc, puisque $\inf_{\epsilon > 0} P(H_\epsilon) = 0$, $U(P) \leq u_G^+$.

b) Le même résultat s'obtient par un raisonnement symétrique lorsque seule la borne supérieure de $u(G)$ est atteinte.

Reste donc le cas où $u(G) = \{u(c) : c \in G\}$ n'atteint ni sa borne inférieure u_G^- ni sa borne supérieure u_G^+ . Il existe alors nécessairement un résultat c_0 tel que $u_G^- < u(c_0) < u_G^+$.

Partitionnons G en $G_0 = \{c \in G : u(c) < u(c_0)\}$ et $H_0 = G \setminus G_0$ et décomposons P en mélange de ses lois conditionnelles aux événements G_0 et H_0 , notées respectivement P^{G_0} et P^{H_0} . La loi P^{H_0} est concentrée sur H_0 et $u(H_0)$ a pour borne inférieure $u(c_0)$; par l'axiome 4, $U(P^{H_0}) \geq u(c_0)$ et l'on montre comme au (a) que $U(P^{H_0}) \leq u_{H_0}^+ = u_G^+$; d'où l'on déduit $u_G^- \leq U(P^{H_0}) \leq u_G^+$. Par un raisonnement symétrique, on montrerait que $u_G^- \leq U(P^{G_0}) \leq u_G^+$; enfin, P étant un mélange de P^{G_0} et de P^{H_0} et U étant linéaire, $U(P)$ satisfait les mêmes inégalités.

ii) Soit maintenant $P \in \mathcal{P}$, quelconque ; posons $u^- = \inf_{c \in \mathcal{C}} u(c)$ $u^+ = \sup_{c \in \mathcal{C}} u(c)$ et divisons l'intervalle $[u^-, u^+]$ en intervalles $A_{i,n} = \{c \in \mathcal{C} : u^- + (i-1)(u^+ - u^-) < u(c) \leq u^- + i(u^+ - u^-)\}$. P peut s'exprimer comme un mélange des ses conditionnelles aux événements $A_{i,n}$, $P = \sum_i P(A_{i,n})P^{A_{i,n}}$. D'après le (i), $u^- + (i-1)(u^+ - u^-) \leq U(P^{A_{i,n}}) \leq u^- + i(u^+ - u^-)$ et donc, par linéarité de U , $\sum_i [u^- + (i-1)(u^+ - u^-)]P(A_{i,n}) \leq U(P) \leq \sum_i [u^- + i(u^+ - u^-)]P(A_{i,n})$.

Lorsque $n \rightarrow \infty$, les deux bornes tendent vers $\int_{\mathcal{C}} u(c)dP$; d'où le résultat annoncé.

3 Décision séquentielle dans le risque

Lorsqu'une situation de décision dans le risque ne fait intervenir qu'un nombre fini de décisions, dont la possibilité de réalisation est éventuellement conditionnelle à celle d'événements, eux-même en nombre fini, il est comode de lui associer une représentation graphique, appelée *arbre de décision*.

3.1 Arbres de décision

Un *arbre de décision* est en réalité une *arborescence* dont la racine est généralement placée à gauche du graphique et les feuilles à droite ; les arcs sont alors tous dirigés de la gauche vers la droite, mais on omet le plus souvent de représenter leurs flèches (on place aussi parfois la racine en haut, les feuilles en bas avec des arcs allant du haut vers le bas).

L'orientation de l'arbre correspond au déroulement du temps, du passé vers l'avenir.

Il existe trois sortes de sommets : les *sommets de décision* représentés par des *carrés*, les *sommets d'information* représentés par des *cercles* et les *feuilles*, représentées par des points.

Chacun des arcs issu d'un sommet de décision est associé à une décision pouvant être prise en ce sommet.

Chacun des arcs issu d'un sommet d'information est associé à l'un des événements d'une partition de l'événement certain décrivant les informations pouvant être reçues à ce sommet.

La racine de l'arbre sera le plus souvent un sommet de décision (sans que ce soit une obligation).

Nous ferons toujours l'*hypothèse de mémoire parfaite* : si le D. apprend à une date t , à un sommet N , qu'un événement B est réalisé, il saura désormais en fait que, plus précisément, est vrai l'événement $B \cap [\cap_{i \in I} A_i]$, où les $A_i, i \in I$ sont les divers événements dont il a appris la réalisation en parcourant le chemin qui va de la racine de l'arbre à N .

A toute feuille F est associée un résultat c , qui est le résultat obtenu par le D. s'ils prend toutes les décisions indiquées le long du chemin $R \rightarrow F$ joignant la racine R à cette feuille et si, de plus, se réalise l'événement $E_{R \rightarrow F}$ intersection de tous les événements figurant sur ce chemin.

L'intersection de toutes les partitions associées aux divers sommets d'information est elle-même une partition de l'événement certain dont les éléments sont appelés *événements élémentaires*. Les événements $E_{R \rightarrow F}$ sont donc soit des événements élémentaires, soit des sur-événements d'événements élémentaires.

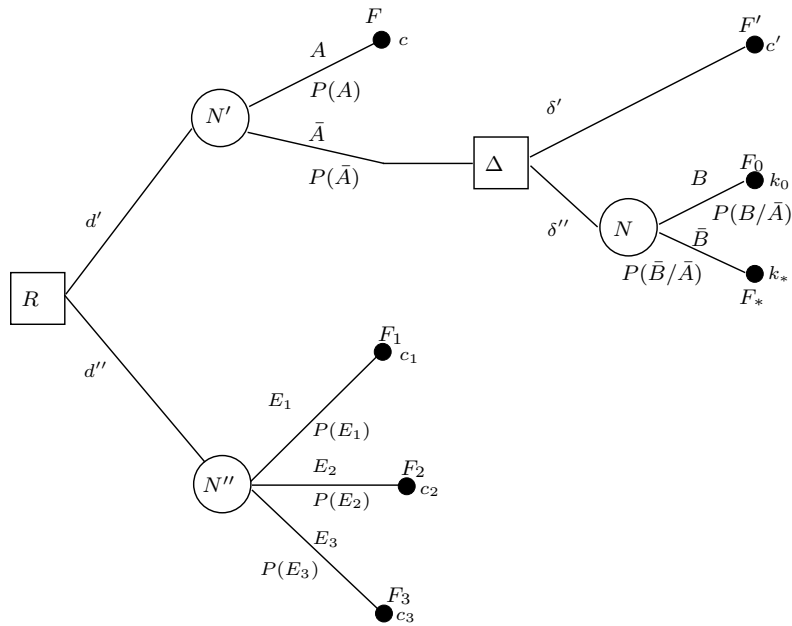


FIG. 4 – Un arbre de décision

Exemple 4. Dans l'arbre de décision de la fig.4, il y a deux sommets de

décision, R (la racine) et Δ , trois sommets d'information, N' , N'' et N , et 7 feuilles.

Les événements élémentaires sont tous ceux des événements de la forme $[A \cup \bar{A}] \cap [B \cup \bar{B}] \cap E_i$ qui sont différents de \emptyset .

En Δ , le D. sait que l'événement \bar{A} est vrai et prévoit donc que s'il prend la décision δ'' , il obtiendra le résultat k_0 avec probabilité $P(B/\bar{A})$ et le résultat k_* avec probabilité $P(\bar{B}/\bar{A})$. En R , s'il a l'intention de prendre immédiatement la décision d' , puis, si \bar{A} est réalisé, la décision δ'' en Δ , il estime avoir une probabilité $P(\bar{A} \cap B) = P(\bar{A}) \times P(B/\bar{A})$ d'obtenir le résultat k_0 ; etc.

3.2 Stratégies et sous-stratégies de décision

3.2.1 Stratégies de décision

Une *stratégie (de décision)* se définit comme la sélection, en tout sommet de décision D de l'arbre, d'une décision d appartenant à l'ensemble \mathcal{D} des décisions réalisables à ce sommet.

N.B. *Comme il peut paraître inutile de préciser les choix que le D. sera sûr de ne pas avoir à faire, on donne parfois une définition différente d'une stratégie n'exigeant la spécification des décisions sélectionnées qu'aux sommets susceptibles d'être atteints, étant donné les décisions antérieures prises par cette stratégie. Il y a de bonnes raisons pour préférer la première définition : elle est plus simple ; elle est compatible avec celle de la Théorie des Jeux ; enfin, les choix en tous les sommets doivent de toute façon être déterminés au cours de la recherche du comportement optimal du D.*

Exemple 5. Dans l'arbre de décision de la fig.4, il y a 4 stratégies : d' en R , δ' en Δ , que nous noterons $[d', \delta']$, ainsi que $[d', \delta'']$, $[d'', \delta']$ et $[d'', \delta'']$. Notons que les deux dernières spécifient ce que le D. devrait faire en Δ , alors qu'il est certain de ne pas y passer ; avec les deux premières, il peut y passer, mais seulement si l'événement \bar{A} est réalisé.

3.2.2 Stratégies de décision

A toute stratégie correspond une loi de probabilité sur l'espace des résultats, \mathcal{C} , dite *loi engendrée par la stratégie*.

Exemple 6. Dans l'exemple précédent, la stratégie $S_1 = [d', \delta']$ engendre sur \mathcal{C} la loi Q_1 caractérisée par : $Q_1(\{c\}) = P(A)$; $Q_1(\{c'\}) = P(\bar{A})$.

La stratégie $S_2 = [d', \delta'']$ engendre la loi Q_2 caractérisée par : $Q_2(\{c\}) = P(A)$; $Q_2(\{k_0\}) = P(\bar{A} \cap B)$; $Q_2(\{k_*\}) = P(\bar{A} \cap \bar{B})$.

Les stratégies $S_3 = [d'', \delta']$ et $S_4 = [d'', \delta'']$ engendrent la même loi $Q_3 (= Q_4)$ caractérisée par : $Q_3(\{c_i\}) = P(E_i)$, $i = 1, 2, 3$.

3.2.3 Sous-arbres et sous-stratégies de décision

On appelle *sous-arbre de décision* le sous-graphe obtenu à partir d'un arbre de décision en ne conservant qu'un sommet (de décision ou d'information) et tous ses descendants.

On appelle *sous-stratégie* de l'arbre de décision toute stratégie d'un de ses sous-arbres.

L'hypothèse de mémoire parfaite entraîne qu'à la racine D d'un sous-arbre les croyances probabilistes sur les événements sont les probabilités conditionnelles à toute l'information reçue le long du chemin $R \rightarrow D$ y menant.

Par exemple, dans le sous-arbre \mathcal{A}_Δ de racine Δ de la fig. 4, les probabilités des événements B et \bar{B} valent respectivement $P(B/\bar{A})$ et $P(\bar{B}/\bar{A})$.

3.3 Utilité linéaire et principe d'optimalité

Nous supposons que le D. se comporte dans le risque conformément au modèle de l'utilité linéaire et indépendamment du contexte ; ceci signifie en particulier qu'en tout sommet de décision ses préférences ne dépendent que des données (probabilités, résultats) dans le sous-arbre dont il est la racine (*conséquentialisme*) et que son utilité de vNM, u , est la même à tous ces sommets (*invariance des préférences*).

Plaçons-nous en un sommet de décision quelconque D (cf. fig. 5) ; soit E l'information recueillie entre la racine et D ; en D la loi de probabilité sur les événements futurs est donc $P(./E)$.

Soit d l'une des décisions possibles en D ; considérons tous les chemins issus de D , commençant par d et se terminant au premier sommet de décision rencontré, Δ , ou, si l'on n'en rencontre pas, par une feuille, F .

Nous noterons $D_i, i \in I$ et $F_j, j \in J$, respectivement les ensembles de sommets de décision et de feuilles rencontrés (I et J peuvent être vides).

Vus de D , les événements de type $H =$ "le chemin $D \rightarrow \Delta$ (ou le chemin $D \rightarrow F$) est réalisé" forment une partition et sont aléatoires de probabilités $P(H/E)$.

Chaque sommet Δ est la racine d'un sous-arbre, \mathcal{A}_Δ , où les événements ont des probabilités conditionnelles $P(./H \cap E)$; ces probabilités déterminent la loi de probabilité Q_σ associée à toute stratégie σ de \mathcal{A}_Δ (qui est aussi une sous-stratégie de \mathcal{A}_D) ; une stratégie optimale σ^* dans \mathcal{A}_Δ est caractérisée par une utilité linéaire $U(Q_{\sigma^*}) \geq U(Q_\sigma), \forall \sigma$.

La stratégie Σ de \mathcal{A}_D commençant par d et se poursuivant, si H_i est réalisé, par une stratégie σ_i dans \mathcal{A}_{Δ_i} ($i \in I$) engendre, en vertu de la

formule des probabilités totales, la loi de probabilité

$$Q = \sum_{i \in I} P(H_i/E) \times Q_{\sigma_i} + \sum_{j \in J} P(H_j/E) \times \delta_{c_j};$$

par linéarité de la fonction d'utilité U ,

$$U(Q) = \sum_{i \in I} P(H_i/E) \times U(Q_{\sigma_i}) + \sum_{j \in J} P(H_j/E) \times u(c_j);$$

de ce que $U(Q_{\sigma_i^*}) \geq U(Q_{\sigma_i}), \forall \sigma_i$, il résulte alors que la stratégie Σ^* de \mathcal{A}_D commençant par d et se poursuivant, si H_i est réalisé, par la stratégie σ_i^* , offre une utilité $U(Q^*) \geq U(Q)$ et est donc optimale parmi toutes celles commençant aussi par d .

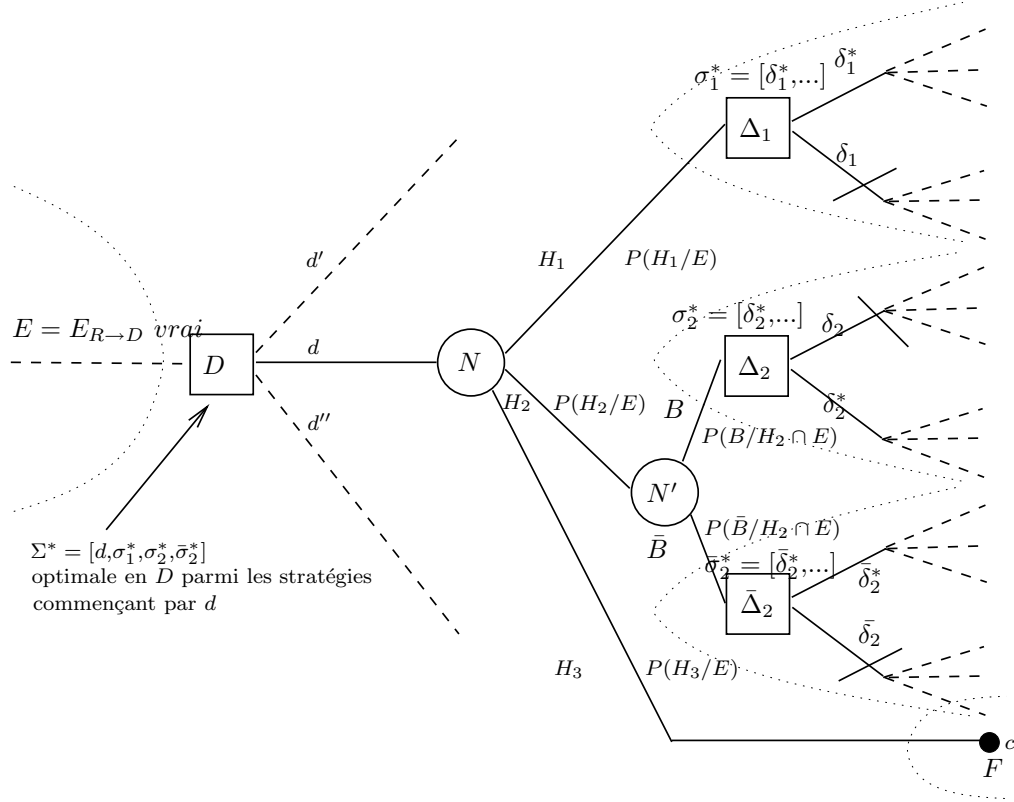


FIG. 5 – principe d'optimalité

Exemple 7. Dans la fig. 5, la décision d en D peut amener en Δ_1 avec probabilité $P(H_1/E)$, en Δ_2 avec probabilité $P(B/E) = P(H_2/E) \times P(B/H_2 \cap E)$, en $\bar{\Delta}_2$ avec probabilité $P(\bar{B}/E) = P(H_2/E) \times P(\bar{B}/H_2 \cap E)$ et enfin en F avec probabilité $P(H_3/E)$.

Une meilleure stratégie en D parmi celles commençant par d consiste à choisir respectivement pour sous-stratégies en Δ_1 , Δ_2 et $\bar{\Delta}_2$ les stratégies σ_1^* , σ_2^* et $\bar{\sigma}_i^*$, optimales dans \mathcal{A}_{Δ_1} , \mathcal{A}_{Δ_2} et $\mathcal{A}_{\bar{\Delta}_2}$.

D'où la validité du *principe d'optimalité* : pour obtenir une stratégie optimale de \mathcal{A}_D il suffit de comparer entre elles les meilleures stratégies commençant par chacune des décisions possibles en D et de prendre celle qui a la plus grande utilité.

Ce n'est autre que la propriété qui sert de fondement à la *programmation dynamique*.

L'application récursive de cette propriété conduit à la procédure suivante :

3.3.1 Détermination d'une stratégie optimale par induction arrière

La procédure opère en remontant des feuilles de l'arbre de décision vers sa racine (*induction arrière*) en évaluant chaque sommet rencontré, c-à-d en déterminant l'utilité maximum réalisable à ce sommet.

Evaluation d'une feuille F

condition : aucune

valeur $U_F = u(c)$, où c est le résultat associé à F

Evaluation d'un sommet d'information N

condition : tous les successeurs directs $S_i, i \in I$ de N sont déjà évalués

valeur : $U_N = \sum_{i \in I} P(H_i/E_{R \rightarrow N}) \times U_{S_i}$ où $E_{R \rightarrow N}$ est l'information connue en N et H_i est l'événement "l'arc $N \rightarrow S_i$ est réalisé"

Evaluation d'un sommet de décision D

condition : tous les successeurs directs $S_i, i \in I$ de D sont déjà évalués

valeur : $U_D = \max_{i \in I} U_{S_i}$

sélection : $i^* = \arg \max_{i \in I} U_{S_i}$; DS_{i^*} est la décision optimale en D

La procédure se termine une fois la racine R évaluée.

4 Attitude par rapport au risque

Nous n'avons pas encore fait d'hypothèse sur la structure de l'ensemble des résultats \mathcal{C} ; nous supposons dorénavant que c'est un *intervalle, borné ou non*, de \mathbb{R} , comme dans les exemples où les résultats sont des sommes d'argent ; cependant certaines des définitions et propositions qui suivent peuvent se généraliser au cas où \mathcal{C} est un sous-ensemble convexe de \mathbb{R}^n . on note toujours \mathcal{P} l'ensemble des lois de probabilité sur \mathcal{C} ; lorsqu'une loi P de \mathcal{P} possède une espérance mathématique, on la désigne par $E(P)$; $E(P) \in \mathcal{C}$ et $\delta_{E(P)}$ est la loi certaine en $E(P)$.

4.1 Attitude globale vis-à-vis du risque

Définition 4. On dit qu'un $D.$ est *adversaire du risque* lorsque, pour toute loi de probabilité de \mathcal{P} possédant une espérance, il préfère être certain d'obtenir le résultat $E(P)$ plutôt que d'obtenir un résultat aléatoire de loi P :

$$\delta_{E(P)} \succsim P ;$$

on dit qu'il est *joueur* s'il manifeste les préférences opposées,

$$\forall P, P \succsim \delta_{E(P)} ;$$

il est dit *neutre vis-à-vis du risque* si

$$\forall P, P \sim \delta_{E(P)} .$$

On peut remarquer que cette définition est *intrinsèque*, en ce sens qu'elle n'est pas liée à un modèle : le $D.$ peut très bien avoir un critère autre qu'une utilité espérée.

Dans le cadre du modèle EU, un $D.$ sera donc :

- adversaire du risque lorsque $\forall P, u(E(P)) \geq U(P)$;
- joueur lorsque $\forall P, u(E(P)) \leq U(P)$; et
- neutre vis-à-vis du risque $\forall P, u(E(P)) = U(P)$.

Il existe alors une caractérisation simple de l'attitude vis-à-vis du risque :

Proposition 7. *Soit un $D.$ ayant pour critère EU avec utilité de vNM u .*
 $D.$ est adversaire du risque $\Leftrightarrow u$ est concave ;
 $D.$ est joueur $\Leftrightarrow u$ est convexe ;
 $D.$ est neutre vis-à-vis du risque $\Leftrightarrow u$ est affine.

Démonstration

(i) u concave $\Leftrightarrow \forall x, y \in \mathcal{C}, u(\frac{1}{2}x + \frac{1}{2}y) \geq \frac{1}{2}u(x) + \frac{1}{2}u(y)$.

Soit donc $x, y \in \mathcal{C}$ et P la loi donnant x et y avec équiprobabilité :

$$P(\{x\}) = P(\{y\}) = \frac{1}{2} ; E(P) = \frac{1}{2}x + \frac{1}{2}y .$$

si D est adversaire du risque, $\delta_{E(P)} \succsim P$ et donc $U(\delta_{E(P)}) \geq U(P)$; or, $U(\delta_{E(P)}) = u(E(P)) = u(\frac{1}{2}x + \frac{1}{2}y)$ et $U(P) = \frac{1}{2}u(x) + \frac{1}{2}u(y)$; u est bien concave (cf la fig.2 plus bas).

(ii) Réciproquement, supposons u concave et soit P une loi possédant une espérance $E(P)$. Par le point $(E(P), u(E(P)))$ de la courbe représentant la fonction $z = u(x)$ il passe une *droite d'appui* (c'est la tangente à la courbe si $u'(E(P))$ existe; n'importe quelle droite de pente comprise entre celle des deux demi-tangentes, qui existent toujours pour u concave, sinon), c-à-d une droite entièrement située au-dessus de cette courbe :

$$\begin{aligned} \forall x \in \mathcal{C}, [z - u(E(P)) = a \cdot [x - E(P)]; u'_-(E(P)) \geq a \geq u'_+(E(P))] \\ \Rightarrow z \geq u(x). \end{aligned}$$

Il en résulte que

$$\forall x \in \mathcal{C}, u(x) - u(E(P)) \leq a \cdot [x - E(P)],$$

d'où en intégrant,

$$\begin{aligned} U(P) - u(E(P)) = \int_{\mathcal{C}} u(x)dP - u(E(P)) \leq \int_{\mathcal{C}} x dP - E(P) = \\ E(P) - E(P) = 0 \end{aligned}$$

et donc $\delta_{E(P)} \succsim P$.

(iii) La preuve de l'équivalence, dans le modèle, entre être joueur et avoir une utilité de vNM convexe est analogue; enfin être neutre vis-à-vis du risque, c'est être à la fois joueur et adversaire du risque, ce qui équivaut à avoir une utilité de vNM à la fois convexe et concave, c-à-d affine.

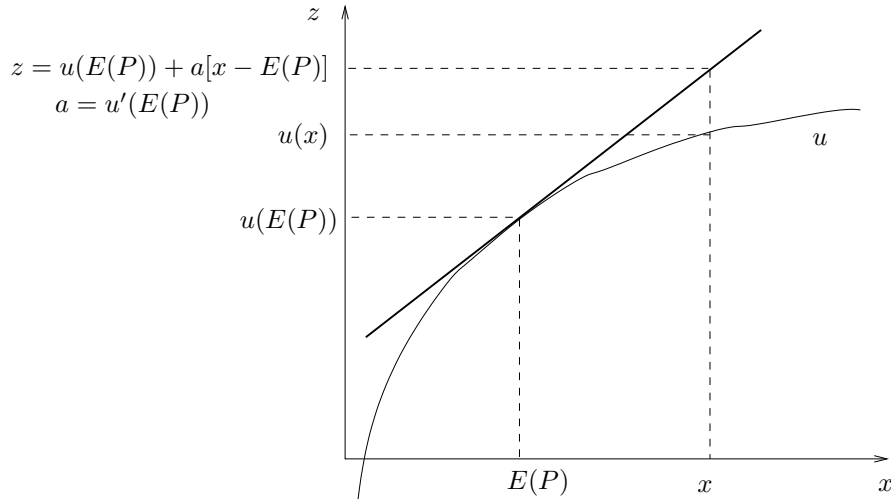


FIG. 6 – Adversaire du risque : utilité de vNM concave

Notation : L'espérance mathématique d'une variable aléatoire X sera notée $E(X)$; lorsque la loi de probabilité qu'elle engendre sur \mathcal{C} sera une loi P on aura donc $E(X) = E(P)$.

D'autre part, on écrira abusivement $U(X)$ pour la valeur de l'utilité linéaire $U(P)$ apportée par sa loi.

On appelle *équivalent-certain* d'une v.a. X , en abrégé $EC(X)$, un résultat tel que le D. soit indifférent entre la perspective aléatoire X et l'obtention de ce résultat avec certitude ; cette définition est intrinsèque ; dans le modèle EU, l'équivalent-certain de X est caractérisé par la relation :

$$u(EC(X)) = U(X) [= U(P)].$$

Il suffit que u soit continue pour que cette équation ait une solution et donc que $EC(X)$ existe ; il sera unique si u est strictement croissante.

On appelle *prime de risque* (associée à la v.a. X) l'écart $\pi(X) = E(X) - EC(X)$; c'est la quantité (algébrique) que le D. est prêt à perdre en moyenne pour éliminer l'incertitude présentée par X .

Dans la partie (ii) de la proposition ci-dessus, nous avons établi la propriété suivante (théorème de JENSEN) ;

$$u \text{ concave} \Leftrightarrow \forall P, u(E(P)) \geq \int u dP$$

Elle entraîne que, dans le cadre du modèle EU, un D. est :

- adversaire du risque $\Leftrightarrow \forall X$, la prime de risque $\pi(X)$ est positive ;
- joueur $\Leftrightarrow \forall X$, la prime de risque $\pi(X)$ est négative ; et
- neutre vis-à-vis du risque $\Leftrightarrow \forall X$, la prime de risque $\pi(X)$ est nulle.

L'existence de compagnies d'assurances est la preuve que les adversaires du risque sont nombreux. Voyons pourquoi.

Supposons pour simplifier que chaque assuré potentiel i n'encourt qu'un seul risque : perdre une somme c_i avec une probabilité p_i ; s'il ne s'assure pas, il est devant une perspective X_i d'espérance $E(X_i) = -p_i c_i (< 0)$ ($-E(X_i) (> 0)$ est la *valeur actuarielle* de ce risque), ayant un équivalent certain $EC(X_i) (< 0)$; s'il est adversaire du risque, $\pi(X_i) = E(X_i) - EC(X_i) > 0$.

S'il s'assure en payant une *prime d'assurance* S_i à l'assureur, il sera devant la perspective non risquée $\delta_{(-S_i)}$; il est donc prêt à s'assurer en payant jusqu'à S_i tel que $u(-S_i) = u(EC(X_i))$, c-à-d $S_i = -EC(X_i) = -E(X_i) + \pi(X_i)$; la *prime d'assurance (maximum)* est la somme de la valeur actuarielle et de la prime de risque.

L'assureur perçoit les primes mais prend en charge tous les aléas et fait donc un profit aléatoire, $\sum_i [X_i + S_i] = \sum_i [X_i - E(X_i) + \pi(X_i)]$, d'espérance, $\sum_i \pi(X_i)$, positive ; si les risques individuels sont indépendants, la loi des grands nombres lui donne une forte probabilité de gagner de l'argent.

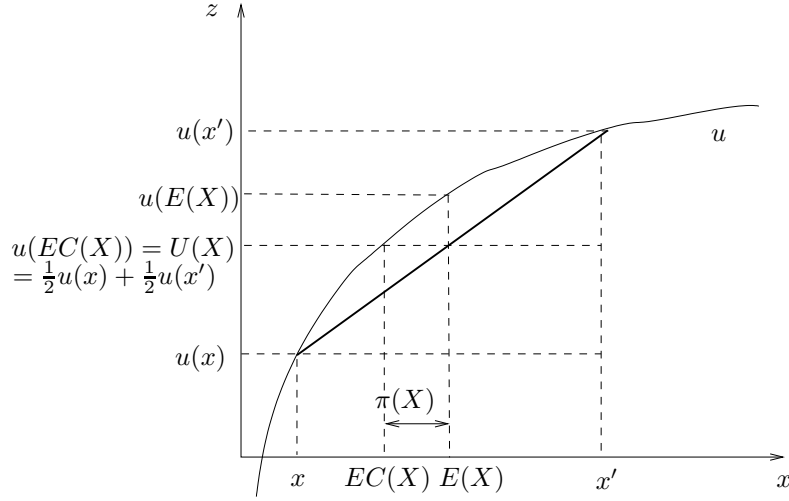


FIG. 7 – Prime de risque d'un adversaire du risque pour $X = \frac{1}{2}\delta_x + \frac{1}{2}\delta_{x'}$

4.2 Attitude locale vis-à-vis du risque

4.2.1 Attitude locale vis-à-vis du risque absolu

Nous allons étudier les comportements face à de “petits risques”, c-à-d à des aléas de faible amplitude par rapport à leur espérance mathématique.

Supposons que le D. a une utilité de vNM, u , de dérivée première u' et seconde u'' . Soit $\{X_h\}$ une famille de v.a. de supports $[m - h, m + h]$, de lois P_h de même espérance $E(X_h) = m$ et de variances $Var(X_h)$ (noter que $Var(X_h) \leq h^2 \rightarrow 0$ quand $h \rightarrow 0$).

Par définition, $U(X_h) = u(EC(X_h))$.

Un développement limité au deuxième ordre de u au voisinage de m s'écrit

$$u(x) = u(m) + (x - m).u'(m) + \frac{1}{2}(x - m)^2.[u''(m) + \epsilon(x - m)],$$

ce qui, en intégrant sur l'intervalle $[m - h, m + h]$, donne

$$U(X_h) = \int_{m-h}^{m+h} u(x)dP_h = u(m) + (m - m)u'(m) + \frac{1}{2}Var(X_h)[u''(m) + \epsilon(h)].$$

Un développement au premier ordre de u au voisinage de m montre par ailleurs que :

$$u(EC(X_h)) = u(m) + (EC(X_h) - m).[u'(m) + \epsilon(EC(X_h) - m)].$$

On en déduit (sous certaines hypothèses techniques sur u'') que

$$\pi(X_h)u'(m) = -(EC(X_h) - m).u'(m) = -\frac{1}{2}Var(X_h).u''(m) + o(h);$$

d'où, pour h infiniment petit,

$$\pi(X_h) \sim -\frac{1}{2} \frac{u''(m)}{u'(m)} Var(X_h).$$

Pour de petits risques X , la prime de risque est donc proportionnelle d'une part à la variance du risque, $Var(X)$, et d'autre part à une caractéristique

locale du comportement du D., son *coefficient (local) d'aversion au risque absolu*, qui en x vaut

$$R_a(x) = -\frac{u''(x)}{u'(x)}.$$

4.2.2 Attitude locale vis-à-vis du risque relatif

On peut penser que le fait qu'un risque doit être considéré comme petit ou non dépend de la richesse du D.

Soit alors $\{Z_h\}$ une famille de v.a. de supports $[1-h, 1+h]$, de lois P_h de même espérance $E(Z_h) = 1$ et de variances $Var(Z_h)$.

On s'intéresse aux v.a. $X_h = m \cdot Z_h$, qui ont donc la même espérance $E(X_h) = m \cdot E(Z_h) = m > 0$ et des variances $Var(X_h) = m^2 \cdot Var(Z_h)$; $Var(Z_h) = \frac{Var(X_h)}{[E(X_h)]^2}$ est la *variance relative* de X_h et $\rightarrow 0$ quand $h \rightarrow 0$; les X_h correspondent donc à des aléas de faible amplitude relative.

Les mêmes calculs que précédemment montrent que, pour h infiniment petit,

$$\pi(X_h) \sim -\frac{1}{2} \frac{u''(m)}{u'(m)} m^2 \cdot Var(Z_h)$$

soit encore :

$$\frac{1}{m} \pi(X_h) \sim -\frac{1}{2} \frac{u''(m)}{u'(m)} m \cdot Var(Z_h).$$

Pour de petits risques relatifs $X (\geq 0)$, la *prime de risque relative* $\frac{1}{E(X)} \pi(X)$ est donc proportionnelle d'une part à la variance relative de ce risque, $\frac{Var(X)}{[E(X)]^2}$ et d'autre part à une caractéristique locale du comportement du D., son *coefficient (local) d'aversion au risque relatif*, qui en x vaut

$$R_r(x) = -x \frac{u''(x)}{u'(x)}.$$

4.2.3 Exemples

$$u(x) = \ln(x) \Rightarrow R_a(x) = \frac{1}{x}, R_r(x) = 1;$$

$$u(x) = k - \frac{1}{x^\gamma} (\gamma > 0) \Rightarrow R_a(x) = \frac{\gamma+1}{x}, R_r(x) = \gamma + 1;$$

$$u(x) = x^\gamma (0 < \gamma < 1) \Rightarrow R_a(x) = \frac{\gamma-1}{x}, R_r(x) = \gamma - 1;$$

$$u(x) = -x^2 + \beta x \Rightarrow R_a(x) = \frac{2}{\beta-2x}, R_r(x) = \frac{2x}{\beta-2x}.$$

Les trois premières fonctions d'utilité correspondent à des coefficients locaux d'aversion au risque absolu décroissants et à des coefficients locaux d'aversion au risque relatif constants, ce qui paraît assez réaliste; ce n'est pas le cas de l'utilité quadratique, où les deux sont croissants.