Notion d'utilité et aversion au risque

Sommaire

1.1 Cadre Général

L'étude du comportement du consommateur se fait en deux étapes. Il faut tout d'abord décrire les préférences des individus, c'est-à-dire comment ils préfèrent tel bien à un autre. Ensuite, le consommateur, ayant des ressources limitées, va rechercher la maximisation de l'utilité sous contrainte budgétaire. La

combinaison des préférences et des contraintes de budget détermine les choix de consommation, et plus précisément quelle combinaison de biens les agents économiques choisiront afin de maximiser leur utilité.
Une grande partie des résultats et définitions de ce chapitre est basée sur le chapitre 2 du livre [34] de Hans Föllmer et Alexander Schied et les travaux de D. Bernoulli [9], John von Neumann et Oskar Morgenstern [110]et J. Leonard Savage [102].

1.2 Les préférences

On définit sur l'ensemble (au sens mathématique) des paniers de consommation la relation de préférence, c'est-à-dire qu'un agent peut exprimer une préférence entre deux paniers de bien.

On suppose que cette relation est :
- **complète** (l'agent est toujours capable de comparer deux paniers de biens).
- **transitive** (si l'agent préfère A à B et B à C, alors il préfère A à C).
- **de comparaison** (si l'agent compare les biens A et B, alors il les considère équivalents).
- **de dominance** (si l'agent préfère plus A à B).
- **de substituabilité** (si l'agent a préféré A par rapport à B à cause de la quantité, par exemple ; alors il est toujours possible de rendre ce dernier indifférent de A en compensant l'insuffisance de B par un surplus de quantité)

De plus, on supposera également qu'un consommateur préfère toujours consommer plus que moins. C'est-à-dire que si on prend un panier puis qu'on augmente la quantité d'un ou plusieurs biens, alors le nouveau panier sera préféré au panier initial (principe de non-satiété).

Cette hypothèse est contestable : on peut en effet penser que le consommateur va "saturer" au bout d'un moment et que la consommation de biens supplémentaires ne lui apporte plus de satisfaction supplémentaire. On va choisir de se placer dans un cadre de long terme (où la saturation est moins probable :

l'agent risque moins de saturer s'il peut répartir sa consommation sur toute une année par exemple). Notons également au passage que la rareté est au coeur de l'analyse économique et que, par conséquent, on s'intéresse plutôt aux situations où les agents sont confrontés à cette rareté et ne peuvent s'offrir tout ce qu'ils désirent.

1.2.1 Représentation numérique d'une relation de préférence

Dans ce paragraphe, on introduit la notion de représentation numérique d'une relation de préférence. Les conditions d'existence d'une telle représentation dans un espace de fonctions $\mathcal{X}$, ont été étudiées et démontrées par J. Leonard Savage [102].
Soit $\mathcal{X}$ un ensemble non vide. Un élément $x \in \mathcal{X}$ est interprété comme étant un choix possible d'un agent économique. Si dans cet ensemble, entre deux éléments x et y, l'agent doit préférer l'un par rapport à l'autre, alors ceci nous amène à la définition suivante

Définition 1.1. *Un ordre de préférence $\succ$ dans $\mathcal{X}$ est une relation binaire, qui a les deux propriétés suivantes :*

- *Asymétrie : si $x \succ y$, alors $y \not\succ x$*
- *transitivité négative :si $x \succ y$ et $z \in \mathcal{X}$, alors soit $x \succ z$ ou $z \succ y$ ou bien les deux en même temps.*

La deuxième assertion s'interprète comme suit : Si l'agent a une nette préférence de x par rapport à y, et si on rajoute un troisième choix alors forcément x reste le plus préféré $x \succ z$ ou bien y est le moins préféré $z \succ y$.

Un ordre de préférence $\succ$ dans $\mathcal{X}$ induit un ordre de préférence faible $\succeq$ défini par

$$x \succeq y \Leftrightarrow y \not\succ x$$

et une relation d'indifférence $\sim$ définie par

$$x \sim y \Leftrightarrow x \succeq y \quad \text{et} \quad y \succeq x$$

$x \succeq y$ veut dire x est préféré à y ou bien il n'y a pas de préférence claire entre les deux.

Remarque 1.1. *Il est bien clair que l'asymétrie et la transitivité négative de $\succ$ sont équivalentes aux deux propriétés suivantes de $\succeq$:*

- *Complètude : $\forall x, y \in \mathcal{X}$ $x \succeq y$ ou $y \succeq x$, ou bien les deux sont vraies.*
- *transitivité : si $x \succeq y$ et $y \succeq z$ alors $x \succeq z$.*

Inversement, toute relation complète transitive $\succeq$ induit un ordre de préférences $\succ$ par la négation de $\succeq$, c-à-d.,

$$x \succ y \Leftrightarrow y \not\succeq x.$$

La relation d'indifférence $\sim$ est une relation d'équivalence, c-à-d, elle est réflexive, symétrique et transitive.

Définition 1.2. *Une représentation numérique d'un ordre de préférence noté $\succ$ est une fonction $U : \mathcal{X} \to \mathbb{R}$ telle que*

$$y \succ x \Leftrightarrow U(x) > U(y).$$

Ceci est clairement équivalent à la relation suivante

$$y \succeq x \Leftrightarrow U(x) \geq U(y).$$

Il faut remarquer qu'une telle représentation numérique U n'est pas unique : si f est une fonction strictement croissante quelconque, alors $\tilde{U}(x) := f(U(x))$ est une représentation numérique différente de U.

Définition 1.3. *Soit $\succ$ une relation de préférence sur $\mathcal{X}$. Un sous-ensemble de $\mathcal{X}$, noté $\mathcal{Z}$, est dit* ordre dense *si pour toute paire x, y dans $\mathcal{X}$ telle que $x \succ y$, il existe un $z \in \mathcal{Z}$ tel que $x \succeq z \succeq y$.*

Le résultat suivant caractérise les relations de préférence admettant une telle représentation numérique.

Théorème 1.1. *Une condition nécessaire et suffisante pour l'existence d'une représentation numérique d'une relation de préférence $\succ$ est l'existence d'un sous-ensemble dénombrable de $\mathcal{X}$, noté $\mathcal{Z}$, à ordre dense. En particulier, tout ordre de préférence sur $\mathcal{X}$ admet une représentation numérique si $\mathcal{X}$ est dénombrable.*

Définition 1.4. *Soit $\mathcal{X}$ un espace topologique. Une relation de préférence $\succ$ est continue si pour tout $x \in \mathcal{X}$*

$$\overline{\mathcal{B}}(x) := \{y \in \mathcal{X} | y \succ x\} \tag{1.1}$$

$$\underline{\mathcal{B}}(x) := \{y \in \mathcal{X} | x \succ y\} \tag{1.2}$$

sont des ensembles ouverts de $\mathcal{X}$.

Remarque 1.2. *Tout ordre de préférence qui admet une représentation numérique continue est lui-même continu. Sous certaines hypothèses sur l'ensemble $\mathcal{X}$, on montrera dans la suite que la réciproque est vraie.*

Définition 1.5 (Espace de Hausdorff)**.** *Un espace séparé ou espace de Hausdorff est un espace topologique dans lequel, pour deux points distincts x et y quelconques, il existe un voisinage de x et un voisinage de y disjoints.*

Proposition 1.1. *Soit $\succ$ un ordre de préférence dans un espace topologique de Hausdorff $\mathcal{X}$. Alors les assertions suivantes sont équivalentes.*

(a) $\succ$ est continue.

(b) L'ensemble $\{(x, y)|y \succ x\}$ est ouvert dans $\mathcal{X} \times \mathcal{X}$.

(c) L'ensemble $\{(x, y)|y \succeq x\}$ est fermé dans $\mathcal{X} \times \mathcal{X}$.

Démonstration. Raisonnons par étapes :

(a)$\Leftarrow$ (b) : On a à démontrer que pour toute paire

$$(x_0, y_0) \in M := \{(x, y)|y \succ x\}$$

il existe deux ouverts U, $V \subset \mathcal{X}$ tels que $x_0 \in U$, $y_0 \in V$, et $U \times V \subset M$. Considérons d'abord le cas où il existe un $z \in \mathcal{B}(x_0) \cap \mathcal{B}(y_0)$. Alors, on a $y_0 \succ z \succ x_0$, et donc $U := \underline{\mathcal{B}}(z)$ et $V := \overline{\mathcal{B}}(z)$ sont deux voisinages ouverts respectivement de x_0 et y_0. De plus, si $x \in U$ et $y \in V$, alors $y \succ z \succ x$, et donc $U \times V \subset M$.

Si $\overline{\mathcal{B}}(x_0) \cap \underline{\mathcal{B}}(y_0) = \emptyset$, alors on pose $U := \underline{\mathcal{B}}(y_0)$ et $V := \overline{\mathcal{B}}(x_0)$. Si $(x, y) \in U \times V$, alors par définition, on a $y_0 \succ x$ et $y \succ x_0$. L'objectif étant de montrer que $y \succ x$ afin d'établir que $U \times V \subset M$. On suppose pour cela que $x \succeq y$. Alors

$y_0 \succ y$ par transitivité négative, et donc $y_0 \succ y \succ x_0$. Mais alors, on obtient la contradiction suivante :$y \in \overline{\mathcal{B}}(x_0) \cap \underline{\mathcal{B}}(y_0) \neq \emptyset$.

(b)$\Leftarrow$ (c) : D'abord, notons que la fonction $\phi(x, y) := (y, x)$ est un homéomorphisme de $\mathcal{X} \times \mathcal{X}$. D'autre part, l'ensemble $\{(x, y)|y \succeq x\}$ est le complémentaire de l'ouvert $\phi(\{(x, y)|y \succ x\})$.

(c)$\Leftarrow$ (a) : Comme $\mathcal{X}$ est un espace topologique de Haussdorff, $\{x\} \times \mathcal{X}$ est fermé dans $\mathcal{X} \times \mathcal{X}$. De plus on a,

$$\{x\} \times \mathcal{X} \cap \{(x, y)|y \succeq x\} = \{x\} \times \{y|y \succeq x\}.$$

Par conséquent, $\{y|y \succeq x\}$ est fermé dans $\mathcal{X}$ ce qui implique que son complémentaire $\{y|x \succ y\}$ est un ouvert.

On montre ainsi par les mêmes arguments que $\{y|y \succ x\}$ est un ouvert. Ce qui achève la preuve. □

Définition 1.6 (Espace connexe.)**.** *Soit un espace topologique $\mathcal{X}$. Les trois propositions suivantes sont équivalentes :*

- *$\mathcal{X}$ n'est pas la réunion de deux ouverts non vides disjoints.*
- *$\mathcal{X}$ n'est pas la réunion de deux fermés non vides disjoints.*
- *Toute application continue $f : \mathcal{X} \to \{0, 1\}$ est constante.*

Dans le cas où ces conditions sont remplies, on dit que l'espace $\mathcal{X}$ est connexe.

Définition 1.7 (Espace séparable.)**.** *On dit qu'un espace topologique $\mathcal{X}$ est séparable s'il possède une partie Z dense au plus dénombrable.*

Proposition 1.2. *Soit $\mathcal{X}$ un espace topologique connexe et soit $\succ$ un ordre de préférences continu dans $\mathcal{X}$. Alors, tout ensemble dense $\mathcal{Z}$ de $\mathcal{X}$ est un ordre dense (définition 1.3). En particulier $\succ$ admet une représentation numérique, si $\mathcal{X}$ est séparable.*

Démonstration. Soient $x, y \in \mathcal{X}$ tel que $y \succ x$, et on considère $\overline{\mathcal{B}}(x)$ et $\underline{\mathcal{B}}(y)$ (définis dans (1.1) et (1.2)). Donc $x \in \underline{\mathcal{B}}(y)$ et $y \in \overline{\mathcal{B}}(x)$. Ceci implique que ni $\overline{\mathcal{B}}(x)$ ni $\underline{\mathcal{B}}(y)$ n'est vide. D'autre part, la négative transitivité de $\succ$ implique que $\mathcal{X} = \overline{\mathcal{B}}(x) \cup \underline{\mathcal{B}}(y)$ et comme $\mathcal{X}$ est connexe alors ces deux ensembles ne peuvent être disjoints, ce qui implique que l'ensemble $\overline{\mathcal{B}}(x) \cap \underline{\mathcal{B}}(y)$ est un ouvert non vide. Comme $\mathcal{Z}$ est dense dans $\mathcal{X}$ alors il existe

$$z \in \mathcal{Z} \cap (\overline{\mathcal{B}}(x) \cap \underline{\mathcal{B}}(y))$$

qui satisfait

$$y \succ z \succ x.$$

Par conséquent, $\mathcal{Z}$ est un ordre dense. $\square$

Théorème 1.2. *Soit $\mathcal{X}$ un espace topologique qui satisfait l'une des assertions suivantes :*

- *$\mathcal{X}$ admet une base dénombrable d'ensembles ouverts.*
- *$\mathcal{X}$ est connexe séparable.*

Alors tout ordre de préférence dans $\mathcal{X}$ admet une représentation numérique continue.

Pour la preuve, voir [16].

Lemme 1.1. *Soit $\mathcal{X}$ un espace métrique connexe qui admet un ordre de préférences continu $\succ$. Si $U : \mathcal{X} \longrightarrow \mathbb{R}$ une fonction continue, dont la restriction à un sous-ensemble $\mathcal{Z}$ dense de $\mathcal{X}$ est la représentation numérique de la restriction de $\succ$ à $\mathcal{Z}$, alors U est une représentation numérique de $\succ$ dans tout l'espace $\mathcal{X}$.*

Démonstration. Soient $x, \quad y \in \mathcal{X}$ tels que $y \succ x$. Alors pour démontrer ce lemme, il suffit de montrer que :

$$y \succ x \iff U(y) > U(x) \tag{1.3}$$

Commençons par démontrer l'implication directe.

($\Longrightarrow$) : Comme $\mathcal{Z}$ est dense dans $\mathcal{X}$, d'après la proposition 1.2, l'ensemble $\mathcal{Z}$ est un ordre dense dans $\mathcal{X}$ ce qui implique l'existence d'un $z_0 \in \mathcal{Z}$ tel que

$$y \succ z_0 \succ x.$$

Par le même raisonnement, on obtient qu'il existe $z_0' \in \mathcal{Z}$ tel que

$$z_0 \succ z_0' \succ x.$$

En itérant la procédure, on construit deux suites $(z_n)_{n \geq 0}$ et $(z_n')_{n \geq 0}$ dans $\mathcal{Z}$ telles que $z_n \to y$ et $z_n' \to x$. De plus

$$z_n \succ z_0 \succ z_0' \succ z_n'$$

ce qui implique que

$$U(z_n) > U(z_0) > U(z'_0) > U(z'_n).$$

Par la suite, par la continuité de U, et vu la convergence des suites $(z_n)_{n\geq 0}$ et $(z'_n)_{n\geq 0}$, on obtient que

$$U(z_n) \to U(y) \text{ et } U(z'_n) \to U(x).$$

Et donc

$$U(y) \geq U(z_0) > U(z'_0) \geq U(z'_n).$$

Ainsi, le premier sens de l'équivalence est prouvé.

Inversement, on suppose que x et y sont tels que $U(y) > U(x)$. On a alors par continuité de U que les ensembles

$$\overline{\mathcal{U}}(x) := \{z \in \mathcal{X} | U(z) > U(x)\}$$
$$\underline{\mathcal{U}}(y) := \{z \in \mathcal{X} | U(z) < U(y)\}$$

sont des ouverts non vides de $\mathcal{X}$. De plus $\mathcal{X} = \overline{\mathcal{U}}(x) \cup \underline{\mathcal{U}}(y)$. Or par hypothèse, $\mathcal{X}$ est un espace connexe, ce qui implique dans la définition 1.6 que $\overline{\mathcal{B}}(x) \cap \underline{\mathcal{B}}(y) \neq \emptyset$. On choisit $z_0 \in \overline{\mathcal{U}}(x) \cap \underline{\mathcal{U}}(y)$. On remarque alors que

$$U(x) < U(z_0) < U(y). \tag{1.4}$$

On définit par la suite et de la même manière $\underline{\mathcal{U}}(z_0)$. Le même raisonnement nous permet alors de déduire que $\overline{\mathcal{U}}(x) \cap \underline{\mathcal{U}}(z_0) \neq \emptyset$, donc $\forall z'_0 \in \overline{\mathcal{U}}(x) \cap \underline{\mathcal{U}}(z_0)$, en utilisant (1.4)

$$U(x) < U(z'_0) < U(z_0) < u(y). \tag{1.5}$$

Par itération, on construit ainsi deux suites $(z_n)_{n\geq 0}$ et $(z'_n)_{n\geq 0}$ dans $\mathcal{Z}$ (rappelons que $\mathcal{Z}$ est dense dans $\mathcal{X}$) telles que $z_n \to y$ et $z'_n \to x$ vérifiant $U(z_n) > U(z_0)$ et $U(z'_n) < U(z'_0)$.

Par la suite, et puisque par hypothèse U est une représentation numérique de $\succ$ dans $\mathcal{Z}$, on a :

$$z_n \succ z_0 \succ z'_0 \succ z'_n.$$

Finalement, par la continuité de $\succ$, ni $z_0 \succ y$ ni $x \succ z'_0$ n'est vraie. Grâce à la transitivité négative, on obtient alors que $y \succ x$. □

1.2.2 Représentation de Von Neumann-Morgenstern

On suppose qu'à chaque choix possible $x \in \mathcal{X}$ que peut faire l'agent correspond une distribution de probabilité dans un ensemble de scénarios possibles. L'ensemble $\mathcal{X}$ peut être donc identifié à un sous ensemble $\mathcal{M}$ de $\mathcal{M}(S, \mathcal{S})$: l'espace de toutes les distributions de probabilités sur un espace mesurable $(S, \mathcal{S})$. Les élément de $\mathcal{M}$ sont appelés *loteries*, et dans toute la suite on fera l'hypothèse que $\mathcal{M}$ est convexe.

Le but de cette section est de caractériser les ordres de préférences notés $\succ$ dans $\mathcal{M}$ possédant une représentation numérique de la forme.

$$U(\mu) = \int u(x)\mu(dx). \tag{1.6}$$

Cette formulation des préférences de loteries en termes d'utilité est due à D. Bernoulli [9].

Définition 1.8 (Représentation de Von Neumann-Morgenstern)**.** *Une représentation numérique U d'un ordre de préférence $\succ$ dans $\mathcal{M}$ est une représentation de Von Neumann-Morgenstern si elle est de la forme (1.6).*

Une représentation de Von Neumann-Morgenstern est dite affine dans $\mathcal{M}$ si et seulement si $\forall \mu,\ \nu \in \mathcal{M}$ et $\alpha \in [0, 1]$, on a

$$U(\alpha\mu + (1-\alpha)\nu) = \alpha U(\mu) + (1-\alpha)U(\nu).$$

Il est facile, par la suite, de voir que cette propriété que la préférence de μ à ν, c-à-d que $\mu \succ \nu$, est préservée dans n'importe quelle combinaison convexe à l'aide d'une troisième loterie $\lambda \in \mathcal{M}$. On appelle cette propriété axiome d'indépendance. Cette théorie axiomatique, que nous détaillerons un peu plus dans la suite, a été initiée par John von Neumann et Oskar Morgenstern dans [110].

Définition 1.9. *Une relation de préférence $\succ$ dans $\mathcal{M}$ satisfait l'axiome d'indépendance si, $\forall \mu,\ \nu \in \mathcal{M}$ la relation $\mu \succ \nu$ implique*

$$\alpha\mu + (1-\alpha)\lambda \succ \alpha\nu + (1-\alpha)\lambda$$

$\forall \lambda \in \mathcal{M}$ et $\alpha \in (0, 1]$.

Cet axiome est aussi appelé *axiome de substitution*. En effet, on considère deux loteries μ et ν de $\mathcal{M}$ telles que $\mu \succ \nu$ et on introduit une troisième loterie $\alpha\mu + (1-\alpha)\lambda$ qui consiste à tirer μ avec une probabilité α ou tirer λ avec une probabilité $1-\alpha$. Si on tire λ, alors dans ce cas il n'y a pas de différence entre considérer la loterie $\alpha\mu + (1-\alpha)\lambda$ ou $\alpha\nu + (1-\alpha)\lambda$ car le tirage réalisé vaut λ. Par contre, si on réalise μ qui est préférée à ν alors là on voit bien qu'on préfère $\alpha\mu + (1-\alpha)\lambda$ à $\alpha\nu + (1-\alpha)\lambda$.

Définition 1.10. *Une relation de préférence $\succ$ dans $\mathcal{M}$ satisfait l'axiome d'Archimède si, pour chaque triplet $\mu \succ \lambda \succ \nu$, il existe $(\alpha, \beta) \in (0,1)^2$ tel que*

$$\alpha\mu + (1-\alpha)\nu \succ \lambda \succ \beta\mu + (1-\beta)\nu$$

Cet axiome dérive de son homologue ; le principe d'Archimède en analyse, c-à-d $\forall \varepsilon > 0$ et $\forall x$ grand on peut trouver $n \in \mathbb{N}$ tel que $n\varepsilon > x$.

Cet axiome est aussi appelé *axiome de continuité*. On suppose que l'ensemble $\mathcal{M}$ est muni d'une topologie sous laquelle les combinaisons convexes de la forme $\alpha\mu + (1-\alpha)\nu$ sont continues, c-à-d convergent vers μ si $\alpha \to 1$ et vers ν si $\alpha \to 0$. Alors, dans ce cadre, la continuité de l'ordre de préférence $\succ$ implique immédiatement cet axiome.

Remarque 1.3. *Cet axiome est moins intuitif que le précédent. En effet, pour mieux le comprendre, on considère les* 3 *loteries déterministes suivantes :*

$$\mu \to 1000\text{€},\ \lambda \to 10\text{€}, \nu \to \text{ décès à coup sûr.}$$

Même pour un paramètre α très petit, il n'est pas clair qu'un agent va préférer la loterie $\alpha\mu + (1-\alpha)\nu$ (qui évoque le décès avec une probabilité non nulle) à la loterie λ qui consiste elle à gagner uniquement 10€ *à coup sûr. Pourtant, la majorité des gens n'hésite pas à faire* 100 *Km de plus par jour pour gagner* 1000€ *supplémentaires lorsqu'ils courent le risque d'un accident mortel.*

Le théorème qui suit prouve que l'axiome d'indépendance et celui d'Archimède impliquent l'existence d'une représentation numérique affine (définition).

Théorème 1.3. *Soit $\succ$ une relation de préférence qui satisfait les deux axiomes d'indépendance et d'Archimède définis ci-dessus. Alors elle admet une représentation numérique affine U. De plus U est unique à une transformation affine près, c-à-d toute autre représentation numérique affine $\tilde{U}$ de $\succ$ est de la forme :*

$$\tilde{U} = aU + b, \quad a > 0, \; b \in \mathbb{R}.$$

Avant de donner la démonstration de ce résultat, on donne deux cas importants dans lesquels une représentation de type Von Neumann-Morgenstern existe. Ceci se traduit par les deux corollaires qui suivent.

Pour le premier corollaire, on a besoin de la notion de *distribution de probabilité simple.*

Définition 1.11. *$\mu \in \mathcal{M}$ est dite une distribution de probabilité simple si c'est une mesure de probabilité dans S qui peut être écrite sous la forme d'une combinaison convexe de masse de Dirac, c-à-d $\exists\, x_1, x_2, \ldots, x_n \in S$ et $\alpha_1, \alpha_2, \ldots, \alpha_n \in (0, 1]$ tel que*

$$\mu = \sum_{i=1}^{n} \alpha_i \delta_{x_i}.$$

Maintenant, énonçons le premier corollaire.

Corollaire 1.1. *Supposons que $\mathcal{M}$ est l'ensemble de toutes les distributions de probabilité simples dans S, et supposons que $\succ$ est un ordre de préférence qui satisfait les deux axiomes d'indépendance et d'Archimède. Alors $\succ$ admet une représentation de type Von Neumann-Morgenstern U. De plus, U et u introduites dans (1.6) sont uniques à une transformation affine près.*

Démonstration. D'après le théorème 1.3, il existe une représentation numérique affine de $\succ$. On définit donc $u(x) := U(\delta_x)$, pour $x \in S$.

Si $\mu \in \mathcal{M}$, alors elle s'écrit par hypothèse sous la forme $\mu = \sum_{i=1}^{n} \alpha_i \delta_{x_i}$. Donc comme U est affine, alors ceci implique que

$$U(\mu) = \sum_{i=1}^{n} \alpha_i U(\delta_{x_i}) = \int u(x)\mu(dx),$$

ce qui prouve le corollaire. □

Corollaire 1.2. *On suppose que $\mathcal{M}$ est l'ensemble de toutes les distributions de probabilités dans un ensemble fini S, et supposons que $\succ$ est un ordre de préférence qui satisfait les deux axiomes d'indépendance et d'Archimède. Alors $\succ$ admet une représentation de type Von Neumann-Morgenstern unique à une transformation affine près.*

Démonstration. Dans un ensemble fini, toutes les mesures de probabilité sont simples. Donc ce corollaire n'est autre qu'un cas particulier du corollaire précédent. □

Lemme 1.2. *Sous les hypothèses du théorème 1.3, les assertions suivantes sont équivalentes :*

(a) *Si $\mu \succ \nu$ et si $0 \le \alpha < \beta \le 1$, alors*

$$\beta\mu + (1-\beta)\nu \succ \alpha\mu + (1-\alpha)\nu.$$

(b) *Si $\mu \succ \nu$ et $\mu \succcurlyeq \lambda \succcurlyeq \nu$ alors il existe un unique $\alpha \in [0,1]$ tel que $\lambda \sim \alpha\mu + (1-\alpha)\nu$.*

(c) *Si $\mu \sim \nu$, alors $\alpha\mu + (1-\alpha)\lambda \sim \alpha\nu + (1-\alpha)\lambda$ pour tout $\alpha \in [0,1]$ et tout $\lambda \in \mathcal{M}$.*

Démonstration. Raisonnons par étapes :

(a) : Soit $\lambda := \beta\mu + (1-\beta)\nu$. D'après l'axiome d'indépendance, on a :

$$\lambda \succ \beta\nu + (1-\beta)\nu = \nu$$

pour $\gamma := \alpha/\beta$. Ainsi, on obtient que

$$\beta\mu + (1-\beta)\nu = (1-\gamma)\lambda + \gamma\lambda \succ (1-\gamma)\nu + \gamma\lambda = \alpha\mu + (1-\alpha)\nu.$$

(b) L'unicité est une conséquence de (a). Pour l'existence, il suffit de se focaliser sur le cas $\mu \succ \lambda \succ \nu$, car les deux autres cas $\mu \sim \lambda$ et $\lambda \sim \nu$ sont triviaux et correspondent à un choix de α qui vaut respectivement 1 et 0 respectivement. Un candidat naturel est le suivant

$$\alpha := \sup\{\gamma \in [0,1] | \ \lambda \succcurlyeq \gamma\mu + (1-\gamma)\nu\}.$$

Dans la suite, on va raisonner par l'absurde. On suppose donc que $\lambda \nsim \alpha\mu + (1-\alpha)\nu$. Alors l'une des assertions suivantes est vraie,

i) $\lambda \succ \alpha\mu + (1-\alpha)\nu$,

ii) $\alpha\mu + (1-\alpha)\nu \succ \lambda$.

Si la première assertion est satisfaite, alors, en appliquant l'axiome d'Archimède, il existe $\beta \in (0,1)$ tel que

$$\lambda \succ \beta[\alpha\mu(1-\alpha)\nu] + (1-\beta)\mu = (1-\gamma)\nu + \gamma\mu, \tag{1.7}$$

pour un paramètre $\gamma = 1-\beta(1-\alpha)$ et vérifiant $\gamma - \alpha = (1-\beta)(1-\alpha) \geq 0$. Donc, par définition de α, on a forcément $\gamma\mu + (1-\gamma)\nu \succ \lambda$, ce qui est en contradiction avec (1.7).

Dans le cas où $\alpha\mu + (1-\alpha)\nu \succ \lambda$, toujours par l'axiome d'Archimède, il existe $\beta \in (0,1)$ tel que

$$\beta[\alpha\mu(1-\alpha)\nu] + (1-\beta)\nu = \beta\alpha\mu + (1-\beta\alpha)\nu \succ \lambda. \tag{1.8}$$

Clairement $\beta\alpha < \alpha$, donc par définition de α, il existe un $\gamma \in (\beta\alpha, \alpha]$ pour lequel $\lambda \succcurlyeq \gamma\mu + (1-\gamma)\nu$. Encore une fois, en utilisant le fait que $\beta\alpha < \alpha$ et (a), on obtient que

$$\lambda \succcurlyeq \gamma\mu + (1-\gamma)\nu \succ \beta[\alpha\mu(1-\alpha)\nu] + (1-\beta)\nu,$$

ce qui est en contradiction avec (1.8).
(c) Il est facile de voir que démontrer (c) est équivalent à montrer que ni l'un ni l'autre des deux cas suivants n'est vrai :

i) $\alpha\mu + (1-\alpha)\lambda \succ \alpha\nu + (1-\alpha)\lambda$,

ii) $\alpha\nu + (1-\alpha)\lambda \succ \alpha\mu + (1-\alpha)\lambda$.

Pour prouver ceci, on supposera qu'il existe $\rho \in \mathcal{M}$ tel que $\rho \nsim \mu \sim \nu$. En effet, dans le cas contraire, on a $\mathcal{M} = \{\mu\}$ et le résultat devient trivial.

On suppose maintenant que $\rho \succ \mu$, le cas $\mu \succ \rho$ étant similaire. Supposons que le premier cas est satisfait. Par l'axiome d'indépendance, et pour tout $\beta \in (0,1)$, on a

$$\beta\rho + (1-\beta)\nu \succ \beta\nu + (1-\beta)\nu = \nu \sim \mu,$$

ce qui implique que

$$\alpha[\beta\rho + (1-\beta)\nu] + (1-\alpha)\lambda \succ \alpha\mu + (1-\alpha)\lambda, \quad \forall\beta \in (0,1).$$

Par la suite, en utilisant l'hypothèse $\alpha\nu + (1-\alpha)\lambda \succ \alpha\mu + (1-\alpha)\lambda$ et (b), on déduit qu'il existe un unique γ (β fixé) tel que :

$$\begin{aligned}\alpha\mu + (1-\alpha)\lambda &\sim \gamma[\alpha(\alpha\mu + (1-\alpha)\lambda) + (1-\alpha\lambda)] + (1-\gamma)\alpha\nu + (1-\alpha)\lambda \\ &= \alpha[\beta\gamma\rho + (1-\beta\gamma)\nu] + (1-\alpha)\lambda \\ &\succ \alpha\mu + (1-\alpha)\lambda\end{aligned}$$

où on a utilisé l'équation précédente en remplaçant β par $\beta\gamma$ en dernière étape. Ceci constitue bien une contradiction.

Le deuxième cas se montre de la même manière. $\square$

Démonstration. (théorème 1.3)
Pour la construction de U, on considère dans un premier temps deux loteries λ et ρ telles que $\lambda \succ \rho$ et on définit l'ensemble

$$\mathcal{M}(\lambda,\rho) := \{\mu \in \mathcal{M} | \lambda \succeq \mu \succeq \rho\}.$$

le résultat est trivial si une telle paire $\lambda \succ \rho$ n'existe pas.

Si $\mu \in \mathcal{M}(\lambda,\rho)$, la partie ($b$) du lemme 1.2 implique qu'il existe un unique $\alpha \in [0,1]$ tel que $\mu \sim \alpha\lambda + (1-\alpha)\rho$. On pose alors $U(\mu) := \alpha$. Pour prouver que U est une représentation numérique de $\succ$ sur $\mathcal{M}(\lambda,\rho)$, on doit montrer que, pour $\nu,\ \mu \in \mathcal{M}(\lambda,\rho)$, on a $U(\mu) > U(\nu)$ si est seulement si $\mu \succ \nu$.
Commençons par l'implication inverse. On applique la partie (a) du lemme 1.2 pour conclure que :

$$\mu \sim U(\mu)\lambda + (1-U(\mu))\rho \succ U(\nu)\lambda + (1-U(\nu))\rho \sim \nu,$$

et donc $\mu \succ \nu$. Inversement, si $\mu \succ \nu$, alors l'argument précédent implique qu'il est impossible d'avoir $U(\nu) > U(\mu)$. Il reste alors à discuter le cas $U(\mu) = U(\nu)$. En effet, si $U(\mu) = U(\nu)$, alors la définition de U implique que $\mu \sim \nu$, ce qui contredit l'hypothèse $\mu \succ \nu$. On conclut alors que U est une représentation numérique de $\succ$ restreinte à $\mathcal{M}(\lambda,\rho)$.
Montrons que $\mathcal{M}(\lambda,\rho)$ est un ensemble convexe. Soient $\mu,\ \nu \in \mathcal{M}(\lambda,\rho)$ et $\alpha \in [0,1]$. Alors

$$\lambda \succeq \alpha\lambda + (1-\alpha)\nu \succeq \alpha\mu + (1-\alpha)\nu.$$

Utilisons l'axiome d'indépendance pour les cas $\lambda \succ \nu$ et $\lambda \succ \mu$, ainsi que la partie (c) du lemme 1.2 pour les cas $\lambda \sim \nu$ et $\lambda \sim \mu$. En se basant sur le même argument, on obtient que

$$\begin{aligned}
&\alpha U(\mu) + (1-\alpha)U(\nu) \\
&\sim \alpha\left(U(\mu)\lambda + (1-U(\mu))\rho\right) + (1-\alpha)\left(U(\nu)\lambda + (1-U(\nu))\rho\right) \\
&= \left[\alpha U(\mu) + (1-\alpha)U(\nu)\right]\lambda + \left[1 - \alpha U(\mu) - (1-\alpha)U(\nu)\right]\rho.
\end{aligned}$$

La définition de U ainsi que l'unicité dans la partie (b) du lemme 1.2 impliquent que

$$U(\alpha\mu + (1-\alpha)\nu) = \alpha U(\mu) + (1-\alpha)U(\nu).$$

Donc U est une représentation numérique affine de $\succ$ sur $\mathcal{M}(\lambda, \rho)$. La dernière étape consiste à montrer que la représentation numérique affine U de $\succ$ sur $\mathcal{M}(\lambda, \rho)$ est unique à transformation affine positive près. Donc, soit $\tilde{U}$ une autre représentation numérique affine de $\succ$ sur $\mathcal{M}(\lambda, \rho)$ et soit, pour $\mu \in \mathcal{M}(\lambda, \rho)$,

$$\hat{U} := \frac{\tilde{U}(\mu) - \tilde{U}(\rho)}{\tilde{U}(\lambda) - \tilde{U}(\rho)}.$$

Alors $\hat{U}$ est une transformation affine positive de $\tilde{U}$ telle que $\hat{U}(\rho) = 0 = \mathcal{U}(\rho)$ et $\hat{U}(\lambda) = 1 = \mathcal{U}(\lambda)$. Ainsi, le caractère affine de $\hat{U}$ et la définition de U impliquent que

$$\hat{U}(\mu) = \hat{U}\left(U(\mu)\lambda + (1-U(\mu))\rho\right) = U(\mu)\hat{U}(\lambda) + (1-U(\mu))\hat{U}(\rho) = U(\mu)$$

pour tout $\mu \in \mathcal{M}(\lambda, \rho)$. Alors $\hat{U} = U$.

Enfin, il reste à prouver que U peut être prolongée en une représentation numérique sur tout l'ensemble $\mathcal{M}$. On considère alors $\tilde{\lambda}$, $\tilde{\rho} \in \mathcal{M}$ tels que $\mathcal{M}(\tilde{\lambda}, \tilde{\rho}) \supset \mathcal{M}(\lambda, \rho)$. En utilisant le même argument qu'au début de cette preuve, on prouve qu'il existe une représentation numérique affine $\tilde{U}$ de $\succ$ sur $\mathcal{M}(\tilde{\lambda}, \tilde{\rho})$ sous l'hypothèse que $\tilde{U}(\lambda) = 1$ et que $\tilde{U}(\rho) = 0$. Dans le cas contraire, on peut s'y ramener par une transformation affine positive de $\tilde{U}$. D'après l'étape précédente, on conclut que $\tilde{U}$ coïncide avec U sur $\mathcal{M}(\lambda, \rho)$, ainsi $\tilde{U}$ s'avère être l'unique prolongement consistant de U. Comme chaque loterie appartient à un ensemble $\mathcal{M}(\tilde{\lambda}, \tilde{\rho})$ donné, la représentation numérique affine U peut être prolongée d'une manière unique à tous les sous-ensembles de $\mathcal{M}$. Ceci achève la démonstration.

□

Remarque 1.4. *Dans la preuve du théorème 1.3 ci-dessus, on n'a pas utilisé le fait que les éléments de $\mathcal{M}$ sont des mesures de probabilité. La clé de la démonstration consiste en la convexité de $\mathcal{M}$ et les deux axiomes d'indépendance et d'Archimède.*

Revenons maintenant au problème de construction d'une représentation numérique de type Von Neumann-Morgenstern d'une relation de préférence dans l'espace des distributions de probabilité. Si $\mathcal{M}$ est l'ensemble de toutes les distributions de probabilité dans un ensemble fini S, alors toute représentation numérique est forcément de ce type comme on a vu dans la preuve du corollaire 1.2. La situation devient un peu plus compliquée et demande plus d'hypothèses dans le cas où l'ensemble S est infini. Comme on peut le voir dans les deux exemples qui suivent, le théorème 1.3 n'est plus vrai de manière générale, c-à-d une représentation numérique de type von Neumann-Morgenstern peut ne pas exister.

Exemple 1.1. *Soit $\mathcal{M}$ l'ensemble des mesures de probabilité dans $S = 1, 2, 3, ...$ dans lequel $U(\mu) := \lim_{k\nearrow+\infty} k\mu(k)$ existe et bien fini. U est évidement affine et induit une relation de préférence qui vérifie les deux axiomes d'indépendance et d'Archimède, pourtant, U n'admet pas une représentation numérique de von Neumann-Morgenstern.*

Exemple 1.2. *Soit $\mathcal{M}$ l'ensemble des mesures de probabilité dans $S = [0, 1]$, et λ la mesure de Lebesgue dans S. Par le théorème de décomposition de Lebesgue, tout $\mu \in \mathcal{M}$ peut être décomposé comme une somme de deux mesures*

$$\mu = \mu_s + \mu_a$$

avec μ_s une mesure singulière par rapport à λ et une deuxième mesure absolument continue μ_a. On définit donc la fonction $U : \mathcal{M} \longrightarrow [0, 1]$ par

$$U(\mu) := \int x\mu_a(dx)$$

qui est bien une fonction affine dans $\mathcal{M}$ et induit une relation de préférence $\succ$ vérifiant les deux axiomes d'indépendance et d'Archimède, mais ne peut admettre la représentation numérique (1.6) car $U(\delta_x) = 0 \ \forall x$ donc la seule possibilité dans (1.6) est de prendre $u \equiv 0$. Donc la relation de préférence $\succ$ induite *par U ne peut être que triviale dans le sens où $\mu \sim \lambda$ quel que soit $\mu \in \mathcal{M}$, ce qui est en contradiction avec le fait $U(\lambda) = \frac{1}{2}$ et $U(\delta_{\frac{1}{2}}) = 0$.*

Supposons que l'ordre de préférence $\succ$ est continu dans le sens de la définition 1.4. Comme on l'a déjà remarqué, l'axiome d'Archimède est automatiquement vérifié dans ce contexte, si on suppose de plus que l'axiome d'indépendance est à son tour vérifié, alors comme l'indique le résultat suivant une représentation numérique de la forme (1.6) existe et est unique dans un sens qu'on précisera.

Théorème 1.4. *Soit $\mathcal{M} := \mathcal{M}_1(S)$ l'ensemble de toutes les mesures de probabilité dans S muni de la topologie faible, et soit $\succ$ un ordre de préférence continu dans $\mathcal{M}$ et vérifiant l'axiome d'indépendance. Alors il existe une représentation numérique de von Neumann-Morgenstern*

$$U(\mu) = \int u(x)\mu(dx) \tag{1.9}$$

où $u : S \longrightarrow \mathbb{R}$ est une fonction bornée et continue. De plus U et u sont uniques à une transformation affine près.

Démonstration. On note par $\mathcal{M}_S$ l'espace de toutes les distributions de probabilités simples dans S. Comme on l'a déjà noté, la continuité de $\succ$ induit l'axiome d'Archimède donc en appliquant le corollaire 1.1 la restriction de $\succ$ dans $\mathcal{M}_S$ admet une représentation numérique de la forme (1.6).
On montre d'abord que la fonction u dans cette représentation est bornée, pour cela on raisonne par l'absurde et on suppose par exemple que cette fonction n'est pas bornée supérieurement, donc il existe une séquence $x_0, x_1, \ldots \in S$ tel que $u(x_0) < u(x_1)$ et $u(x_n) > n$. Notons maintenant

$$\mu_n := (1 - \frac{1}{\sqrt{n}})\delta_{x_0} + \frac{1}{\sqrt{n}}\delta_{x_n}$$

et remarquons, $\mu_n \to \delta_{x_0}$ si $n \to +\infty$. la continuité de $\succ$ plus le fait que $\delta_{x_1} \succ \delta_{x_0}$ impliquent qu'à partir d'un certain rang on préférera δ_{x_1} à μ_n, c-à-d $\delta_{x_1} \succ \mu_n$. Or

$$U(\mu_n) = (1 - \frac{1}{\sqrt{n}})u(x_0) + \frac{1}{\sqrt{n}}u(x_n) > \frac{1}{\sqrt{n}}u(x_n) > \sqrt{n}, \quad \forall n$$

Ce qui est contradictoire avec $\delta_{x_1} \succ \mu_n$ pour n grand.

Continuité : de la même manière, on va raisonner par l'absurde en supposant que u n'est pas continue ce qui impliquera en particulier l'existence d'un $x \in S$ et d'une suite $x_n \to x$ si $n \to +\infty$ et tel que $u(x_n) \nrightarrow u(x)$. Par la suite, il est facile de montrer qu'il existe une sous-suite de $(x_n)_{n\geq 1}$ notée encore $(x_n)_{n\geq 1}$ tel que $u(x_n) \to l \neq u(x)$, on suppose que $u(x) - l := \varepsilon > 0$. Par hypothèse, il existe $m \in \mathbb{N}$ telque

$$|u(x_n) - l| < \frac{\varepsilon}{3}, \quad \forall n \geq m$$

Soit $\lambda := \frac{1}{2}(\delta_x + \delta_{x_m})$. Alors $\forall n \geq m$,

$$U(\delta_x) = l + \varepsilon > l + \frac{2\varepsilon}{3} > \frac{1}{2}(u(x) + u(x_m)) = U(\lambda) > l + \frac{\varepsilon}{3} > U(\delta_{x_n}$$

Ce qui implique, clairement, $\delta_x > \lambda > \delta_{x_n}$, or $\delta_{x_n} \to \delta_x$ quand $n \to +\infty$, car, par hypothèse, $x_n \to_{n\to+\infty} x$, ceci nous donne alors la contradiction voulue, et donc la continuité de u.

Pour achever la démonstration du théorème, il reste à prouver que

$$U(\mu) := \int u(x)\mu(dx), \quad \forall \mu \in \mathcal{M}$$

est bien une représentation numérique de $\succ$ dans l'ensemble $\mathcal{M}$. Comme u est bornée continue, ceci implique la continuité U dans $\mathcal{M}$ muni de la topologie faible. De plus par le théorème $A.22$ de [34] page 387, l'ensemble $\mathcal{M}_S$ est dense dans $\mathcal{M}$. La démonstration est donc achevée en appliquant le lemme 1.1. $\square$

Dans tout ce qui précède, on a présenté la théorie générale autour des relations de préférences et les fonctions d'utilités, notamment les deux axiomes d'indépendance et d'Archimède qu'un ordre de préférence doit vérifier, en plus de la continuité, pour pouvoir être représenté sous une forme de type Von Neumann-Morgenstern. Ceci est la théorie, mais d'une manière générale et en pratique, les gens ne raisonnent pas exactement de cette manière, comme le montre bien l'exemple suivant, dû à M.Allais [3].

Exemple 1.3. *("Allais Paradox") Cet exemple est basé sur les deux observations suivantes :*

– *Soit ν_1 la loterie donnée par*

$$\nu_1 = 0.33\delta_{2500} + 0.66\delta_{2400} + 0.01\delta_0$$

on gagne 2500 € *avec une probabilité* 0.33*,* 2400 € *avec une probabilité* 0.66 *et enfin rien avec une probabilité* 0.01*, et on note* μ_1 *la loterie qui nous fait gagner* 2400 *euros à coup sûr. En posant la question aux gens, on remarque immédiatement que la pluspart préfère gagner* 2400 € *à coup sûr, malgré que l'espérance des gains de la première loterie égale à* 2409 *soit plus grande que l'espérance de la deuxième* 2400*.*

– *On considère maintenant les deux loteries suivantes :*

$$\mu_2 = 0.34\delta_{2400} + 0.66\delta_0$$

$$\nu_2 = 0.33\delta_{2500} + 0.67\delta_0$$

Dans ce cas, et contrairement au premier, les gens préfèrent la loterie ν_2 *à* μ_2 *en accord avec leurs espérances qui valent* 825 € *pour* ν_2 *et* 816 *pour* μ_2*.*

Cette observation due à M.Allais [3] a été confirmée par D.Kahnemann et A.Tversky [106] en faisant des tests empiriques, où 82% *des personnes questionnés préfèrent* μ_1 *à* ν_1 *et* 83% *choisissent* ν_2 *plutôt que* μ_2*, ce qui veut dire que* 65% *des gens choisissent* $\mu_1 \succ \nu_1$ *et* $\nu_2 \succ \mu_2$*. Allais a donc remarqué que ce choix simultané est paradoxal dans le sens d'inconsistance avec le modèle de von Neumann-Morgenstern, plus précisément, toute relation de préférence pour laquelle* $\mu_1 \succ \nu_1$ *et* $\nu_2 \succ \mu_2$ *ne vérifie forcement pas l'axiome d'indépendance.*

En effet, si l'axiome d'indépendance est satisfait alors, nécessairement, on a

$$\alpha\mu_1 + (1-\alpha)\nu_2 \succ \alpha\nu_1 + (1-\alpha)\nu_2 \succ \alpha\nu_1 + (1-\alpha)\mu_2.$$

Ceci étant vrai pour tout $\alpha \in (0,1)$*, en particulier pour* $\alpha = \frac{1}{2}$ *, on obtient*

$$\frac{1}{2}(\mu_1 + \nu_2) \succ \frac{1}{2}(\nu_1 + \mu_2).$$

Or par des simples calculs

$$\begin{aligned} \frac{1}{2}(\mu_1 + \nu_2) &= 0.5\delta_{2400} + 0.165\delta_{2500} + 0.335\delta_0 \\ &= \frac{1}{2}(\mu_1 + \nu_2) = 0.5\delta_{2400} + 0.165\delta_{2500} + 0.335\delta_0 \end{aligned}$$

D'ou la contradiction. Et donc, l'axiome d'indépendance est violé par 65% *des gens questionnés.*

Remarque 1.5. *À ce stade, nous avons étudié les conditions d'existence d'une représentation numérique u d'un ordre de préférence. Dans le paragraphe* 2.4 *(p.74) du chapitre* 2 *de [34], les auteurs se penchent sur la question d'une relation de préférence $\succ$ qui satisfait : $\mu \succ \nu$ implique $\int u d\mu \geq \int u d\nu$ et ce quel que soit la fonction u : c'est la notion de préférence uniforme. De même, dans le paragraphe* 2.5 *(p.89), les auteurs introduisent la notion de préférence robuste. Mais nous ne détaillons pas ces notions dans ce chapitre car ceci n'a pas de liens avec nos travaux futurs.*

1.2.3 Une définition de l'utilité

Définition 1.12. *Soit U une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$. U est une représentation de la relation de préférence $\succeq$ si et seulement si quels que soient $X = (x_1, x_2, ..., x_n)$ et $Y = (y_1, y_2, ..., y_n)$ on a*

$$U(x_1, x_2, ..., x_n) \geq U(y_1, y_2, ..., y_n) \Longleftrightarrow X \succeq Y$$

La fonction U est dite alors la fonction d'utilité. Dans le cas où U est dérivable, la dérivée de U par rapport à la richesse x, est appelée utilité marginale.

Remarque 1.6. *L'intérêt de travailler avec des fonctions plutôt qu'avec des relations de préférence est immédiat. On peut facilement utiliser tout ce que l'on sait en analyse. De plus, et c'est bien le but recherché, on peut chiffrer la satisfaction par rapport au gain. Mais la construction de cette fonction n'est pas seulement un agrément mathématique. Il s'agit d'un véritable choix méthodologique qui peut d'ailleurs être contesté.*

1.2.4 Propriétés de la fonction d'utilité

Il est assez intuitif de supposer que, pour la quasi-totalité des biens, une augmentation de la quantité d'un bien dans un panier augmente ou laisse inchangée l'utilité retirée de ce panier. C'est pourquoi nous imposons à la fonction d'utilité d'être croissante en chacun de ses arguments :

$$\forall i, \frac{\partial U}{\partial x_i} \geq 0$$

En revanche, nous pouvons également penser que cette augmentation n'est pas indépendante de la quantité de ce bien déjà disponible dans le panier. Ainsi, si un premier gain procure une très grande satisfaction, le second est déjà moins bon. Cela signifie que l'utilité marginale de chaque nouveau gain est inférieure à celle du précédent : l'utilité marginale est décroissante,

$$\forall i, \frac{\partial^2 U}{\partial x_i^2} \leq 0.$$

Nous obtenons alors des fonctions d'utilité concaves, et donc des courbes d'indifférence définissant des ensembles convexes.
Si de plus nous supposons que chaque nouveau gain procure une satisfaction qui ne devient jamais nulle, propriété dite de non-saturation, alors dans ce cas la fonction d'utilité est strictement croissante :

$$\forall i, \frac{\partial U}{\partial x_i} > 0.$$

Le bien-fondé de cette hypothèse repose sur la rationalité de l'agent : si l'utilité est bien définie, l'agent ne perdra jamais son temps à consommer quelque chose qui est dommageable pour lui ou qui ne lui apporte rien.
Nous pouvons, enfin se restreindre aux domaines où l'utilité marginale est strictement décroissante :

$$\forall i, \frac{\partial^2 U}{\partial x_i^2} < 0.$$

1.2.5 Équivalent certain

Dans toute la suite, on va supposer que la relation de préférence considérée admet une représentation de *Neumann-Morgenstern* càd il existe une fonction u (appelé aussi fonction d'utilité : strictement concave, strictement croissante et continue) telle que pour toute variable aléatoire X admettant une distribution μ on a :

$$U(X) = \int u d\mu.$$

On définit par la suite le nombre :

$$m(X) := \int x\mu(dx).$$

Remarque 1.7. – *Pour un sous-jacent X aléatoire (actualisé) dont la distribution μ est connu, la valeur $m(X)$ est appelée "the fair price" de X.*
- *Pour un contrat financier dont μ est la distribution de ses payements futurs à percevoir avec un risque donné, la valeur $m(X)$ est dite " the fair premium".*

On considère maintenant une relation de préférence donnée $\succeq$, admettant une représentation de *Neumann-Morgenstern* :

$$U(X) = \int u d\mu$$

Appliquons par la suite le théorème des valeurs intermédiaires à la fonction u (strictement croissante et continue), il existe alors un unique réel $c(X)$ vérifiant l'identité :

$$u(c(X)) = U(X) = \int u d\mu_X \tag{1.10}$$

Définition 1.13. *L'équivalent certain du payoff aléatoire X est défini par le nombre $c(X)$, et la quantité :*

$$\rho(X) := m(X) - c(X)$$

est appelée le " risk premium" de X ou encore de la loterie μ_X.

Remarque 1.8.

- *L'équivalent certain $c(X)$ peut être interprété comme " upper bound" pour tout prix acceptable de X pour un agent de fonction d'utilité u (ou encore U).*
- *Le risk premium doit être interprété comme le montant qu'un agent est prêt à payer pour répliquer X par son espérance.*

Exemple 1.4. *("Le Paradoxe de St Petersburg")*
Considèrons la distribution suivante

$$\mu = \sum_{n=1}^{+\infty} 2^{-n} \delta_{2^{n-1}}$$

qui est le payoff du jeu suivant : on tire au pile et face jusqu'à la première fois où face apparaît. Si face apparaît au n-ième tirage on gagne 2^{n-1} euros. Jusqu'au 18ème siècle le prix de cette loterie était considéré comme étant égal à son espérance ou encore son "fair price" $m(\mu) = +\infty$. Donc le prix de ce jeu valait l' ∞ alors qu'en demandant aux gens, il est très difficile de trouver quelqu'un qui est prêt à débourser 20 euros pour ce jeu. En vue de ce paradoxe qui a été mis en évidence par Nicholas Bernoulli en 1713, Gabriel Cramer et Daniel Bernoulli [9] ont introduit indépendamment l'idée de déterminer le prix de ce jeu comme l'équivalent certain par rapport aux deux fonctions d'utilités, respectives, suivantes

$$u(x) = \sqrt{x} \text{ et } v(x) = \log(x)$$

Ces équivalents certains sont calculés à partir de l'équation (1.10) et valent :

$$C^u(\mu) = (2 - \sqrt{2})^{-2} \approx 2.91, \quad C^v(\mu) = 2$$

qui sont bien dans la fourchette de prix que les gens sont prêts à payer pour accéder à ce jeu.

Remarque 1.9. *Notons que, pour toute fonction d'utilité non bornée supérieurement, on peut modifier le payoff de telle sorte que le paradoxe réapparaisse. Il suffit juste remplacer, pour $n > 1000$, le payoff 2^n par $u^{-1}(2^n)$, pour que*

$$\int u(x)\mu(dx) = +\infty$$

Donc en choisissant une fonction d'utilité bornée on peut surmonter cette difficulté, bien que ceci en fasset apparaître d'autres.

1.2.6 Aversion au risque et équivalent certain

Pour une fonction d'utilité u suffisamment régulière, en appliquant le développement de Taylor autour des deux points $x = c(X)$ et $m = m(X)$ on obtient d'une part :

$$u(c(X)) \approx u(m) + u'(m)(c(X) - m) = u(m) - u'(m)\rho(X),$$

d'autre part :

$$\begin{aligned}
u(c(X)) &= \int u(x)\mu(dx) \\
&= \int [u(m) + u'(m)(x-m) + \frac{1}{2}u''(m)(x-m)^2 + r(x)]\mu(dx) \\
&\approx u(m) + \frac{1}{2}u''(m)var(X)
\end{aligned}$$

avec r le reste du développement de Taylor de la fonction u, on peut par la suite écrire la nouvelle identité suivante :

$$\rho(X) \approx -\frac{u''(m)}{2u'(m)}var(X) =: \frac{1}{2}ARa(m)var(X). \tag{1.11}$$

$ARa(m)$ peut être interprété comme le facteur par lequel un agent de fonction d'utilité u, multiplie son risque mesuré par $\frac{1}{2}var(X)$, dans le but de déterminer (calculer) le "risk premium" qu'il est prêt à payer.

Définition 1.14. $ARa(m)$ *est appelé* aversion au risque absolue.

On considère un agent de fonction d'utilité notée par U et de richesse initiale x. Une question qu'on peut poser est la suivante :
Quel sera l'impact d'une variation de cette richesse initiale sur le comportement de l'agent face à son risque ?
Il est relativement naturel d'imaginer que, quand sa richesse augmente, l'individu devient moins sensible au risque quand il joue à des loteries.
Comment sa fonction d'utilité peut-elle décrire son comportement ?
Comment peut-on savoir que l'agent devient $\pm$ sensible (averse au risque) ou indiffèrent ?
Si U est de classe C^2, le coefficient d' aversion absolue au risque défini ci-dessus par

$$ARa(x) = -\frac{U''(x)}{U'(x)} \tag{1.12}$$

permet de donner les réponses à nos questions.

– L'agent a une aversion non croissante pour le risque $\Longleftrightarrow$ ARa est non croissante.

- L'agent a une aversion non décroissante pour le risque $\Longleftrightarrow$ ARa est non décroissante.
- L'agent a une aversion constante pour le risque $\Longleftrightarrow$ ARa est une fonction constante de $x \Longrightarrow U$ est une transformation positive de la fonction $-e^{-\int_0^x ARa(y)dy}$.

Remarque 1.10. *Ara n'est rien d'autre que la valeur absolue de l'indice de courbure de la fonction U.*
On dira alors que le consommateur A de fonction d'utilité U a au moins autant d'aversion au risque que le consommateur B de fonction d'utilité V si U est plus concave que V en tout x càd

$$-\frac{U''(x)}{U'(x)} \geq -\frac{V''(x)}{V'(x)}$$

Comme les fonctions d'utilités ne sont uniques qu'à une transformation linéaire affine près, une mesure qui reste constante et décrit le comportement d'un agent est utile. Cette mesure est le coefficient d'aversion au risque absolue.
En effet si U est l'utilité d'un agent alors $V = aU + b$ est encore une fonction d'utilité pour ce même individu, on dérive alors une fois pour faire disparaître b, puis une deuxième fois et on prend le quotient pour faire disparaître a, ce qui nous donne :

$$V' = aU', V'' = aU'' \Longrightarrow \frac{U''(x)}{U'(x)} = \frac{V''(x)}{V'(x)}$$

Cette mesure décrit bien le comportement d'un agent sur un même marché. Elle est constante pour toute les utilités d'un même marché pour un agent donné. Si on change de marché, cette mesure ne peut nous décrire le lien entre les deux marchés.
On peut encore décrire le lien entre la prime de risque et les caractéristiques de la fonction d'utilité par l'*aversion relative au risque*, on définit alors

$$AR(x) = -\frac{xU''(x)}{U'(x)} \tag{1.13}$$

comme étant le coefficient d'aversion relative au risque. Il représente une modification multiplicative du risque absolu par rapport à un risque de référence. Ce coefficient reste encore une mesure de l'aversion au risque.

Interprétation

- L'aversion absolue au risque est la somme que l'investisseur est prêt à payer pour éviter une augmentation de la variance de sa richesse d'une unité.
- L'aversion relative au risque est la proportion de sa richesse que l'investisseur est prêt à payer pour éviter une augmentation de la variance de sa richesse d'une unité.

Exemples : Les classes de fonctions d'utilités suivantes et les coefficients d'aversion au risque correspondants sont des exemples standard.
Aversion au risque constante (CARA) : $AR(x) = \gamma \Longrightarrow u$, à une transformation affine près, est de la forme $u(x) = a - be^{-\gamma x}$, $\;a \in \mathbb{R}$, $\;b \in \mathbb{R}_+$
Aversion au risque proportionelle à $\frac{1}{x}$ (HARA) : $AR(x) = \frac{(1-\gamma)}{x}$, $\;x \in (0, +\infty)$, $\;\gamma \in [0, 1)$, toujours à une transformation affine près, on a

$$u(x) = \log(x), \text{ si } \gamma = 0$$

$$u(x) = \frac{x^\gamma}{\gamma}, \text{ si } 0 < \gamma < 1$$

Le cas $\gamma = 1$ correspond à une fonction d'utilité affine.
Une des propriétés de l'aversion au risque est donnée dans le résultat suivant.

Proposition 1.3. *Soient u et v deux fonctions différentiables dans un intervalle S. En notant par AE^u et AE^v les coefficients d'aversion au risque correspondent, on a équivalence entre les assertions suivantes :*
(a) $AE^u(x) \geq AE^v(x)$, $\;\forall x \in S$
(b) Il existe une fonction F concave strictement croissante telle que $u = F \circ v$
(c) Les risk premium associés ρ^u et ρ^v sont tels que $\rho^u(X) \geq \rho^v(X)$, $\;\forall X$.

Démonstration. (a) $\Rightarrow$ (b) Comme v est strictement croissante, on peut définir son inverse w et par conséquent la fonction $F(y) := u \circ w(y)$ est croissante, deux fois différentiable et satisfait $u = F \circ v$. En effet,

$$w' = \frac{1}{v' \circ w}, \quad w'' = AE^v \circ \frac{w}{(v' \circ w)^2}$$

donc les deux dérivées, première et seconde de F sont données par

$$F' = u'(w)w' = \frac{u'(w)}{v'(w)} > 0$$

et

$$\begin{aligned} F'' &= \frac{v'(w)u''(w) - u'(w)v''(w)}{(v' \circ w)^3} \\ &= \frac{u'(w)}{(v'(w))^2}(AE^v - AE^u) \leq 0 \end{aligned}$$

ce qui prouve que F est concave strictement croissante.

(b) $\Rightarrow$ (c) La preuve de cette implication est basée essentiellement sur l'inégalité de Jensen qui, appliquée à l'équivalent certain dans (1.10), nous montre que

$$\begin{aligned} u(c^u(\mu)) &= \int u d\mu = \int F \circ v d\mu && (1.14) \\ &\leq F(\int v d\mu) = F(v(c^v(\mu))) = u(c^v(\mu)). && (1.15) \end{aligned}$$

Par la suite $\rho^u(\mu) = m(\mu) - c^u(\mu) \geq m(\mu) - c^v(\mu) = \rho^v(\mu)$, d'où le résultat.

(c) $\Rightarrow$ (a) On suppose (c), en raisonnant par l'absurde : (a) est fausse implique l'existence d'un intervalle $\mathcal{O} \subset S$, tel que $AE^u(x) < AE^v(x), \quad \forall x \in \mathcal{O}$. On note $\mathcal{O}^v := v(\mathcal{O})$, et notons encore w l'inverse de v. D'aprés la formule de la dérivée seconde de $F = u \circ w$ établie ci-dessus, F est strictement convexe dans l'ouvet $\mathcal{O}^v$. Si, par la suite, μ est une mesure à support dans $\mathcal{O}$, alors l'inégalité (1.14) est inversée et devient stricte, à moins que μ est concentrée en un point. Ce qui implique que $\rho^u(x) < \rho^v(x)$, ce qui est contradictoire avec (c). $\square$

1.3 Un premier problème d'optimisation de Portefeuille

Etant donnée une relation de préférence dans un ensemble $\mathcal{M}$, on peut essayer de déterminer la distribution maximale de $\succ$ dans $\mathcal{M}$. Le problème d'optimisation de portefeuilles qu'on considère ici est très simple : soit X une variable aléatoire, majorée par une constante a, dans un espace de probabilité

$(\Omega, \mathcal{F}_T, \mathbb{P})$ et dont la distribution μ n'est pas dégénérée. La question est alors la suivante : quelle est la meilleure combinaison

$$X_\lambda := \lambda c + (1-\lambda)X$$

de payoff risqué X et d'actif sans risque (constant) $c \in S$?
Si on évalue X_λ par le critère de l'utilité espérée $\mathbb{E}[u(X_\lambda)]$, et si on note par μ_λ la distribution de cette variable, alors on est en train de chercher le maximum de la fonction f, définie par

$$f(\lambda) := U(\mu_\lambda) = \int u d\mu_\lambda \tag{1.16}$$

qui est une fonction strictement concave qui atteint son maximum en un unique point $\lambda^* \in [0,1]$.

Proposition 1.4. *Dans le cadre du problème d'optimisation de portefeuille décrit ci-dessus, la proportion optimale à investir dans l'actif sans risque est donnée par*

(a) on a

$$\lambda^* = 1, \quad \text{si } \mathbb{E}(X) \leq c$$

$$\lambda^* > 0, \quad \text{si } c \geq c(\mu)$$

(b) si $u \in \mathcal{C}^1(\mathbb{R})$, alors

$$\lambda^* = 1 \Longleftrightarrow \mathbb{E}(X) \leq c$$

$$\lambda^* = 0 \Longleftrightarrow c \leq \frac{\mathbb{E}[Xu'(X)]}{\mathbb{E}[u'(X)]}$$

Démonstration. (a) En appliquant l'inégalité de Jensen à f, on obtient

$$f(\lambda) \leq u(\mathbb{E}(X_\lambda)) = u((1-\lambda)X + \lambda c)$$

avec égalité si et seulement si $\lambda = 1$. Par conséquent $\lambda^* = 1$ si la quantité à droite de cette inégalité est croissante en λ, c-à-d si $\mathbb{E}(X) \leq c$.
De plus part la stricte concavité de u implique

$$\begin{aligned} f(\lambda) &\geq \mathbb{E}[(1-\lambda)u(X) + \lambda u(c)] \\ &= (1-\lambda)u(c(\mu)) + \lambda u(c) \end{aligned}$$

avec égalité si et seulement si $\lambda \in \{0,1\}$. Le terme à droite de cette inégalité est croissant en λ si $c \geq c(\mu)$, et ceci implique $\lambda^* > 0$.

(b)On a clairement $\lambda^* = 0$ si et seulement si la dérivée à droite f'_+ de f satisfait $f'_+(0) \leq 0$.

Notons u_+ et u_- les dérivés à droite et à gauche de u, et remarquons que le terme

$$\frac{u(X_\lambda) - u(X)}{\lambda} = \frac{u(X_\lambda) - u(X)}{X_\lambda - X}(c - X)$$

est presque sûrement borné par

$$u'_+(a)|c - X| \in \mathcal{L}^1(\mathbb{P}) \quad si \quad a \leq c \wedge X$$

et, donc, converge quand $\lambda \downarrow 0$ vers

$$u'_+(X)(c - X)_+ - u'_-(X)(c - X)_-.$$

Par la suite, le théorème de Lebesgue implique

$$f'_+(0) = \mathbb{E}[u'_+(X)(c - X)_+] - \mathbb{E}[u'_-(X)(c - X)_-].$$

Si maintenant, on suppose que l'une des deux assertions suivantes est vraie :

- La fonction u est dans $\mathbb{C}^1(\mathbb{R})$.
- L'ensemble $\{x | u'_+(x) \neq u'_-(x)\}$ est μ-mesure nulle,

alors la dernière égalité devient,

$$f'_+(0) = \mathbb{E}[u'(X)(c - X)]$$

c-à-d, $f'_+(0) \leq 0$ si et seulement si

$$c \leq \frac{\mathbb{E}(Xu'(X))}{\mathbb{E}(u'(X))}.$$

En répétant maintenant le même raisonnement on peut montrer

$$f'_-(1) = u'_-(c)\mathbb{E}[(c - X)_-] - u'_+(c)\mathbb{E}[(c - X)_+]$$

donc, si u est dérivable en c, alors on obtient

$$f'_-(1) = u'(c)(c - \mathbb{E}[X])$$

ce qui implique $f'_-(1) < 0$ et donc $\lambda^* < 1$ si et seulement si $\mathbb{E}(X) > c$.

Remarque 1.11. *Si $u \in \mathbb{C}^1(\mathbb{R})$ alors par concavité, en appliquant le développement de Taylor-Lagrange, l'équivalent certain est tel que*

$$u(c(X)) \leq u(X) + (c(X) - X)u'(X).$$

Comme $u(c(X) = \mathbb{E}(u(X))$ par définition, on déduit en prenant l'espérance dans cette dernière inégalité que

$$u(c(X)) \leq u(c(X)) + \mathbb{E}[(c(X) - X)u'(X)]$$

ce qui implique que

$$c(X) \geq \frac{\mathbb{E}(Xu'(X))}{\mathbb{E}(u'(X))}$$

ceci d'une part mais d'autre part, toujours par concavité,

$$u(c) \leq u(c(X)) + \mathbb{E}[(c - X)u'(X)]$$

ce qui amène à

$$\mathbb{E}[(c - X)u'(X)] \geq 0$$

pour $c = c(X)$.

□

Exemple 1.5. *Soit X un actif risqué dont le prix aujourd'hui vaut π. Étant donnée une richesse initiale w, un agent dont la fonction d'utilité u est dans $\mathbb{C}^1(\mathbb{R})$ peut investir la fraction $(1 - \lambda)w$ dans l'actif risqué et le reste λw dans l' actif sans risque, dont le taux d'intérêt est noté par r. Son payoff en maturité est alors donné par*

$$X_\lambda := \lambda wr + \frac{(1 - \lambda)w}{\pi}(X - \pi). \tag{1.17}$$

D'après la proposition précédente, il n'y a aucun intérêt d'investir dans l'actif sans risque , c-à-d $\lambda^* = 1$ si et seulement si

$$\mathbb{E}(\frac{X - \pi}{\pi}) \leq r \tag{1.18}$$

ce qui est équivalent à

$$\mathbb{E}(\frac{X}{1 + r}) \leq \pi. \tag{1.19}$$

En d'autres terme, le prix π aujourd'hui de l'actif risqué doit être supérieur à l'espérance du prix actualisé de cet actif à maturité, donc l'optimum pour l'investisseur est de tout placer en cash.