le choix risqué

introduction aux sciences de la décision

M. Cozic & B. Hill

le risque

- situation d'incertitude lato sensu : l'agent ne sait pas quelle sera la conséquence $c \in C$ qui adviendra s'il choisit l'action $a \in A$
- situation de risque: situation où le décideur a connaissance de relations probabilistes entre actions et conséquences. Il sait pour toute action a et pour toute conséquence c avec quelle probabilité l'action a produit la conséquence c.
- deux tickets de loterie:

le risque

- en théorie du choix risqué, on identifie les actions aux distributions de probabilités qu'elles induisent (par hypothèse) sur les conséquences.
- si C est fini, on peut voir une action P comme une fonction $P : C \rightarrow [0, 1]$ telle que $\sum_{c \in C} P(c) = 1$.
- remarques:
  - les conséquences $c \in C$ doivent se concevoir comme mutuellement exclusives.
  - les conséquences $c \in C$ ne sont pas nécessairement monétaires ; dans le cas particulier où elles le sont toutes, on parlera de "loterie monétaire".

1. l'espérance de gain

la logique de Port-Royal d'Arnauld et Nicole (1662)

LA LOGIQVE
OV
L'ART DE PENSER:

Contenant, outre les Regles communes, plusieurs obſeruations nouvelles propres à former le iugement.

A PARIS,
Chez { IEAN GVIGNART le pere, au premier Pillier de la grand' Sale du Palais, au Sacrifice d'Abel.
CHARLES SAVREVX, au pied de la Tour de Nôtre-Dame.
IEAN DE LAVNAY, ſous le Porche des Eſcoles de Sorbonne. }

M. DC. LXII.
AVEC PRIVILEGE DV ROY.

- Q: comment déterminer la valeur d'une action risquée ? comment choisir entre différentes actions risquées ?
- proposition fondamentale: (dans le cas de loteries monétaires) pondérer les différents gains possibles par la probabilité qu'ils arrivent
  - Arnauld & Nicole, *Logique de Port-Royal* (1662) : "...pour juger de ce que l'on doit faire pour obtenir un bien ou pour éviter un mal, il ne faut pas seulement considérer le bien ou le mal en soi, mais aussi la probabilité qu'il arrive ou n'arrive pas, et regarder géométriquement la proportion que toutes ces choses ont ensembles"

exemple

$EG(L_1) = (0.6 \times 5) + (0.4 \times 3) = 4.2$
$EG(L_2) = (0.1 \times 10) + (0.9 \times 2) = 2.8$

l'espérance de gain

- évaluation des options: l'espérance de gain d'une action identifiée à la loterie monétaire P définie sur C est

$$EG(P) = \sum_{x \in C} P(x).x$$

- règle de décision: choisir l'action dont l'espérance de gain est maximum.

le paradoxe de Saint-Petersbourg

Une pièce non-biaisée est lancée de manière répétée jusqu'à ce qu'elle tombe sur face (F). Les paiements (et probabilités induites) sont comme suit :

série	F	PF	PPF	...	P...PF	...
proba.	1/2	1/4	1/8	...	$1/2^n$	...
gain	2	4	8	...	2^n	...

Question : quelle est la valeur du pari de St-Petersbourg?
Selon le (MEG), $EG(StP) = 1 + 1 + 1 + ... = \infty$!

le paradoxe de Saint-Petersbourg

- prop de Daniel Bernouilli (1738): distinguer le gain monétaire (la conséquence, disons x) et la satisfaction que l'agent en retire ou l'utilité (disons $u(x)$). On peut alors substituer $u(x)$ à x dans l'espérance de gain pour obtenir l'espérance d'utilité :

$$EU(P) = \sum_{x \in C} P(x).u(x)$$

- comment est-ce cela permet de résoudre le Paradoxe de Saint-Petersbourg ? En imposant certaines conditions sur $u(.)$: on suppose que l'argent a une utilité marginale décroissante.
  - exemple $u(x) = log(x)$. Dans ce cas, l'espérance d'utilité $EU(StP) = log(4)$ et la valeur monétaire du Pari est (environ) 4 euros.

remarque sur l'utilité

- la notion d'utilité change de signification par rapport à celle qui servait simplement à représenter les préférences; elle reflète désormais l'intensité de ces préférences (*mais voir discussion ultérieure*)
- exemple
  $L_1 = (5\$, 0.6; 3\$, 0.4)$ vs. $L_2 = (10\$, 0.1; 2\$, 0.9)$
  $L_1 = (5\$, 0.6; 3\$, 0.4)$ vs. $L_3 = (10000\$, 0.1; 2\$, 0.9)$

  les deux situations de choix partagent la même structure du point de vue du classement ; mais (pour une partie d'entre nous), L_1 est préférable à L_2 mais pas à L_3.

2. l'espérance d'utilité

notations

- options = loteries. Si l'on suppose C fini, alors l'ensemble **P** des loteries se ramène à l'ensemble des fonctions $P : C \to [0, 1]$ telles que

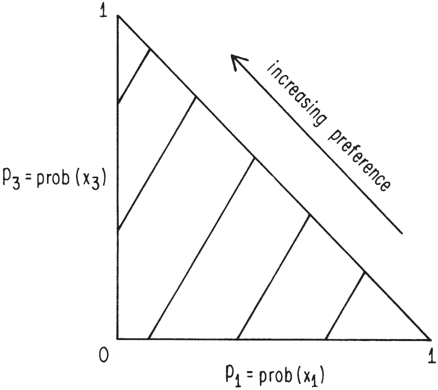
$$\sum_{x \in C} P(x) = 1$$

- si l'on numérote les conséquences $C = \{c_1, ..., c_n\}$, alors on peut écrire une loterie P de la manière suivante :

  $P = (c_1, p_1; ...; c_n, p_n)$ ou $P = (P(c_1), ..., P(c_n))$ ou $P = (p_1, ..., p_n)$

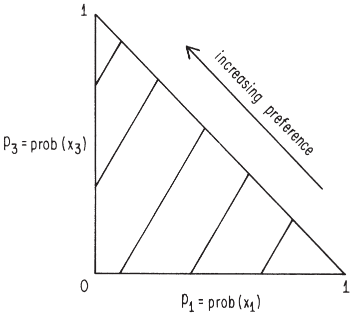
- Les lotéries peuvent également être représentées par des arbres de décision ('de profondeur un').

le triangle Marschak-Machina



représentation géométrique de l'ensemble **P** des loteries pour $|C| = 3$. Les trois extrémités sont les loteries dégénérées $\delta_1 = (1, 0)$, $\delta_2 = (0, 0)$ et $\delta_3 = (0, 1)$. p_3 est en ordonnée et p_1 en abscice.
hypothèse: $\delta_3 \succ \delta_2 \succ \delta_1$

courbes d'indifférence dans le triangle MM



les **courbes d'indifférence** sont les loteries entre lesquelles l'agent est indifférent. Deux distributions de probabilités (p_1, p_3) et (p'_1, p'_3) appartiennent à la même courbe d'indifférence ssi l'agent est indifférent entre elles ssi $EU(p_1, p_3) = EU(p'_1, p'_3)$. Une courbe d'indifférence est l'ensemble des distributions de probabilités qui ont la même espérance d'utilité (disons $\bar{u}$) - si l'agent obéit au MEU !

linéarité des courbes d'indifférence

les courbes d'indifférence sont données par des équations de la forme :

$$\bar{u} = \sum_{i=1}^{3} u(c_i).p_i = u(c_1)p_1 + u(c_2)(1 - p_1 - p_3) + u(c_3).p_3$$

Toutes les courbes d'indifférence sont des droites de pente

$$u(c_2) - u(c_1)/u(c_3) - u(c_2)$$

En effet,
$(u(c_2) - u(c_1)).p_1 + \bar{u} = (u(c_3) - u(c_2)).p_3$
$p_3 = u(c_2) - u(c_1)/u(c_3) - u(c_2).p_1 + \bar{u}/u(c_3) - u(c_2)$

mixage de loteries

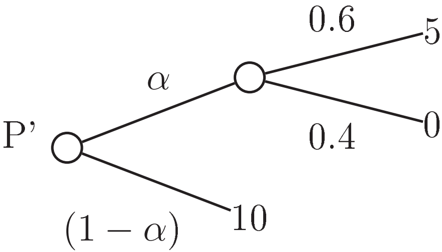
- Définition
  Soit P et Q deux distributions de probabilités sur C ; pour $a \in [0, 1]$, on appelle α-mixage de P et Q la distribution de probabilité notée $\alpha P \oplus (1-\alpha)Q$ et définie ainsi :

  $$\forall x \in C, \alpha P \oplus (1-\alpha)Q(x) = \alpha P(x) + (1-\alpha)Q(x).$$

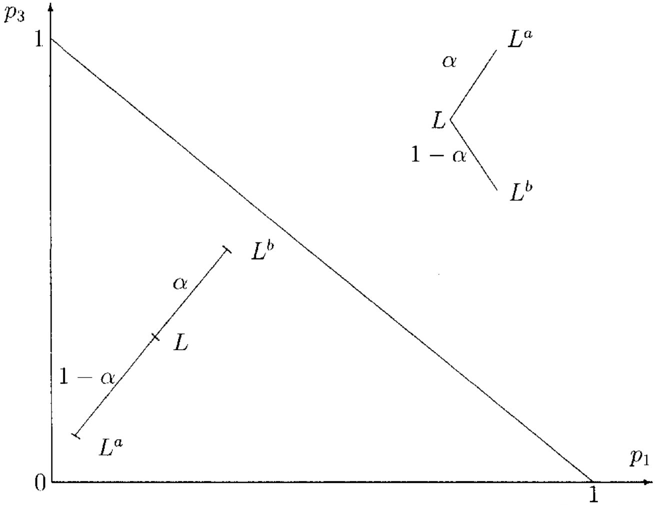
- on peut concevoir les α-mixages de distributions de probabilité comme des loteries composées où la loterie P est tirée avec probabiltié α et la loterie Q avec probabilité $(1-\alpha)$.
  Cette conception des α-mixages repose sur l'hypothèse de réduction de loteries composées, selon laquelles les gens 'réduisent' les loteries composées en loteries simples en appliquant le calcul des probabilités.

mixage de loteries et loteries composées, exemple

exemple: une manière de représenter l'α-mixage P' de P et de la loterie dégénérée qui donne 10 euros avec probabilité 1



mixage de loteries et triangle MM



mixage de loteries et triangle MM

- chaque loterie peut être conçue comme un mixage de loteries dégénérées, i.e. de loteries qui produisent l'une des conséquences $x \in C$ avec certitude. On note génériquement une telle loterie δ_x.
- soit $P \in \mathbf{P}$ une loterie. Alors on peut vérifier que

$$P = \sum_{x \in C} P(x).\delta_x$$

où la "somme" se comprend comme un mixage à $|C|$ arguments.

la préservation par mixage

- Proposition
  EU satisfait la propriété de préservation par mixage (ou linéarité):
  $EU(\alpha P \oplus (1-\alpha)Q) = \alpha EU(P) + (1-\alpha)EU(Q)$

  in french: l'espérance d'utilité d'un α-mixage de P et Q est la somme pondérée (par α et $(1-\alpha)$) de l'espérance d'utilité de P et de celle de Q.

  Preuve.
  $EU(\alpha P \oplus (1-\alpha)Q) = \sum_{x \in C}[\alpha P(x) + (1-\alpha)Q(x)]u(x)$
  $= \sum_{x \in C} \alpha P(x).u(x) + \sum_{x \in C}(1-\alpha)Q(x).u(x)$
  $= \alpha EU(P) + (1-\alpha)EU(Q)$ ♠.

- la réciproque est vraie : si $V : \mathbf{P} \to \mathbb{R}$ est une fonction d'utilité qui satisfait la préservation par mixage, alors $V(.)$ a une forme d'utilité espérée.

  Preuve:
  - on peut voir toute loterie P comme $\sum_{x \in C} P(x).\delta_x$,
  - si l'on applique ensuite la préservation par mixage on obtient $V(P) = \sum_{x \in C} P(x).V(\delta_x)$
  - il suffit alors de prendre comme fonction d'utilité $v : C \to \mathbb{R}$ la fonction $v(x) = V(\delta_x)$.

utilité espérée et préservation par mixage

- **Définition**
  Une fonction d'utilité $U : \mathbf{P} \rightarrow \mathbb{R}$ a une forme d'utilité espérée (ou est une fonction d'utilité VN-M) s'il existe une fonction d'utilité sur les conséquences (ou fonction d'utilité de Bernouilli) $u : C \rightarrow \mathbb{R}$ telle que pour toute loterie P,

$$U(P) = \sum_{x \in C} u(x).P(x)$$

- **Proposition**
  Une fonction d'utilité sur les loteries $U : \mathbf{P} \rightarrow \mathbb{R}$ a une forme d'utilité espérée ssi elle satisfait la préservation par mixage.

3. le théorème de von Neumann-Morgenstern

espérance d'utilité

Le principe de la théorie de l'espérance d'utilité peut être écrit formellement comme suit:

$$P \preceq Q \text{ ssi } \sum_{c \in C} P(c).u(c) \leq \sum_{c \in C} Q(c).u(c)$$

pour tout $P, Q \in \mathbf{P}$.

Si c'est le cas, on dit que u **EU-représente** $\preceq$.

Quand est-ce que les préférences peuvent être représentées ainsi?

Indépendance Pour tout $P, Q, R \in \mathbf{P}$ et tout $\alpha \in (0, 1)$,

$$P \preceq Q \text{ si et seulement si } \alpha P + (1-\alpha)R \preceq \alpha Q + (1-\alpha)R$$

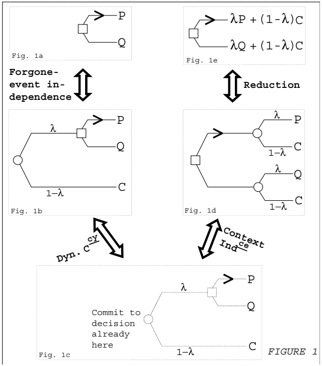
Cette propriété:

- paraît naturelle quand formulée en termes d'arbres de décision
- peut être justifiée à partir des principes de conséquentialisme et de cohérence dynamique
- est souvent satisfaite

C'est la propriété centrale de la théorie de l'espérance d'utilité (dans le risque.

(NB: des versions plus faibles, mais moins intuitives, auraient suffi pour les résultats techniques qui suivent.)

l'argument dynamique pour l'indépendance



Continuité Pour tout P, Q, R dans **P**, si $P \succ Q \succ R$, il existe $\alpha, \beta \in (0, 1)$ tels que

$$\alpha P + (1 - \alpha)R \succ Q \succ \beta P + (1 - \beta)R$$

Cette propriété:

- dit qu'une très petite probabilité d'obtenir une très mauvaise loterie ne gâche pas forcément une bonne loterie (et vice versa)
- est généralement considérée comme une condition "technique".

les axiomes

(vNM1) complétude et transitivité

(vNM2) indépendance

(vNM3) continuité

théorème de représentation

- Proposition (théorème de représentation)
  Une relation de préférence $\succeq$ sur **P** satisfait (vNM 1)-(vNM 3) ssi il existe une fonction d'utilité (de Bernouilli) $u : C \rightarrow \mathbb{R}$ telle que

  $$P \succeq Q \text{ ssi } \sum_{c \in C} u(c)P(c) \geq \sum_{c \in C} u(c)Q(c)$$

  En outre, u' se substitue à u dans cette représentation ssi $u'(.) = a.u(.) + b$ où $a > 0$ (ou encore : u est unique à une transformation affine positive près).

théorème de représentation vNM

- quels usages pour le théorème de représentation ? Principalement trois:
  - usage normatif : l'axiomatisation sert à évaluer normativement le modèle d'espérance d'utilité. En effet, pourquoi après tout obéir à la maximisation de l'espérance d'utilité ? Pourquoi ne pas suivre une autre règle ? Le théorème suggère que si l'on trouve les axiomes sur les préférences plausibles, alors l'espérance d'utilité s'en trouve justifiée.
  - usage descriptif : les préférences entre loteries sont quasiment observables ; par conséquent, une axiomatique comme celle que nous venons de présenter permet de déterminer les implications observables du modèle. Typiquement, quand on veut tester le modèle d'espérance d'utilité, on sélectionne un ou plusieurs axiomes (testables) et on observe le comportement des agents.

théorème de représentation vNM

- quels usages pour le théorème de représentation ? Principalement trois:
  - usage méta-scientifique ou fondationnel: la notion d'utilité n'est pas observable, et ainsi on pourrait se demander ce que veut dire avoir une fonction d'utilité. Le théorème fournit une réponse en termes des propriétés des préférences qui, elles, sont (supposémment) observables. On dit que le théorème donne un *fondement* pour la notion d'utilité.
    En outre, le théorème dit quelles propriétés de la fonction d'utilité ont du sens. C'est l'importance de la partie concernant l'unicité: les seules propriétés des fonctions d'utilités qui ont du sens sont celle qui sont "préservées" quand on remplace une fonction d'utilité qui représente les préférences par une autre.

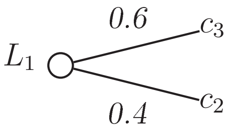
unicité à une transformation affine croissante près

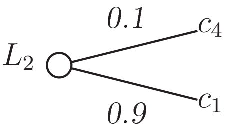
- Rappel de la partie "unicité" du théorème
  Soit $u : C \to \mathbb{R}$ une fonction d'utilité de Bernouilli qui EU-représente la relation de préférences $\succeq$ sur **P**. Alors la fonction d'utilité de Bernouilli $v :\to \mathbb{R}$ EU-représente également la relation de préférences $\succeq$ ssi il existe $a > 0$ et b tels que pour tout $x \in C$,

$$v(x) = au(x) + b$$

ordinalité et cardinalité

- on dit souvent que les fonctions d'utilité $u(.)$ ne sont pas ordinales mais cardinales: elles contiennent plus d'information que la seule façon dont l'individu classe les conséquences, elle représentent aussi l'intensité des préférences
- exemple:





ordinalité et cardinalité

soient deux fonctions d'utilité de Bernouilli $u(.)$ et $v(.)$:

	u	v
c_1	2	2
c_2	3	3
c_3	5	5
c_4	10	50

$u(.)$ et $v(.)$ classent de la même façon les conséquences. Mais, si l'on se base sur l'espérance d'utilité, elles ne classent pas de la même façon les loteries :

$EU_u(L_1) = (0.6 \times 5) + (0.4 \times 3) = 4.2$;

$EU_u(L_2) = (0.1 \times 10) + (0.9 \times 2) = 2.8$

$EU_v(L_1) = (0.6 \times 5) + (0.4 \times 3) = 4.2$;

$EU_v(L_2) = (0.1 \times 50) + (0.9 \times 2) = 6.8$

celsius et fahrentheit

Les mesures Celsius et Fahrenheit de température sont des échelles d'intervalle. Notons $(x)C$ la température du corps x en Celcius
- l'ordre est évidemment préservée: $(a)C > (b)C$ ssi $(a)F > (b)F$.
- les ratios ne sont pas préservés : il se peut par exemple que $(a)C/(b)C \neq (a)F/(b)F$

$\Rightarrow$ pas sens de dire que la température de a est deux fois plus élevée que celle de b. Si $(a)F = 100$ et $(b)F = 50$, alors $(a)C = 37.38$ et $(b)C = 10$.
- mais les ratios de différences de température sont préservés :

$$[(a)C - (b)C/(b)C - (c)C] = [(a)F - (b)F/(b)F - (c)F]$$

les invariants

- Proposition (invariance des ratios de différences)
  Soit deux fonctions d'utilité de Bernouilli u et v qui EU-représentent $\succ$. Soit quatre conséquences x_1, x_2, x_3, x_4 avec $x_3 \neq x_4$. Les ratios de différences sont invariants de u à v.

  $$(v(x_1) - v(x_2))/(v(x_3) - v(x_4)) =$$
  $$(u(x_1) - u(x_2))/(u(x_3) - u(x_4))$$

l'unicité et la révélation des utilités

L'unicité garantit qu'il fait sens de mesurer ou révéler les utilités sur les conséquences d'un individu qui obéit aux axiomes VNM. Voici comment on peut procéder pour mesurer

- Etape de calibration: puisque l'utilité est unique à une transformation affine croissante près, deux valeurs de la fonction d'utilité sont conventionnelles. Ainsi, prenons n'importe quels $\overline{x}, \underline{x} \in C$ tel que $\overline{x} \succ \underline{x}$ et leur donner n'importe quelles valeurs différentes (où la valeur donnée à $\overline{x}$ et supérieur à celle donnée à $\underline{x}$).
- Par exemple, on prend typiquement comme $\overline{x}$ la conséquence préférée et comme $\underline{x}$ la moins bonne. On dit que $u(\overline{x}) = 1$ et $u(\underline{x}) = 0$.

l'unicité et la révélation des utilités

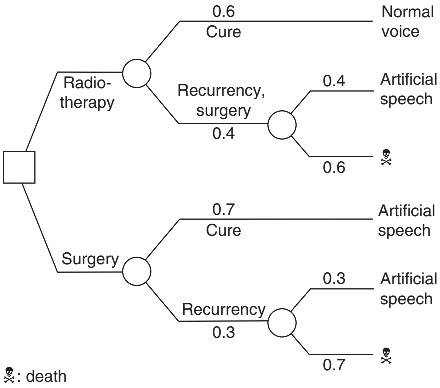
L'unicité garantit qu'il fait du sens de mesurer ou révéler les utilités sur les conséquences d'un individu qui obéit aux axiomes VNM. Voici comment on peut procéder pour mesurer

- Etape de mesure: pour savoir quelle utilité le décideur attribue à x, on lui demande de déterminer α_x t.q.

$$\delta_x \sim (\alpha_x : \bar{x}; (1 - \alpha_x) : \underline{x})$$

- on appelle α_x la standard-gamble (SG) probability
- α_x mesure l'utilité de x: il s'ensuit de la formule d'espérance d'utilité que $u(x) = \alpha_x$.

Example (from Wakker, 2010): help a patient with laryngeal cancer in stage T3 to choose between surgery (in which case the voice is lost) and radiotherapy by measuring her utility for a life with an artificial speech.



la mesure de l'utilité: la SGM

- Pour quelle valeur de p êtes vous indifférent entre la vie avec voix artificielle à coup sûr, et avoir une vie normale avec proba. p et la mort avec proba $(1 - p)$?

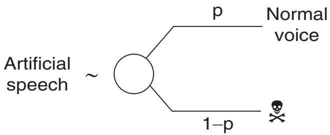


Figure: de Wakker (2010)

si par exemple la réponse est $p = 0.8$, alors c'est la radiothérapie qui a la meilleure espérance d'utilité et qui est recommandée.

preuve du théorème de représentation (existence)

D'ailleurs, une preuve du théorème de représentation procède en démontrant que ce procédé de mesure "marche" toujours: qu'il donne une valeur unique d'utilité pour chaque élément de C et que ces valeurs représentent les préférences selon le critère d'espérance d'utilité.

preuve du théorème de représentation

Nous aurons besoin du fait préliminaire suivant:

Lemme (*) Pour tout $P, Q \in L$ avec $P \succ Q$, et pour tout $\alpha, \beta \in [0, 1]$,

$\beta P \oplus (1-\beta)Q \succ \alpha P \oplus (1-\alpha)Q$ ssi $\beta > \alpha$

($\Leftarrow$). Supp. $\beta > \alpha$. On peut réécrire
$\beta P \oplus (1-\beta)Q = \gamma P \oplus (1-\gamma)[\alpha P \oplus (1-\alpha)Q]$ où $\gamma = (\beta-\alpha)/(1-\alpha)$
Par ailleurs, Indépendance implique
$P = \alpha P \oplus (1-\alpha)P \succ \alpha P \oplus (1-\alpha)Q$
et donc en ré-appliquant Indépendance
$\gamma P \oplus (1-\gamma)[\alpha P \oplus (1-\alpha)Q] \succ \alpha P \oplus (1-\alpha)Q)$
$\beta P \oplus (1-\beta)Q \succ \alpha P \oplus (1-\alpha)Q)$
($\Rightarrow$). Laissé à votre sagacité. ♠

preuve du théorème de représentation

Notation

on suppose qu'il existe une "meilleure" conséquence $\bar{x}$ et une "pire" conséquence $\underline{x}$ telles que pour tout $y \in C$, $\bar{x} \succeq y \succeq \underline{x}$. Si C fini et les préférences sont rationnelles, une telle configuration est garantie.

On démontrera que:

- le procédé de mesure décrit précédemment donne toujours une valeur pour chaque élément de C
- la valeur qu'elle donne est unique
- la fonction d'utilité ainsi définie représente les préférences selon le critère de maximisation d'espérance d'utilité.

preuve

Pour les deux premiers points, on établit le résultat suivant.

Lemme (calibration)

Si $\succ$ satisfait (A1)-(A3), alors pour toute conséquence y il existe un unique α_y t.q.[1]

$$y \sim \alpha_y \bar{x} \oplus (1 - \alpha_y)\underline{x}$$

(a) Preuve de l'existence.

Soit $\alpha_y = sup\{\alpha \in [0, 1] : \alpha \bar{x} \oplus (1 - \alpha)\underline{x}) \preceq y\}$. On va montrer qu'$\alpha_y$ ainsi défini satisfait bien la propriété désirée : $y \sim \alpha_y \bar{x} \oplus (1 - \alpha_y)\underline{x}$. Etant donné les hypothèses, il existe trois cas possibles :

[1] Pour simplifier la notation, on écrit désormais y plutôt que δ_y, et de même pour $\bar{x}$, $\underline{x}$.

preuve

(i). $\alpha_y \bar{x} \oplus (1-\alpha_y)\underline{x} \succ y \succ \underline{x}$. L'axiome de continuité nous conduit à une contradiction. Notons d'abord que pour tout $\alpha < \alpha_y$, $y \succ \alpha\bar{x} \oplus (1-\alpha)\underline{x}$. Or, par l'axiome de continuité, le cas (i) implique l'existence de $\beta \in (0,1)$ tel que $\beta(\alpha_y \bar{x} \oplus (1-\alpha_y)\underline{x}) \oplus (1-\beta)\underline{x} \succ y$. Mais $\beta(\alpha_y \bar{x} \oplus (1-\alpha_y)\underline{x}) \oplus (1-\beta)\underline{x} = \beta\alpha_y \bar{x} \oplus (1-\beta\alpha_y)\underline{x}$. Puisque $\beta\alpha_y < \alpha_y$, contradiction.

(ii). $\bar{x} \succ y \succ \alpha_y \bar{x} \oplus (1-\alpha_y)\underline{x}$. L'axiome de continuité nous conduit derechef à une contradiction par un argument analogue.

(iii). $y \sim \alpha_y \bar{x} \oplus (1-\alpha_y)\underline{x}$ est le seul cas restant.

(b) Preuve de l'unicité. Une conséquence immédiate du Lemme (*): si α et β avec $\alpha > \beta$ satisfont la condition dans la définition de la mesure, alors $\beta\bar{x} \oplus (1-\beta)\underline{x} \sim y \sim \alpha\bar{x} \oplus (1-\alpha)\underline{x} \succ \beta\bar{x} \oplus (1-\beta)\underline{x}$, ce qui contredit la transitivité des préférences.

preuve I

Grace au Lemme de calibration, on a une fonction $u : C \to \Re$ définie de manière suivante: pour tout $y \in C$, $u(y) = \alpha_y$. Venons maintenant au troisième point: il faut démontrer que u représente $\preceq$. C'est-à-dire, il faut démontrer que

Pour tout $P, Q \in L$, $P \succeq Q$ ssi $\sum_{y \in C} P(y).u(y) \geq \sum_{y \in C} Q(y).u(y)$.

Preuve.

- Par le Lemme de calibration, pour chaque y, $y \sim \alpha_y \bar{x} \oplus (1-\alpha_y)\underline{x}$. Soit P la loterie $(y_1, p_1; \ldots; y_n, p_n)$. Notez que $P = p_1 \delta_{y_1} \oplus (1-p_1)R_1$, où $R_1 = (y_2, \frac{p_2}{1-p_1}; \ldots; y_n, \frac{p_n}{1-p_1})$. Il s'ensuit de l'axiome d'indépendance que $P \sim p_1(\alpha_{y_1}\bar{x} \oplus (1-\alpha_{y_1})\underline{x}) \oplus (1-p_1)R_1 = (y_2, \frac{p_2}{1-p_1}; \ldots; y_n, \frac{p_n}{1-p_1}; \bar{x}, p_1\alpha_{y_1}; \underline{x}, p_1(1-\alpha_{y_1}))$. En répétant ce raisonnement pour tout y_i, on démontre que $P \sim P'$, où P' est telle que $P'(\bar{x}) = \sum_{y \in C} P(y)\alpha_y$, $P'(\underline{x}) = \sum_{y \in C} P(y)(1-\alpha_y)$.[2]

preuve II

- De même, $Q \sim Q'$, où $Q'(\bar{x}) = \sum_{y \in C} Q(y)\alpha_y$, $Q'(\underline{x}) = \sum_{y \in C} Q(y)(1-\alpha_y)$
- $P \succeq Q$ ssi

  $\sum_{y \in C} P(y)\alpha_y \cdot \bar{x} \oplus \sum_{y \in C} P(y)(1-\alpha_y) \cdot \underline{x} \succeq$
  $\sum_{y \in C} Q(y)\alpha_y \cdot \bar{x} \oplus \sum_{y \in C} Q(y)(1-\alpha_y) \cdot \underline{x}$,
  donc par le Lemme (*), $P \succeq Q$ ssi $\sum_{y \in C} P(y)\alpha_y \geq \sum_{y \in C} Q(y)\alpha_y$
- Par la définition de u, il s'ensuit que $P \succeq Q$ ssi $\sum_{y \in C} P(y).u(y) \geq \sum_{y \in C} Q(y).u(y)$. ♠.

---

[2] En traçant les arbres de décision, la logique de cette étape sort plus clairement. Tracez tout premièrement la loterie composée obtenue en remplaçant y_1 par $\alpha_{y_1}\bar{x} \oplus (1-\alpha_{y_1})\underline{x}$. Par indépendance, cette loterie est indifférente à P. Répétant pour tous les y, on obtient une loterie composée qui est indifférente à P. Par notre supposition concernant le rapport entre des loteries composées et les loteries simples (la "Réduction des Loteries Composées"), cette loterie peut être considérée comme le même qu'une loterie simple obtenu en "multipliant les probabilités": c'est la loterie P'.

4. attitudes par rapport au risque

illustration par le triangle MM

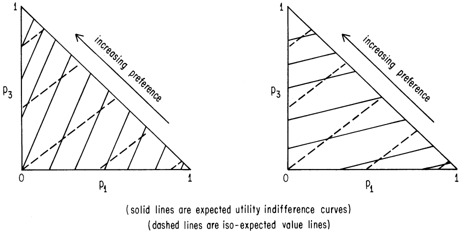


Fig. 3a. Relatively steep indifference curves of a risk averter

Fig. 3b. Relatively flat indifference curves of a risk lover

courbes d'iso-espérance de gain = courbes qui relient les points qui ont la même espérance de gain. Décrites par des équations de la forme suivante :

$$\bar{g} = \sum_{i=1}^{3} c_i.p_i = c_1.p_1 + c_2.(1 - p_1 - p_3) + c_3.p_3$$

illustration par le triangle MM

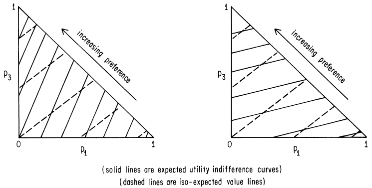


Fig. 3a. Relatively steep indifference curves of a risk averter

Fig. 3b. Relatively flat indifference curves of a risk lover

considérons l'origine 0 c'est-à-dire la loterie (dégénérée) qui donne c_2 avec certitude. Un décideur est risquophode s'il préfère cette loterie à une loterie qui a la même espérance de gain, en particulier à la loterie L_0, intersection de la droite d'iso-espérance de gain qui passe par 0 et de la droite (1,1).

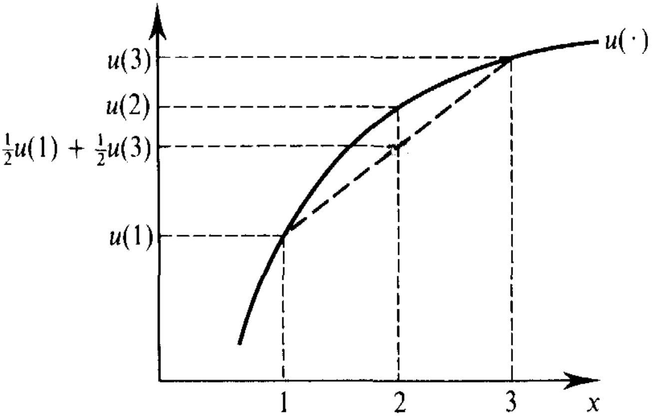
attitudes, définition

- Notation
  Soit $P \in \mathbf{P}$ une loterie monétaire et soit $EG(P)$ l'espérance de gain de P. On note $\delta_{EG(P)}$ la loterie dégénérée qui produit $EG(P)$ avec certitude
- une relation de préférence est
  - risquophobe (ou averse au risque) si pour toute loterie $P \in \mathbf{P}$, $\delta_{EG(P)} \succeq P$
  - neutre à l'égard du risque si pour toute loterie $P \in \mathbf{P}$, $\delta_{EG(P)} \sim P$
  - risquophile si pour toute loterie $P \in \mathbf{P}$, $\delta_{EG(P)} \preceq P$

formes de $u(.)$ et attitudes

- **Définition**
  Soit $f : C \to \mathbb{R}$;
  - f est concave si pour tous $x, y \in C$ et $a \in [0, 1]$, $f(ax + (1-a)y) \geq af(x) + (1-a)f(y)$
  - f est affine s'il existe a, b tels que $f(x) = ax + b$
  - f est convexe si pour tous $x, y \in C$ et $a \in [0, 1]$, $f(ax + (1-a)y) \leq af(x) + (1-a)f(y)$
- **Proposition**
  Soit $\succ$ une relation de préférence qui se laisse EU-représenter par une fonction sur les conséquences $u(.)$;
  - $\succ$ est risquophobe ssi $u(.)$ est concave
  - $\succ$ est neutre à l'égard du risque ssi $u(.)$ est affine
  - $\succ$ est risquophile ssi $u(.)$ est convexe

concavité et risquophobie



liens (désirables ?) entre aversion pour le risque et utilité marginale décroissante

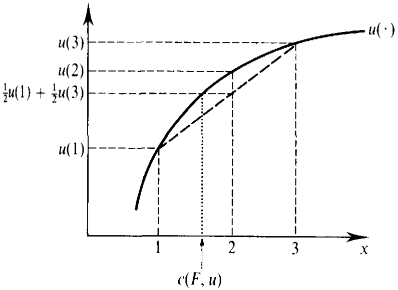
- utilité marginale décroissante (UMD) $\approx$ plus l'agent a de richesse (monétaire), moins l'incrément d'une unité monétaire lui procure d'utilité. $u'(.)$ est non croissante ou encore $u''(.)$ est non positive. Si $u(.)$ est concave, alors (étant donné les conditions mathématiques appropriées), $u'(.)$ est non croissante.
- l'UMD signifie que, à un niveau donné de richesse, l'augmentation d'utilité procurée par le gain d'1 euro est inférieure à la perte d'utilité infligée par la perte d'1 euro. C'est ce que met en évidence la figure pour un niveau de départ de 2 euros : $u(3) - u(2) < u(2) - u(1)$ soit $u(2) > (u(3) + u(1))/2$. De cela suit que l'agent qui a cette fonction d'utilité préfère le statu quo à la loterie $L = 1/2(\delta_3) \oplus 1/2(\delta_1)$. En effet, $EU(\delta_2) = u(2)$ tandis que $EU(L) = (u(3) + u(1))/2$. Donc $EU(\delta_2 > EU(L)$. Graphiquement, $EU(L)$ correspond au milieu du segment $(u(3), u(1))$.

équivalent-certain et prime de risque

- Définition
  Soit P une loterie ; l'ensemble des équivalents-certains de P, noté $C(P)$, est l'ensemble des conséquences $x \in C$ telles que l'agent est indifférent entre la loterie dégénérée qui produit x avec certitude et P : $C(P) = \{x \in C : \delta_x \sim P\}$
- sous certaines conditions, cet ensemble a un seul élément et on parle alors de **l'** équivalent-certain $C(P)$
- sous cette hypothèse, on peut définir la prime de risque $R(P) = EG(P) - C(P)$

équivalent-certain, exemple

équivalent certain de la loterie $L = 1/2(\delta_3) \oplus 1/2(\delta_1)$. On sait que l'agent est risphophobe donc $C(L) < EG(L)$. $C(L)$ s'obtient en prenant l'inverse de $(u(3) + u(1))/2$ (l'utilité espérée de L) par la fonction $u(.)$: $C(L) = u^{-1}((u(3) + u(1))/2)$.



coefficient d'aversion au risque

- définition (Arrow-Pratt)
  Supposons que les préférences de l'individu soient représentables par une fonction d'utilité de Bernouilli deux fois dérivable $u(.)$. Le coefficient absolu d'aversion au risque au point $x \in C$ est défini comme

$$\rho(x) = -u''(x)/u'(x)$$

- Fait
  $\rho(x)$ est indépendant de la fonction d'utilité choisie.

famille CARA

famille exponentielle ou *Constant Absolute Risk Aversion*:

pour $\theta > 0$, $u(x) = 1 - e^{-\theta x}$
pour $\theta = 0$, $u(x) = x$
pour $\theta < 0$, $u(x) = e^{-\theta x} - 1$

Fait

$\rho(x) = \theta =$ constante

famille CARA

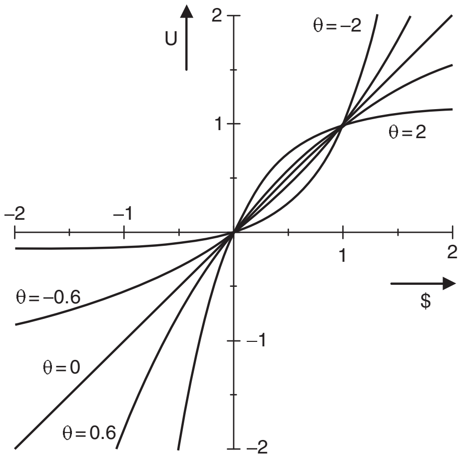


Figure 3.5.2 Exponential utility, normalized at 0 and 1.

Figure: De Wakker (2010)

5. critique 1: comment EU rend compte du risque

assurance et jeux d'argent I

Il y a des gens qui assurent leurs biens mais qui parient. Comment la théorie d'EU peut-elle rendre compte des deux comportements?

- acheter de l'assurance indique une aversion au risque.
- acheter des paris indique une risquophilie.

Vu ce qui a été dit dans la section précédente, il paraît que la seule manière de rendre compte de ces comportements est avec une fonction d'utilité de ce la forme suivante:

assurance et jeux d'argent II

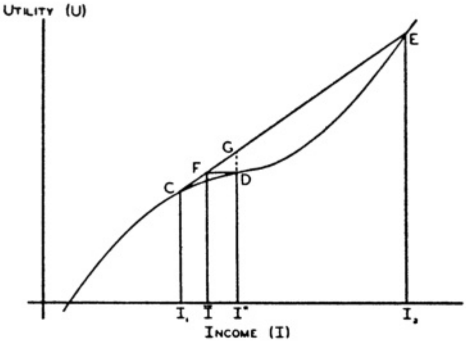


FIG. 2.—Illustration of utility function consistent with willingness of a low-income consumer unit both to purchase insurance and to gamble.

assurance et jeux d'argent III

C'est ce qui a été proposé par Savage et Friedman en 1948.

Or:

- pourquoi tout le monde est sur le point d'inflexion entre risquophilie et risquophobie?
- et si l'on gagne la loterie, et donc accroît tout d'un coup sa richesse, devient-on risquophile? Cesse-t-on par exemple d'assurer ses biens?

la dépendance pa rapport à la référence I

Qu'est-ce qui est à la source de ce problème?

Une suggestion:

- les conséquences (les objets de la fonction d'utilité) ne sont pas des valeurs absolues de richesse, comme on le suppose dans l'analyse standard du risque: ce sont des changements de richesse par rapport à un point de référence. Autrement dit, ce sont des gains et des pertes relativement à un point de référence.

la dépendance pa rapport à la référence II

Ainsi

- Une fonction d'utilité qui est convexe dans les gains et concave dans les pertes expliquerait l'assurance et les paris. De plus, le point d'inflexion change avec le niveau de richesse, sous l'hypothèse que le niveau de richesse est le point de référence.

La notion de point de référence a été introduite dans la théorie de la décision par Kahneman & Tversky (1979).

la dépendance pa rapport à la référence III

Problème

Comme le notent K & T, l'explication qui vient d'être proposée pour l'assurance et les jeux d'argent n'est pas tout à fait convaincante. Les gens tendent à choisir:

- $(500, 1)$ plutôt que $(1000, 0.5; 0, 0.5)$ [indiquant une risquophobie dans les gains]
- $(-1000, 0.5; 0, 0.5)$ plutôt que $(-500, 1)$ [indiquant une risquophilie dans les pertes]

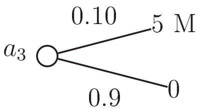
Ainsi la fonction d'utilité serait concave dans les gains et convexe dans les pertes ...l'inverse de ce qu'il faudrait pour rendre compte de l'assurance et des jeux d'argent!

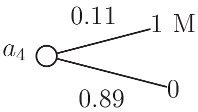
6. critique 2: le paradoxe d'Allais et les anomalies empiriques

Maurice Allais (1911-2010)

Allais, choix 1

Allais, choix 2





Allais, résultats

Comportement modal : $a_1 \succ a_2$ et $a_3 \succ a_4$ [[$(0, 100\%)$, $(100\%, 0)$]]

Intuition : dans le Choix 1, il ne vaut pas la peine d'avoir une chance (même faible) de repartir avec 0 euro pour avoir une chance (seulement) modérée d'obtenir 5M. Donc a_1. Dans le Choix 2, on a dans tous les cas une très grande chance de repartir avec 0 euro. Dans ce cas, autant sacrifier une petite partie de sa probabilité de gagner pour gagner vraiment beaucoup dans l'éventualité où l'on gagne. Donc a_3.

Allais et l'espérance d'utilité

- Qu'est-ce qui ne va pas (du point de vue du MEU) avec le comportement modal ? On peut concevoir $a_1, ..., a_4$ comme des loteries composées. Considérons en effet les loteries P, Q, Φ_0, Φ_1 :
  $P = \Phi_1 = (1M, 1)$
  $\Phi_0 = (0, 1)$
  $Q = (5M, 10/11; 0, 1/11)$
- on peut facilement vérifier que
  $\triangleright a_1 = 0.11P \oplus 0.89\Phi_1$
  $\triangleright a_2 = 0.11Q \oplus 0.89\Phi_1$
  $\triangleright a_3 = 0.11Q \oplus 0.89\Phi_0$
  $\triangleright a_4 = 0.11P \oplus 0.89\Phi_0$

  $\Rightarrow$ **le comportement modal viole l'axiome d'indépendance**

- Une première représentation des quatre loteries :

	0.01	0.89	0.1
a_1	1M	1M	1M
a_2	0	1M	5M
a_4	1M	0	1M
a_3	0	0	5M

Allais dans le triangle MM

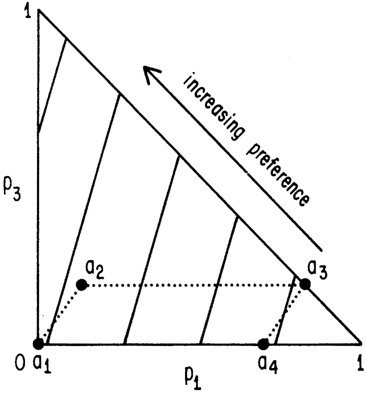


Fig. 4a. Expected utility indifference curves and the

Allais dans le triangle MM

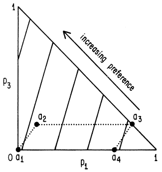


Fig. 4a. Expected utility indifference curves and the Allais Paradox

(a_1, a_2) et (a_3, a_4) sont des droites parallèles:
- la droite (a_1, a_2) est décrite par $y = 10x$
- la droite (a_3, a_4) est décrite par $y = 10x - 8.9$

Allais dans le triangle MM

- point crucial : si les préférences de l'agent obéissent aux axiomes du MEU, alors quelles qu'elles soient, si a_1 est préféré à a_2, alors a_4 doit être préféré à a_3 (et réciproquement). Pourquoi ?
- Parce que dans le triangle MM les courbes d'indifférence d'un agent se conformant au MEU sont (1) linéaires et (2) parallèles entre elles.
- Les courbes d'indifférences "observées" correspondent plutôt à la figure suivante, où les courbes ne sont pas parallèles mais partent dans plusieurs directions (*fan out*). a_1 est "au-dessus" de a_2 tandis que a_4 est "en-dessous" de a_3.

Allais dans le triangle MM

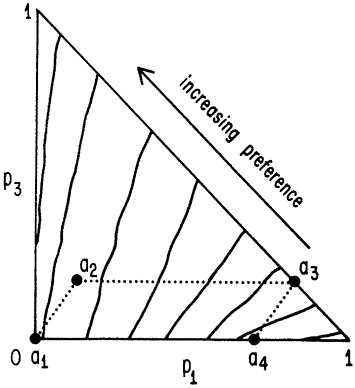


Fig. 4b. Indifference curves which 'fan out' and the Allais Paradox

Figure: D. Kahneman

Figure: A. Tversky

common consequence effect

- le PA est une instance du **common consequence effect** :

choix 1	$a_1 = \alpha\delta_c \oplus (1-\alpha)L_3$	$a_2 = \alpha L_1 \oplus (1-\alpha)L_3$
choix 2	$a_4 = \alpha\delta_c \oplus (1-\alpha)L_2$	$a_3 = \alpha L_1 \oplus (1-\alpha)L_2$

où

(i) δ_c est la loterie (dégénérée) qui produit la conséquence c avec certitude [dans le PA, il s'agit de P soit 1M].
(ii) L_1 est une loterie qui comprend au moins une issue meilleure que c et une issue moins bonne que c [dans le PA, Q dont les conséquences sont 0M et 5M]
(iii) L_3 domine stochastiquement L_2 [dans PA, Φ_0 dom.stoch. Φ_1]

common consequence effect

choix 1	$a_1 = \alpha\delta_c \oplus (1-\alpha)L_3$	$a_2 = \alpha L_1 \oplus (1-\alpha)L_3$
choix 2	$a_4 = \alpha\delta_c \oplus (1-\alpha)L_2$	$a_3 = \alpha L_1 \oplus (1-\alpha)L_2$

• on parle de “conséquence commune” parce que a_1 et a_2 (resp. a_3 et a_4) ont pour conséquence commune L_3 (resp. L_2).

• interprétation de Machina (1987, p. 10) : concevons ces loteries composées comme issues d'un tirage d'une pièce qui tombe sur le côté α ou sur le côté $(1-\alpha)$. Ce qui se passe, c'est que l'aversion au risque des décideurs sur ce qui peut se passer si α change selon ce qui se passe si $(1-\alpha)$. Meilleure est la loterie obtenue si $(1-\alpha)$, plus le décideur est risquophile. (“déjà, si α, je perds beaucoup, alors je ne veux pas prendre de risque.”)

common ratio effect

- une autre anomalie qui met en jeu l'axiome d'indépendance :

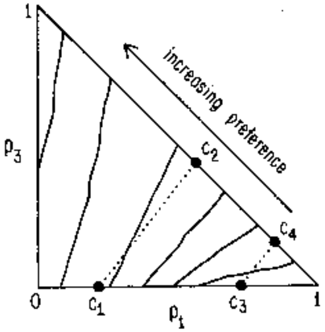
choix 1	$c_1 = (p, X; 1-p, 0)$	$c_2 = (q, Y; 1-q, 0)$
choix 2	$c_3 = (rp, X; 1-rp, 0)$	$c_4 = (rq, Y; 1-rq, 0)$

- exemple (KT 79)

choix 1	$c_1 = (0.8, 4000)$	$c_2 = (1, 3000)$
choix 2	$c_3 = (0.2, 4000)$	$c_4 = (0.25, 3000)$

common ratio effect et triangle MM

- comme pour le common consequence effect, on peut voir avec le triangle de Marschack-Machina que le MEU implique $c_1 \succ c_2$ ssi $c_3 \succ c_4$ (suit de l'axiome d'indépendance).



Indifference curves which fan out and the

résultats

choix 1	$c_1 = (0.8, 4000)$	$c_2 = (1, 3000)$
choix 2	$c_3 = (0.2, 4000)$	$c_4 = (0.25, 3000)$

Comportement modal : $c_1 \prec c_2$ (80 %) et $c_3 \succ c_4$ (65 %)

Interprétation : “Apparently, reducing the probability of winning from 1.0 to 0.25 has a greater effect than the reduction from 0.8 to .2”

reflection effet

choix 1	$c_1 = (0.8, 4000)$	$c_2 = (1, 3000)$
choix 2	$c'_1 = (0.8, -4000)$	$c'_2 = (1, -3000)$

Comportement modal : $c_1 \prec c_2$ et $c'_1 \succ c'_2$

Interprétation : $c_1 \prec c_2$ manifeste risquophobie pour les gains ; mais $c'_1 \succ c'_2$ manifeste risquophilie pour les pertes.

attitudes par rapport au risque

- les choses se compliquent :

(a) risquophilie pour les petites probabilités de gain

choix 1	$d_1 = (1, 500)$	$d_2 = (0.5, 1000)$
choix 2	$d'_1 = (1, 5)$	$d'_2 = (0.001, 5000)$

Comportement modal : $d_1 \succ d_2$ (84 %) et $d'_1 \prec d'_2$ (72 %)

attitudes par rapport au risque

- les choses se compliquent :

(b) ...et risquophobie pour les petites probabilités de pertes !

choix 1	$e_1 = (1, -500)$	$e_2 = (0.5, -1000)$
choix 2	$e'_1 = (1, -5)$	$e'_2 = (0.001, -5000)$

Comportement modal : $e_2 \succ e_1$ (65 %) et $e'_1 \succ e'_2$ (83 %)

synthèse

- on obtient donc la structure quaternaire suivante d'attitudes par rapport au risque :

	petites proba.	proba. moyenne à forte
gain	risquophile	risquophobe
perte	risquophobe	risquophile

Que faire?

- Un ingrédient de base:

(1) une fonction de pondération $\pi : [0,1] \rightarrow [0,1]$ qui déforme les probabilités ("probability weighting function") :

“decision weights are inferred from choices between prospects much as subjective probabilities are inferred from preferences in the Ramsey-Savage approach. However, decision weights are not probabilities : they do not obey the probability axioms and they should not be interpreted as measures of degree of belief...Decision weights measure the impact of events on the desirability of prospects and not merely the perceived likelihood of these events.”

Que faire.

- Un ingrédient de base:

(1) une fonction de pondération $\pi : [0,1] \to [0,1]$
$\pi(0) = 0$ et $\pi(1) = 1$. La fonction de pondération sur-pondère les petites probabilités et sous-pondère les probabilités moyennes à fortes.

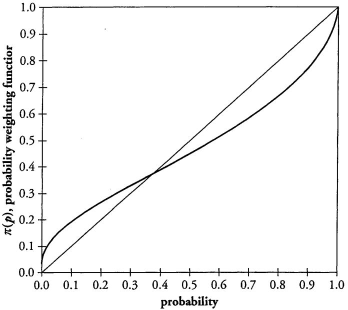
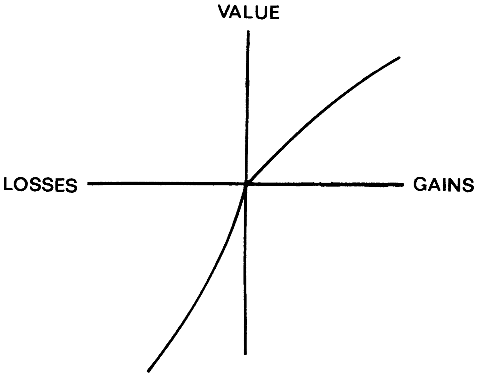


Figure 20.2 A characteristic probability weighting function

Que faire?

- L'autre ingrédient de base (voir section précédente):

(2) une **fonction de valeur** $v(.)$ est
(i) concave pour les gains
(ii) convexe pour les pertes
(iii) *steeper* pour les pertes que pour les gains (*loss aversion*)



le point de référence

- propriété très importante de la fonction de valeur: la dépendance par rapport à un point de référence
  - "An essential feature of the present theory is that the carriers of value are changes in wealth or welfare rather than final states. This assumption is compatible with basic principles of perception and judgment. Our perceptual apparatus is attuned to the evaluation of changes or differences rather than to the evaluation of absolute magnitudes."

Ce sont les deux ingrédients principaux de la théorie des perspectives proposée par K & T en 1979.

évaluation des loteries

A partir de la fonction de pondération et de la fonction de valeur, on obtient la fonction $V(.)$ qui évalue les loteries :

$$(1)V(P) = \pi(p_1)v(c_1) + \pi(p_2)v(c_2) \tag{1}$$

Cela vaut pour c_1, c_2 ni strictement positifs tous les deux, ni strictement négatifs. Si l'on est dans l'un ou l'autre de ces cas de figure, alors, supposons que c_2 soit la composante la moins risquée (la plus proche de 0),

$$(2)V(P) = v(c_2) + \pi(p_1)[v(c_1) - v(c_2)] \tag{2}$$

L'équation (1) se réduit à l'équation (2) si $\pi(p_1) + \pi(1 - p_1) = 1$ - ce qui n'est pas le cas général.

attitudes par rapport au risque

choix 1	$f_1 = (0.25, 6000)$	$f_2 = (0.25, 4000; 0.25, 2000)$
choix 2	$f_1' = (0.25, -6000)$	$f_2' = (0.25, -4000; 0.25, -2000)$

Comportement modal : $f_1 \prec f_2$ et $f_1' \succ f_2'$
On applique l'équation (1) :
$V(f_1) < V(f_2) \Leftrightarrow$
$\pi(0.25).v(6000) < \pi(0.25).[v(4000) + v(2000)] \Leftrightarrow$
$v(6000) < [v(4000) + v(2000)]$ - ce dont on peut rendre compte par v concave pour les gains.

$V(f_1') > V(f_2') \Leftrightarrow$
$\pi(0.25).v(-6000) > \pi(0.25).[v(-4000) + v(-2000)] \Leftrightarrow$
$v(-6000) < [v(-4000) + v(-2000)]$ - ce dont on peut rendre compte par v convexe pour les pertes.

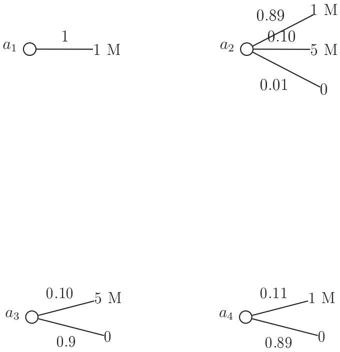
attitude par rapport au risque

- attention : dans le MEU, l'attitude par rapport au risque dépend exclusivement de la fonction d'utilité. Dans la TP, ce n'est plus le cas : la fonction de pondération participe également à l'attitude par rapport au risque.

choix 1	$g_1 = (0.001, 5000)$	$g_2 = (1, 5)$
choix 2	$g_1' = (0.001, -5000)$	$g_2' = (1, -5)$

Comportement modal : $g_1 \succ g_2$ et $g_1' \prec g_2'$ (risquophilie pour gain et risquophobie pour perte).
$g_1 \succ g_2 \Leftrightarrow \pi(0.001).v(5000) > v(5)$ donc $\pi(0.001) > v(5)/v(5000) > 0.001$ si v est concave. π grossit donc les petites probas.

common consequence effect



common consequence effect

comportement modal : $a_1 \succ a_2$ et $a_3 \succ a_4$.
On en infère $v(1M) > \pi(0.89).v(1M) + \pi(0.10).v(5M)$ et $\pi(0.10).v(5M) > \pi(0.11).v(1M)$. Donc

$$v(1M) > \pi(0.89).v(1M) + \pi(0.11).v(1M)$$
$$v(1M) > [\pi(0.89) + \pi(0.11)].v(1M)$$
$$1 > \pi(0.89) + \pi(0.11)$$

propriété de subcertainty est la suivante : $\pi(p) + \pi(1-p) < 1$.
"...subcertainty captures an essential element of people's attitudes to uncertain events, namely that the sum of the weights associated with complementary events is typically less than the weight associated with the certain event."

common ratio effect

choix 1	$c_1 = (0.8, 4000)$	$c_2 = (1, 3000)$
choix 2	$c_3 = (0.2, 4000)$	$c_4 = (0.25, 3000)$

Comportement modal : $c_1 \prec c_2$ (80 %) et $c_3 \succ c_4$ (65 %)

$\pi(1).v(3000) > \pi(0.8).v(4000)$ $\pi(0.2).v(4000) > \pi(0.25).v(3000)$

$\pi(1)/\pi(0.8) > v(4000)/v(3000)$ $v(4000)/v(3000) > \pi(0.25)/\pi(0.2)$

$\Rightarrow \pi(1)/\pi(0.8) > \pi(0.25)/\pi(0.2)$

pour $p > q$ et $r \in (0, 1)$, la propriété de subproportionality est la suivante : $\pi(p)/\pi(q) > \pi(rp)/\pi(rq)$.

références

- Gilboa, I. (2009) *Theory of Decision under Uncertainty*, Cambridge: CUP
- Gilboa, I. (2010) *Rational Choice*, Cambridge, Mass: MIT Press {chap. 4 (informel) et App. B4 (formel) }
- Kreps, D. (1988), *Notes on the Theory of Choice*, Westview Press {chap. 1-3 }
- Wakker, P. (2010), *Prospect Theory*, {chap. 1-3 }