Master IAD
Module NI214 – MODE

Transparents de cours

année 2012–2013

Christophe Gonzales

MODE — cours 1 : décision dans le risque et l'incertain

Christophe Gonzales

LIP6 – Université Paris 6, France

Plan du cours 1

1. Les modèles mathématiques décisionnels
2. La décision dans le risque
3. La décision dans l'incertain
4. L'attitude vis à vis du risque

1 Modèles mathématiques décisionnels

Décision sans incertitude

Diverses décisions $\Rightarrow$ préférences sur les décisions : $d \succsim_{\mathcal{D}} d'$

Exemple

deux enveloppes : la 1ère contient 100 €, la 2ème 200 €

- d_1 = «prendre la première enveloppe»
- d_2 = «prendre la deuxième enveloppe»

$\Longrightarrow$ décideur : $d_2 \succsim_{\mathcal{D}} d_1$

Pourquoi ? $\Longrightarrow$ conséquence(d_2) préférée à conséquence(d_1)

Préférences sur les conséquences des décisions

$$d_1 \succsim_{\mathcal{D}} d_2 \Longleftrightarrow x(d_1) \succsim_{\mathcal{X}} x(d_2).$$

Décision avec incertitudes

En pratique : les conséquences sont incertaines

Exemple

enveloppe 2 choisie parmi une pile de 100 enveloppes dont 3 contiennent 1000 € et 97 contiennent 1 €

$$\Longrightarrow \begin{cases} \text{enveloppe 1} = 100\,€ \\ \text{enveloppe 2} = \text{3 chances sur 100 d'avoir 1000 € et} \\ \qquad\qquad\qquad\ \text{97 chances sur 100 d'avoir 1 €} \end{cases}$$

$\Longrightarrow$ décideur : $d_1 \succsim_{\mathcal{D}} d_2$

Pourquoi ? $\Longrightarrow$ trop de risque d'avoir 1 € avec d_2

$\Longrightarrow$ décision = conséquences + incertitudes

Terminologie (1/3)

Exemple

Préparation d'une omelette
Déjà 5 œufs cassés, 6ème œuf intact
3 alternatives :

- casser l'œuf dans l'assiette contenant les cinq œufs ;
- le casser dans une autre assiette pour l'inspecter ;
- ne pas vous servir de cet œuf.

Les conséquences

- conséquences : grande omelette, 5 œufs perdus, assiette inutilement salie, petite omelette
- $\mathcal{X}$: ensemble des conséquences
- Alternatives jugées en fonction de leurs conséquences

Terminologie (2/3)

Exemple

Préparation d'une omelette
Déjà 5 œufs cassés, 6ème œuf intact
3 alternatives :

- casser l'œuf dans l'assiette contenant les cinq œufs ;
- le casser dans une autre assiette pour l'inspecter ;
- ne pas vous servir de cet œuf.

Les incertitudes

- incertitudes : œuf bon, œuf mauvais
- $\mathcal{S}$ = ensemble des états de la nature
  = événements élémentaires
- $\mathcal{A} = \{S \subseteq \mathcal{S}\}$ = ensemble des événements

Terminologie (3/3)

Exemple

Préparation d'une omelette
Déjà 5 œufs cassés, 6ème œuf intact
3 alternatives :

- casser l'œuf dans l'assiette contenant les cinq œufs ;
- le casser dans une autre assiette pour l'inspecter ;
- ne pas vous servir de cet œuf.

Les actes

- acte v = fonction $\mathcal{S} \mapsto \mathcal{X}$
- acte $\equiv$ description d'une alternative / décision
- $\mathcal{V} = \{\text{actes}\}$

$$d_1 \succsim_{\mathcal{D}} d_2 \Longleftrightarrow v(d_1) \succsim_{\mathcal{V}} v(d_2)$$

Contextes de prise de décision

Décision dans le certain

- $\forall$ état de la nature, un acte $\Rightarrow$ toujours la même conséquence
- décision dans le certain : les préférences sur les actes correspondent aux préférences sur les conséquences

Décision dans le risque (von Neumann-Morgenstern 1944)

- alternative $\Rightarrow$ peut avoir plusieurs conséquences suivant la réalisation d'un événement ou d'un autre
- on suppose qu'il existe une distribution ≪objective≫ de probabilité connue sur $(\mathcal{S}, \mathcal{A})$

Décision dans l'incertain

- alternative $\Rightarrow$ peut avoir plusieurs conséquences suivant la réalisation d'un événement ou d'un autre
- on ne suppose pas l'existence d'une distribution de probabilité sur $(\mathcal{S}, \mathcal{A})$

Décision dans le certain

Fonction d'utilité (sur les décisions)

fonction $U : \mathcal{D} \mapsto \mathbb{R}$ telle que :

$$d_1 \succsim_{\mathcal{D}} d_2 \Longleftrightarrow U(d_1) \geq U(d_2).$$

Fonction d'utilité (sur les conséquences)

fonction $f : \mathcal{X} \mapsto \mathbb{R}$ telle que :

$$x_1 \succsim_{\mathcal{X}} x_2 \Longleftrightarrow f(x_1) \geq f(x_2).$$

- *Avantages* :
  - problème décisionnel = problème d'optimisation $\Longrightarrow$ utilisation d'un solveur
  - préférences transitives $\Longrightarrow$ optimum facile à justifier
  - souvent, justifications axiomatiques simples
- *Inconvénients* :
  - nécessité d'avoir des préférences transitives
  - élicitation des préférences

Décision dans le risque ou l'incertain (à la Savage)

Le modèle d'espérance d'utilité

- Incertitudes $=$ distribution de probabilité P sur $\mathcal{A}$
- $U : \mathcal{D} \mapsto \mathbb{R}$: Fonction d'utilité sur les décisions
- actes constants : $\forall a \in \mathcal{A} \mapsto$ le même $x \in \mathcal{X}$
- actes constants $\Longrightarrow f : \mathcal{X} \mapsto \mathbb{R}$ fonction d'utilité sur les conséquences

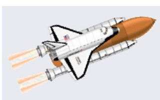
Modèle d'espérance d'utilité :

$$U(d) = \sum_{x(d)} P(x(d)) f(x(d)).$$

Exemple des enveloppes :

$\Longrightarrow U(\text{env } 1) = f(100\,€)$ et $U(\text{env } 2) = \frac{3}{100} f(1000\,€) + \frac{97}{100} f(1\,€)$

Quelques applications de ces modèles décisionnels

-  Arrimage de la navette spatiale
-  décision médicale
-  gestion de l'énergie ou de ressources critiques
-  domaine militaire
-  détermination de prix d'assurances

Quelques questions sur le modèle EU

- Les probas représentent-elles bien les incertitudes ?
- Raisonne-t-on avec des probas ?
- D'où sortent les probas ?
- Les préférences sur $\mathcal{X}$ doivent-elles être représentées par des fonctions d'utilité ?
- Est-ce que EU est un bon critère de décision ?

Réponse : ça dépend

Cours de MODE $\Longrightarrow$ voir dans quels cas EU est un modèle raisonnable

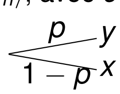
2 Décision dans le risque

Axiomatique de von Neumann-Morgenstern (1/5)

- $\mathcal{V}$: ensemble des actes = applications de $\mathcal{S}$ dans $\mathcal{X}$
- $\succsim_{\mathcal{D}}$: relation de préférence sur $\mathcal{V}$
- $\succ_{\mathcal{D}}$: relation de préférence stricte, $\sim_{\mathcal{D}}$: indifférence
- $\succsim_{\mathcal{D}} \Longrightarrow \succsim_{\mathcal{X}}$ sur $\mathcal{X}$ grâce aux actes constants

hypothèse du risque (incertain probabilisé) : $\exists P : 2^{\mathcal{S}} \mapsto \mathbb{R}$

- $\mathcal{A} = 2^{\mathcal{S}}$, $(\mathcal{S}, \mathcal{A}, P)$: espace probabilisé
- $f \in \mathcal{V} \Longrightarrow P_f$ sur $(\mathcal{X}, 2^{\mathcal{X}})$
- $\mathcal{L}$: loteries, ensemble des lois à support fini sur $\mathcal{X}$
  $f \Longrightarrow P_f = \langle c_1, p_1; \ldots; c_n, p_n \rangle$, avec $c_1 \precsim_{\mathcal{X}} c_2 \precsim_{\mathcal{X}} \cdots \precsim_{\mathcal{X}} c_n$



- $\succsim_{\mathcal{D}} \Longrightarrow \succsim$ sur $\mathcal{L}$

Axiomatique de von Neumann-Morgenstern (2/5)

Théorème de von Neumann-Morgenstern (44)

représentation de $\succsim$ sur $\mathcal{L}$ par EU

Rappel sur la terminologie (1/3)

Exemple de Savage

Préparation d'une omelette
Déjà 5 œufs cassés, 6ème œuf intact
3 alternatives/décisions :

- d_1 : casser l'œuf dans l'assiette contenant les cinq œufs ;
- d_2 : le casser dans une autre assiette pour l'inspecter ;
- d_3 : ne pas vous servir de cet œuf.

- $\mathcal{S}$: ensemble des états de la nature $=$ {œuf (B)on,œuf (M)auvais}
- $\mathcal{A}$: ensemble des événements $= \{\emptyset, \{B\}, \{M\}, \{B, M\}\}$
- $\mathcal{X}$: ensemble des conséquences $=$ {grande omelette, 5 œufs perdus, assiette salie, petite omelette}
- $\mathcal{D}$: ensemble des décisions $= \{d_1, d_2, d_3\}$
- $\mathcal{V}$: ensemble des actes = {fonctions $\mathcal{S} \mapsto \mathcal{X}$}

Rappel sur la terminologie (2/3)

Exemple de Savage

Préparation d'une omelette
Déjà 5 œufs cassés, 6ème œuf intact
3 alternatives/décisions :

- d_1 : casser l'œuf dans l'assiette contenant les cinq œufs ;
- d_2 : le casser dans une autre assiette pour l'inspecter ;
- d_3 : ne pas vous servir de cet œuf.

- décision d_1
- acte représentant la décision $d_1 = A_{d_1} : \mathcal{S} \mapsto \mathcal{X}$
  A_{d_1}(œuf bon) $=$ grand omelette
  A_{d_1}(œuf mauvais) $=$ 5 œufs perdus
- Si $P(\text{bon}) = 0.3$ et $P(\text{mauvais}) = 0.7$
  Acte $A_{d_1} \Longrightarrow P(\text{gde omelette}) = 0.3$ et $P(\text{œufs perdus}) = 0.7$

Rappel sur la terminologie (3/3)

Exemple de Savage

Préparation d'une omelette
Déjà 5 œufs cassés, 6ème œuf intact
3 alternatives/décisions :

- d_1 : casser l'œuf dans l'assiette contenant les cinq œufs ;
- d_2 : le casser dans une autre assiette pour l'inspecter ;
- d_3 : ne pas vous servir de cet œuf.

- $\mathcal{L}$: ensemble des loteries $= \{$lois de proba sur $\mathcal{X}\}$
- Si $P(\text{bon}) = 0.3$ et $P(\text{mauvais}) = 0.7$

  loterie de $d_1 \Longrightarrow P(\text{grande omelette}) = 0.3$
  $P(\text{œufs perdus}) = 0.7$

Décisions, actes et loteries

Actes et loteries

- Acte : résume une décision en ne tenant compte que des couples (états de la nature,conséquences)
- Loterie : résume un acte en ne tenant compte que des probabilités des conséquences

événement	proba	acte $A1$	acte $A2$
e_1	0.3	cons. $C1$	cons. $C2$
e_2	0.3	cons. $C2$	cons. $C1$
e_3	0.4	cons. $C2$	cons. $C2$

2 actes, 1 seule loterie : $L = \{P(C1) = 0.3\ ;\ P(C2) = 0.7\}$

$\Longrightarrow$ info décision $\geq$ info acte $\geq$ info loterie

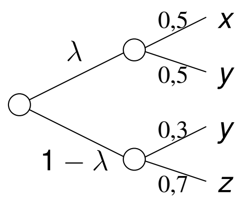
Axiomatique de von Neumann-Morgenstern (3/5)

Théorème de von Neumann-Morgenstern (44)

représentation de $\succsim$ sur $\mathcal{L}$ par EU

Mixage de lois

$\forall P, Q \in \mathcal{L}, \forall \lambda \in [0,1], \lambda P + (1-\lambda) Q \in \mathcal{L}$: loterie



$\langle x, 0.5\lambda; y, (0.5\lambda + 0.3 \times (1-\lambda)); z, 0.7 \times (1-\lambda) \rangle$

Axiomatique de von Neumann-Morgenstern (4/5)

Axiome 1 : préordre total

$\succsim$ est un préordre total (réflexif, transitif, complet) sur $\mathcal{L}$ non trivial ($\exists P, Q \in \mathcal{L}$ t.q. $P \succ Q$).

Axiome 2 : continuité

$\forall P, Q, R \in \mathcal{L}$ tels que $P \succ Q \succ R$, $\exists \alpha, \beta \in]0,1[$ tels que :

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

Axiome 3 : indépendance

$\forall P, Q, R \in \mathcal{L}$, $\forall \alpha \in]0,1]$:

$$P \succsim Q \Longleftrightarrow \alpha P + (1-\alpha)R \succsim \alpha Q + (1-\alpha)R.$$

Axiomatique de von Neumann-Morgenstern (5/5)

Théorème de von Neumann-Morgenstern

Les deux assertions suivantes sont équivalentes :

1. $\succsim$ sur $\mathcal{L}$ vérifie les axiomes 1,2,3.
2. $\succsim$ est représentable par U t.q.

$$U(P) = \sum_{i=1}^{n} p_i u(x_i),$$

où $u(x_i) = U(\langle x_i, 1\rangle)$.

- $u : \mathcal{X} \mapsto \mathbb{R}$ utilité de von Neumann-Morgenstern
- u unique à une transformation affine strictement positive près

3 Décision dans l'incertain

Critère de Wald

Définition

- Choisir l'acte dont la pire conséquence est la meilleure :

$$\text{Argmax}_{f\in\mathcal{V}} \min_{s\in\mathcal{S}} u(f(s)).$$

- idée : principe de prudence.

	s_1	s_2	s_3	min
f_1	20	10	-30	-30
f_2	-10	30	10	-10
f_3	10	20	-5	-5

	s_1	s_2	s_3	min
f_1	20	10	-10	-10
f_2	200	300	-11	-11

Critère d'Hurwicz

Définition

- Choisir l'acte avec le meilleur compromis entre meilleure et pire conséquence :

$$\text{Argmax}_{f\in\mathcal{V}} \left[\alpha \min_{s\in\mathcal{S}} u(f(s)) + (1-\alpha) \max_{s\in\mathcal{S}} u(f(s))\right],$$

avec $\alpha \in [0,1]$
- idée : compromis entre prudence et optimisme.

	s_1	s_2	s_3	s_4	Hurwicz
f_1	200	0	0	-10	$\alpha \times (-10) + (1-\alpha) \times 200$
f_2	200	200	200	-11	$\alpha \times (-11) + (1-\alpha) \times 200$

Min Max Regret (1/2)

Définition

- Choisir l'acte dont on regrettera le moins la conséquence :

$$\text{Argmin}_{f\in\mathcal{V}} \max_{s\in\mathcal{S}} R(f,s),$$

avec $R(f,s) = \max_{g\in\mathcal{V}} u(g(s)) - u(f(s))$.

	s_1	s_2	s_3
f_1	20	10	-30
f_2	-10	30	10
f_3	10	20	-5

	s_1	s_2	s_3	max
R_1	0	20	40	40
R_2	30	0	0	30
R_3	10	10	15	15

Min Max Regret (2/2)

	s_1	s_2	s_3
f_1	25	0	10
f_2	9	30	30
f_3	10	15	15

	s_1	s_2	s_3	max
R_1	0	30	20	30
R_2	16	0	0	16
R_3	15	15	15	15

	s_1	s_2
f_1	8	0
f_2	2	4
f_3	1	7

	s_1	s_2	max
R_1	0	7	7
R_2	6	3	6
R_3	7	0	7

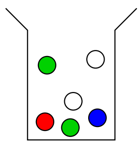
Critère de Laplace

Définition

- Choisir l'acte ayant la conséquence moyenne la plus élevée :

$$\mathrm{Argmax}_{f \in \mathcal{V}} \sum_{s \in \mathcal{S}} \frac{1}{|\mathcal{S}|} u(f(s)).$$

- idée : les événements sont équiprobables.



	R	V	B	$\sum$
f_1	100	0	0	$33,3$
f_2	0	99	99	$66,0$

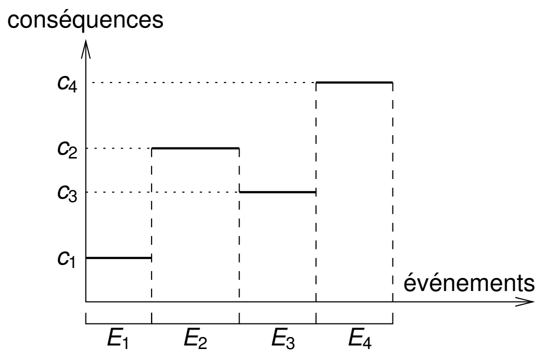
	R	$\neg R$	$\sum$
f_1	100	0	$50,0$
f_2	0	99	$49,5$

Axiomatique de Savage (1/9)

acte : fonction de $\mathcal{S} \mapsto \mathcal{X}$

Acte simple en escalier

f est un acte simple en escalier s'il existe une partition finie $\{E_i, i \in I\}$ de $\mathcal{S}$, telle que $f(E_i) = \{c_i\}$.



Axiomatique de Savage (2/9)

Acte simple en escalier

f est un acte simple en escalier s'il existe une partition finie $\{E_i, i \in I\}$ de $\mathcal{S}$, telle que $f(E_i) = \{c_i\}$.

Acte en escalier

f est un acte en escalier s'il existe une partition dénombrable $\{E_i, i \in I\}$ de $\mathcal{S}$, telle que $f(E_i) = \{c_i\}$.

Acte constant

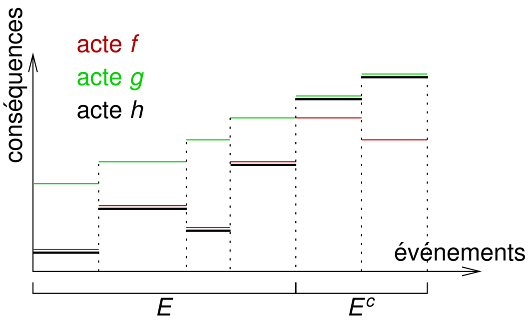
δ_c est un acte constant si $\delta_c(\mathcal{S}) = \{c\}$.

Axiomatique de Savage (3/9)

Greffe

Soit f, g deux actes. Soit $E \subseteq \mathcal{S}$ un événement.

$$\text{acte } h = fEg = \begin{cases} h(s) = f(s) \text{ pour } s \in E, \\ h(s) = g(s) \text{ pour } s \in E^c. \end{cases}$$



Axiomatique de Savage (4/9)

Axiome P1 : préordre total sur les actes

1. $\mathcal{A}$, l'ensemble des événements, est une σ-algèbre.
2. L'ensemble des actes contient l'ensemble des actes en escalier et est fermé pour l'opération de greffe.
3. $\succsim$ est un préordre total.

Axiomatique de Savage (5/9)

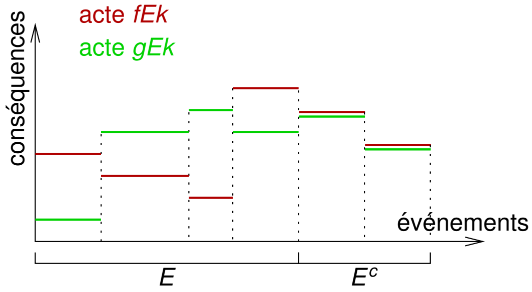
Axiome clé de Savage :

Axiome P2 : Sure thing principle

$\forall$ actes $f, g, h, k \in \mathcal{X}^{\mathcal{S}}$ et $\forall E \subseteq \mathcal{S}$:

$$fEh \succsim gEh \Longleftrightarrow fEk \succsim gEk.$$

Une modification commune d'une partie commune à deux actes ne modifie pas les préférences entre ces actes.



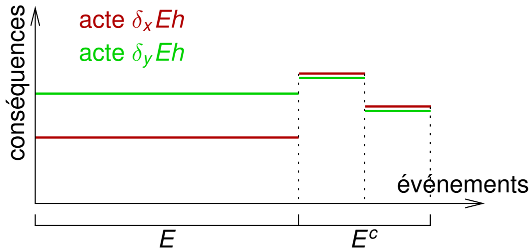
Axiomatique de Savage (6/9)

P2 : $f \succsim_E g \Longleftrightarrow$ pour tout h, $fEh \succsim gEh$

E non négligeable si $\exists f, g$ t.q. $f \succ_E g$

Axiome P3 : existence de préférences dans le certain

$\forall x, y \in \mathcal{X}$, $\forall E \subseteq \mathcal{S}$ non négligeable, $\delta_x \succsim_E \delta_y \Longleftrightarrow x \succsim_{\mathcal{X}} y$, où δ_x, δ_y sont des actes constants.



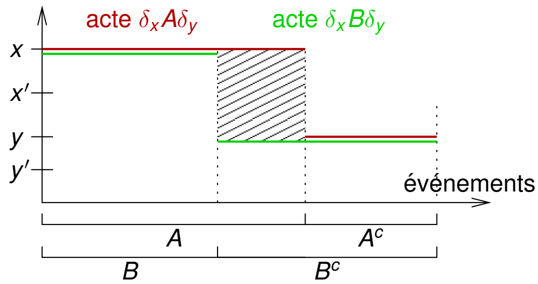
$\Longrightarrow$ un traumatisme émotionnel ne peut bouleverser la hiérarchie des valeurs

Axiomatique de Savage (7/9)

Axiome P4 : préférences sur les événements

$\forall x, x', y, y' \in \mathcal{X}$ tels que $x \succ_{\mathcal{X}} y$ et $x' \succ_{\mathcal{X}} y'$, et $\forall A, B \subseteq \mathcal{S}$,

$$\delta_x A \delta_y \succsim \delta_x B \delta_y \Longleftrightarrow \delta_{x'} A \delta_{y'} \succsim \delta_{x'} B \delta_{y'}.$$



interprétation : puisque ça ne dépend pas des conséquences, c'est que l'on pense que A a plus de chances de se produire que B.

Axiomatique de Savage (8/9)

Axiome P5 : non trivialité des préférences dans le certain

Il existe $x, y \in \mathcal{X}$ tels que $\delta_x \succ \delta_y$.

Axiome P6 : continuité

$\forall$ actes $f, g \in \mathcal{X}^{\mathcal{S}}$ tels que $f \succ g$, $\forall x \in \mathcal{X}$,
$\exists E = \bigcup_{i=1}^{n} E_i$, $E_i \subseteq \mathcal{S}$, tel que :

$$\forall i \in \{1, \ldots, n\}, \quad \delta_x E_i f \succ g \text{ et } f \succ \delta_x E_i g.$$

interprétation : chaque E_i est jugé suffisamment peu probable pour que la modification de f sur E_i ne renverse pas les préférences
$\Longrightarrow$ les $\delta_x E_i f$ sont très «proches» les uns des autres

Axiomatique de Savage (9/9)

Théorème de Savage (54)

Si P1 à P6 sont vérifiés, $\succsim$ est représentable par EU :

$$U(f) = \sum_{s \in \mathcal{S}} p(s) u(f(s)).$$

Remarque : on n'a jamais supposé l'existence d'une loi de proba

$\Longrightarrow$ elle découle des axiomes

$\Longrightarrow$ probabilités subjectives

$\Longrightarrow$ Subjective Expected Utility (SEU)

4 Attitude vis à vis du risque

La notion d'équivalent certain

Définition pour les loteries

- Soit L une loterie quelconque $\Longrightarrow U(L) =$ l'utilité de L
- Notation : $\texttt{CE}_L =$ équivalent certain de L
- $\texttt{CE}_L =$ la loterie $\langle x, 1\rangle$ telle que $U(\texttt{CE}_L) = U(L)$

Définition pour les actes

- Soit f un acte quelconque $\Longrightarrow U(f) =$ l'utilité de f
- Notation : $\texttt{CE}_f =$ équivalent certain de f
- $\texttt{CE}_f =$ l'acte constant $\delta_{U(f)}$

Prix de vente d'une loterie (1/3)

- Loterie $X = \langle x_1, p_1; \ldots; x_n, p_n\rangle$

  $=$ une distribution de probabilité sur $\mathcal{X}$

  $\equiv$ variable aléatoire sur $\mathcal{X}$

- Hypothèse pour la suite : $\mathcal{X} = \mathbb{R}$
- Richesse initiale w_0 fixée
- Richesse finale $W_f = \langle w_0 + x_1, p_1; \ldots; w_0 + x_n, p_n\rangle = w_0 + X$ où $X = \langle x_1, p_1; \ldots; x_n, p_n\rangle$

Exemple :

On possède w_0 € et on en investit une partie à la bourse

Prix de vente d'une loterie (2/3)

- Richesse initiale w_0
- Richesse finale $W_f = w_0 + X$, où $X =$ loterie
- Équivalent certain de $W_f = \texttt{CE}_f$
- $\texttt{CE}_f =$ loterie certaine $\sim W_f$

  $\Longrightarrow \texttt{CE}_f - w_0$ revient à se débarrasser de X

Définition

Prix de vente de la loterie X : $p_v = \texttt{CE}_f - w_0$

Prix de vente d'une loterie (3/3)

Interprétation :

- $U(\text{CE}_f) = U(w_0 + X)$

  $\Longrightarrow p_v =$ prix minimum exigé pour vendre la loterie X

Application : souscription d'une assurance $\Longrightarrow p_v < 0$

Neutralité vis-à-vis du risque

Théorème

- Richesse initiale w_0, richesse finale $W_f = w_0 + X$
- Décideur maximisateur d'espérance d'utilité
- Si l'utilité de von Neumann-Morgenstern est linéaire alors :

$$p_v = E(X)$$

$E(X) =$ espérance de X, pas l'espérance d'utilité !

$\Longrightarrow$ 2 loteries de même espérance $\Longrightarrow$ même prix

$\Longrightarrow \langle x_1, \frac{1}{2}; x_2; \frac{1}{2} \rangle \sim \langle \frac{x_1+x_2}{2}, 1 \rangle$

$\Longrightarrow$ le décideur ne prend pas en compte le risque

Neutralité au risque

Un décideur, maximisateur d'EU, est neutre vis-à-vis du risque si son utilité de von Neumann-Morgenstern est linéaire.

Aversion au risque / goût pour le risque

Aversion au risque

Un décideur, maximisateur d'EU, est adversaire du risque si

$$p_v < E(X)$$

Goût pour le risque

Un décideur, maximisateur d'EU, a du goût pour le risque si

$$p_v > E(X)$$

Théorème

Soit une fonction d'utilité de von Neumann-Morgenstern strictement croissante, alors :

- si elle est strictement concave, $p_v < E(X)$
- si elle est strictement convexe, $p_v > E(X)$

Résumé : attitude face au risque

Aversion faible pour le risque – Arrow(65), Pratt(64)

Un agent a de l'aversion faible pour le risque si, pour toute loterie L, il préfère l'acte certain d'espérance $E(L)$ à l'acte L :

$$\forall L \in \mathcal{L},\ \langle E(L), 1\rangle \succsim L.$$

L'agent a du goût faible pour le risque si : $\forall L \in \mathcal{L},\ L \succsim \langle E(L), 1\rangle$.

L'agent est neutre vis à vis du risque si : $\forall L \in \mathcal{L},\ L \sim \langle E(L), 1\rangle$.

$\Longrightarrow$ concavité de u = aversion au risque

Risque : deux questions

concavité de $u \Longleftrightarrow$ aversion faible au risque

$\Longrightarrow$ deux questions :

1. Peut-on caractériser si un décideur a plus ou moins d'aversion au risque ?
2. Peut-on caractériser si une situation est plus ou moins risquée ?

Coefficient d'aversion absolue au risque (1/2)

- Richesse initiale w_0 ; richesse finale $W_f = w_0 + X$
- X : loterie d'espérance μ et de variance σ^2
- Équivalent certain de $W_f = \mathrm{CE}_f$
- Prix de vente de la loterie X : $p_v = \mathrm{CE}_f - w_0$
- u : fonction d'utilité de von Neumann-Morgenstern

$\Longrightarrow U(\mathrm{CE}_f) = U(W_f) = U(w_0 + p_v)$

Formule de Taylor autour de $w_0 + \mu \Longrightarrow p_v - \mu \approx \dfrac{\sigma^2}{2}\dfrac{u''(w_0+\mu)}{u'(w_0+\mu)}$

Coefficient d'aversion absolue au risque (2/2)

$$p_v - \mu \approx \frac{\sigma^2}{2} \frac{u''(w_0 + \mu)}{u'(w_0 + \mu)}$$

- $\frac{\sigma^2}{2}$: indicateur du risque véhiculé par la loterie X
- $\frac{u''}{u'}$: allure de la fonction d'utilité ($\Longrightarrow$ attitude vis-à-vis du risque)

Coeff d'aversion absolue pour le risque – Arrow(65), Pratt(64)

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

Il est généralement admis que $R_A(x)$ décroît avec x

Remarque : plus $R_A(x)$ est élevé, plus on est adversaire du risque $\Longrightarrow$ comparaisons entre agents

Quantité du risque

- Mesure de quantité de risque ?
- Variance ? Arrow-Pratt : $p_v - \mu \approx \frac{\sigma^2}{2} \frac{u''(w_0 + \mu)}{u'(w_0 + \mu)}$
- Rothschild & Stiglitz (1970) :

  La variance n'est pas le meilleur des indicateurs :

Utiliser la notion de ≪Mean Preserving Spread≫

Mean Preserving Spread

x	$P(x)$
-2	0.09
4	0.30
10	0.40
16	0.21

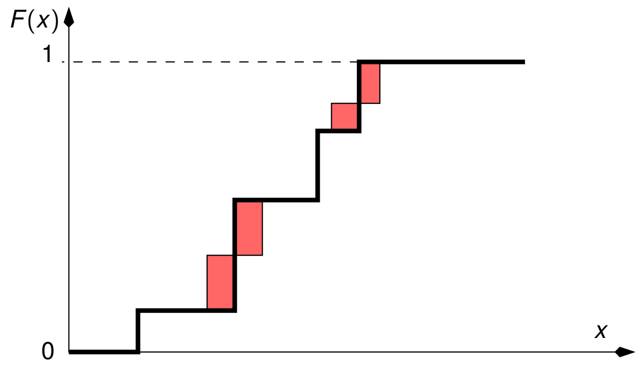
z	$Q(z)$
-2	0.09
3	0.15
5	0.15
10	0.40
12	0.07
18	0.14

- Z a la même espérance que X
- Z déduit de X en remplaçant des valeurs certaines par des loteries $\Longrightarrow$ Z plus risqué que X

$Z = X +$ un bruit blanc = Mean preserving spread

interprétation graphique du MPS

$X \Longrightarrow P(x) \Longrightarrow F(x) = P(z \leq x)$



$$Y = \mathrm{MPS}(X) \Longrightarrow \int_{-\infty}^{T} F_Y(x)dx \geq \int_{-\infty}^{T} F_X(x)dx \text{ pour tout } T$$

Dominance stochastique et Mean Preserving Spread

Définition : dominance stochastique d'ordre 2

- X et Y deux loteries
- X domine stochastiquement Y, noté $X \succsim_{DS2} Y$, si :

$$\int_{-\infty}^{T} F_Y(x)dx \geq \int_{-\infty}^{T} F_X(x)dx \text{ pour tout } T$$

Mean Preserving Spread

- X et Y deux loteries
- on dit que $Y = \mathrm{MPS}(X)$ (autrement dit, Y est un accroissement de risque à moyenne fixée de X) si :
  1. $E(X) = E(Y)$
  1. $X \succsim_{DS2} Y$

Attitude face au risque – Rothschild & Stiglitz (1970)

Aversion forte pour le risque – Rothschild & Stiglitz (1970)

- Un agent a de l'aversion forte pour le risque si : pour toutes loteries X, Y telles que $Y = \mathrm{MPS}(X)$, $X \succsim Y$
- Un agent a un goût fort pour le risque si : pour toutes loteries X, Y telles que $Y = \mathrm{MPS}(X)$, $Y \succsim X$
- Un agent est neutre vis-à-vis du risque si : pour toutes loteries X, Y telles que $Y = \mathrm{MPS}(X)$, $Y \sim X$

$\forall X,\ X = \mathrm{MPS}(E(X)) \Longrightarrow$ aversion forte $\Longrightarrow$ aversion faible

Réciproque ?

Conclusion sur l'attitude face au risque

Proposition de Rothschild & Stiglitz (1970-71)

- agent maximisateur d'espérance d'utilité
- les 3 assertions suivantes sont équivalentes :
    1. l'agent a de l'aversion faible pour le risque
    2. l'agent a de l'aversion forte pour le risque
    3. la fonction d'utilité de von Neumann-Morgenstern de l'agent est concave

Quelques références sur l'axiomatique d'EU

Références

1. J. von Neumann & O. Morgenstern (1947)
   "Theory of Games and Economic Behaviour",
   Princetown University Press
2. L. J. Savage (1954)
   "The Foundations of Statistics", Dover
3. A. Chateauneuf, M. Cohen & J.-M. Tallon (2006)
   "Décision dans le risque : mesure du risque, aversion pour le risque, modèle classique d'utilité espérée, paradoxe d'Allais, modèles à niveaux de sécurité et de potentiel", in *Concepts et méthodes pour l'aide à la décision*, Vol 2, chapitre 1, Hermes
4. A. Chateauneuf, M. Cohen & J.-Y. Jaffray (2006)
   "Décision dans l'incertain : les modèles classiques", in *Concepts et méthodes pour l'aide à la décision*, Vol 2, chapitre 2, Hermes

Quelques références sur l'aversion au risque

Références

5. L. Eeckhoudt & C. Gollier (1992)
   "Les risques financiers – Evaluation, Gestion, Partage",
   Ediscience International
6. J. Pratt (1964)
   "Risk aversion in the small and in the large",
   Econometrica, Vol. 32, pp. 122–136
7. K.J. Arrow (1965)
   "The theory of risk aversion", *in Aspects of the Theory of Risk Bearing*, pp.90–120, Yrjo Jahnsson Fondation, Helsinki.
8. M. Rotschild & J. Stiglitz (1970)
   " Increasing risk I : A definition",
   Journal of Economic Theory, Vol. 2, pp.225–243
9. M. Rotschild & J. Stiglitz (1971)
   " Increasing risk II : Its economic consequences",
   Journal of Economic Theory, Vol. 3, pp.66–84

MODE — cours 2 : modèles non linéaires

Christophe Gonzales

LIP6 – Université Paris 6, France

Plan du cours 2

1. Limites descriptives de EU
2. Prospect Theory
3. Rank Dependent Utility
4. Choquet Expected Utility

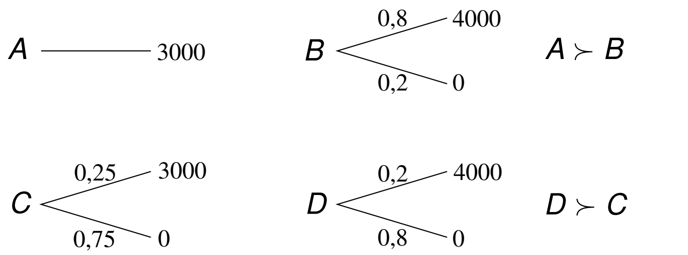
1 Limites descriptives de EU

Limites descriptives de EU (1/4)

- Les axiomes d'EU ne sont pas si ≪naturels≫
- Double interprétation de la fonction d'utilité u :
  - attitude vis à vis du risque (concavité = aversion)
  - préférences dans le certain
    (concavité = utilité marginale décroissante de la richesse)
  - impossible de combiner les 2
    (goût pour le risque + utilités marginales décroissantes)
- commensurabilité des préférences et des incertitudes
- manque de flexibilité pour rendre compte de divers types d'aversion au risque
- dans vNM : les probabilités ≪objectives≫ ne sont pas toujours celles perçues par l'agent

Limites descriptives de EU (2/4)

Kahneman & Tversky :

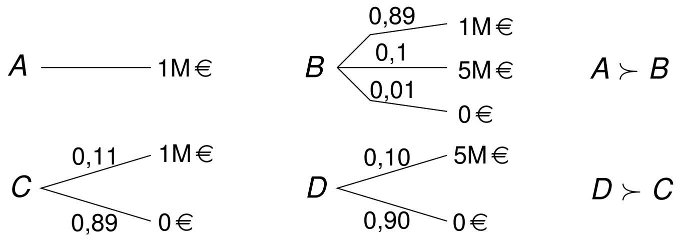


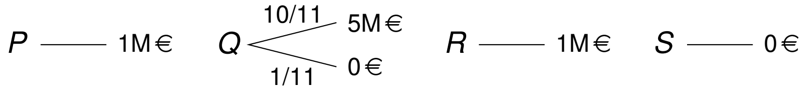
$C = 0,25 \times A + 0,75 \times 0$ $\qquad D = 0,25 \times B + 0,75 \times 0$

violation de l'axiome d'indépendance / effet de certitude

Limites descriptives de EU (3/4)

Le paradoxe d'Allais (53) :



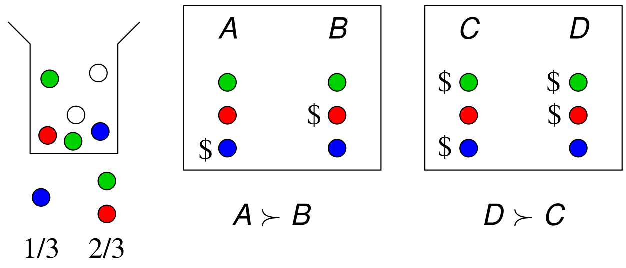


$A = 0,11 \times P + 0,89 \times R$ $\qquad C = 0,11 \times P + 0,89 \times S$
$B = 0,11 \times Q + 0,89 \times R$ $\qquad D = 0,11 \times Q + 0,89 \times S$

violation de l'axiome d'indépendance

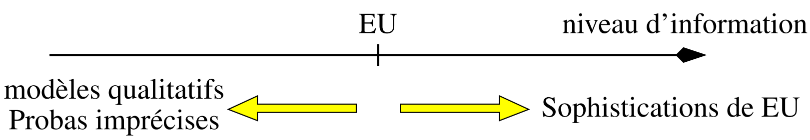
Limites descriptives de EU (3/3)

L'urne d'Ellsberg (1961) :



$\Longrightarrow$ Violation du Sure thing principle

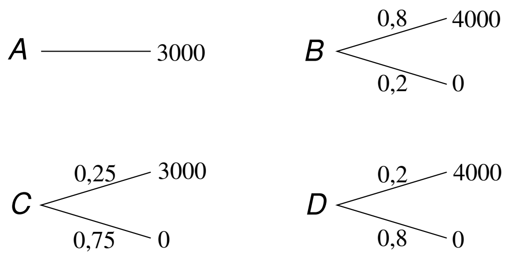
Deux directions...



2 Modèles non linéaires :

La Prospect Theory

Limites descriptives de EU : Kahneman & Tversky



$A \succ B$

$D \succ C$

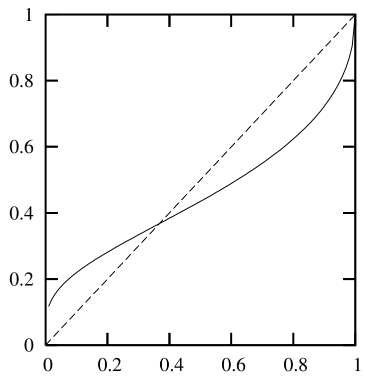
$C = 0,25 \times A + 0,75 \times 0$ $\qquad$ $D = 0,25 \times B + 0,75 \times 0$

$\Longrightarrow$ violation de l'axiome d'indépendance

Déformation des probabilités

Kahneman & Tversky (79)

$\mu(A) = e^{-\sqrt{-\ln(p(A))}}$



Transformation de probabilités

Transformation : idée

Utiliser une capacité non décomposable obtenue à partir d'une déformation des probabilités :

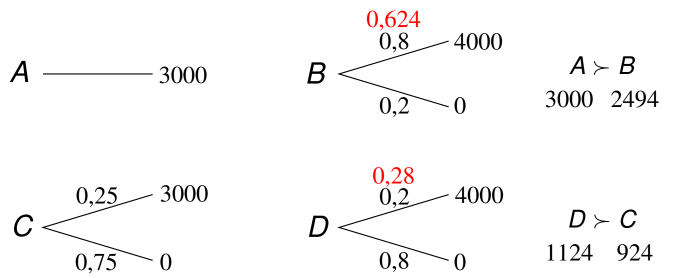
$$\mu(A) = \varphi(P(A)), \text{ où } P(A) = \text{ proba de } A.$$

Prospect Theory – Kahneman & Tversky (79)

- $L = \langle x_1, p_1; \ldots; x_n, p_n \rangle$
- idée : remplacer $EU(L)$ par $PT(L)$:

$$PT(L) = \sum_{i=1}^{n} \varphi(p(x_i))u(x_i)$$

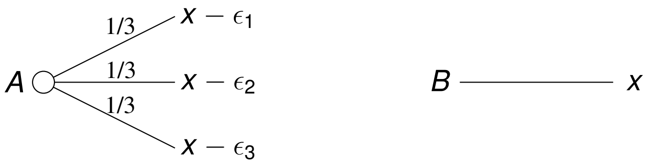
Kahneman & Tversky et le paradoxe d'Allais



Représentable par la Prospect Theory

Problème des déformations de probabilité

Supposons que $\varphi(1/3) > 1/3$



$A \succ B$ ou $B \succ A$?

$\Longrightarrow$ besoin d'un modèle plus sophistiqué : RDU

3 Modèles non linéaires :

Rank Dependent Utility

De EU à RDU

Loterie $L = \langle x_1, p_1; x_2, p_2; x_3, p_3 \rangle$

Critère EU :

Loterie $L \Longrightarrow U(L) = p_1 u(x_1) + p_2 u(x_2) + p_3 u(x_3)$

supposons que $u(x_2) < u(x_1) < u(x_3)$:

$U(L) = (p_1 + p_2 + p_3) u(x_2) + (p_1 + p_3)[u(x_1) - u(x_2)] + p_3[u(x_3) - u(x_1)]$

Critère RDU :

$U(L) = \varphi(p_1 + p_2 + p_3) u(x_2) + \varphi(p_1 + p_3)[u(x_1) - u(x_2)] + \varphi(p_3)[u(x_3) - u(x_1)]$

φ = transformation (de perception) de probabilité

Rank Dependent Utility (RDU)

Définition

- μ : capacité non décomposable obtenue à partir d'une déformation des probabilités :

$$\mu(A) = \varphi(P(A)), \text{ où } P(A) = \text{proba de } A.$$

- x_i triés par utilités croissantes :

$$u(x_1) \leq u(x_2) \leq \cdots \leq u(x_n)$$

- Rank dependent utility :

$$RDU(x) = u(x_1) + \sum_{i=2}^{n} \varphi\left(\sum_{k=i}^{n} p(x_k)\right)(u(x_i) - u(x_{i-1}))$$

Interprétation de RDU

interprétation de RDU

$$RDU(x) = \sum_{i=1}^{n} \varphi\left(\sum_{k=i}^{n} p(x_k)\right)[u(x_i) - u(x_{i-1})]$$

- Le décideur évalue l'utilité minimale $u(x_1)$
- puis pondération des accroissements d'utilité $u(x_i) - u(x_{i-1})$ par une transformation de la proba d'avoir au moins x_i
- $\varphi(p) \leq p \Longrightarrow$ sous-estimation des accroissements
  $\Longrightarrow$ pessimisme dans le risque

Exemple de calcul de RDU

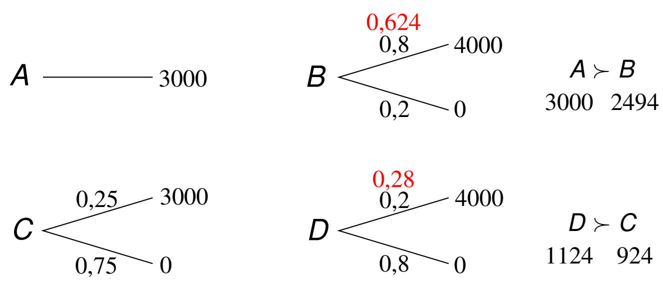
- utilité de von Neumann-Morgenstern $u(x) = x/2$
- transformation de proba $\varphi(x) = x^2$
- loterie $L = \langle 3, 0.2;\ 10, 0.4;\ 5, 0.1;\ 9, 0.3 \rangle$

---

1. trier les conséquences de L par ordre de préférence
   $\Longrightarrow L = \langle 3, 0.2;\ 5, 0.1;\ 9, 0.3;\ 10, 0.4 \rangle$
2. appliquer la formule de RDU :
   $RDU(L) = \varphi(1)\frac{3}{2} + \varphi(0.8)\left[\frac{5}{2} - \frac{3}{2}\right] + \varphi(0.7)\left[\frac{9}{2} - \frac{5}{2}\right] + \varphi(0.4)\left[\frac{10}{2} - \frac{9}{2}\right]$

RDU et le paradoxe d'Allais

RDU avec $\mu(A) = e^{-\sqrt{-\ln(p(A))}}$



Représentable par RDU

Définition précise de RDU

Définition

préférences sur $(\mathcal{L}, \succsim)$ représentables par RDU :

- $u : \mathbb{R} \mapsto \mathbb{R}$, croissante, cardinale, qui représente $\succsim_{\mathcal{X}}$
- $\varphi : [0, 1] \mapsto [0, 1]$ telle que $\varphi(0) = 0$, $\varphi(1) = 1$, croissante
- variables aléatoires X et Y, $X \succsim Y \Leftrightarrow U(X) \geq U(Y)$ où :

$$U(X) = \int_{-\infty}^{0} [\varphi(P(u(X) > t)) - 1]dt + \int_{0}^{\infty} \varphi(P(u(X) > t))dt.$$

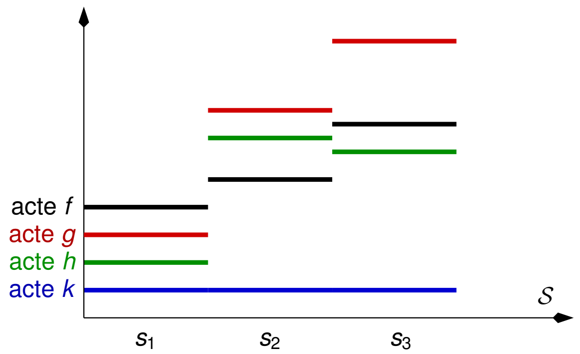
Axiomatique de RDU : la comonotonie (1/2)

Actes comonotones

- $\mathcal{X}$: ensemble des conséquences
- $\mathcal{S}$: ensemble des états de la nature
- acte v = fonction $\mathcal{S} \mapsto \mathcal{X}$
- 2 actes f et g sont comonotones s'il n'existe pas d'événements $s, s' \in \mathcal{S}$ tels que :

$$f(s) \succ_{\mathcal{X}} f(s') \text{ et } g(s) \prec_{\mathcal{X}} g(s')$$

Axiomatique de RDU : la comonotonie (2/2)



- Actes comonotones : (f, g), (g, k), (k, h)
- Actes non comonotones : (g, h)
- Comonotonie non transitive : (g, k) et (k, h) mais pas (g, h)

Caractérisation axiomatique de RDU (1/2)

Principe : Remplacer le ≪Sure Thing Principle≫ par le ≪Comonotonic Sure Thing Principle≫

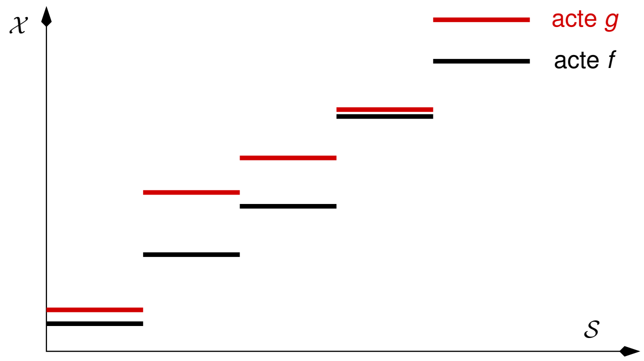
Principe de la chose sûre comonotone

- $\{A_1, \ldots, A_n\}$ partition de $\mathcal{S}$
- acte X : $A_i \mapsto x_i$ avec $x_1 \leq x_2 \leq \cdots \leq x_n$
- acte Y : $A_i \mapsto y_i$ avec $y_1 \leq y_2 \leq \cdots \leq y_n$
- il existe i_0 tel que $x_{i_0} = y_{i_0}$
- 2 actes X' : $A_i \mapsto x'_i$ et Y' : $A_i \mapsto y'_i$ tels que :

$$\begin{cases} x'_{i_0} = y'_{i_0}; \quad x'_i = x_i \text{ et } y'_i = y_i \text{ pour tout } i \neq i_0 \\ x'_1 \leq \cdots \leq x'_n \text{ et } y'_1 \leq \cdots \leq y'_n \end{cases}$$

- Alors : $X \succsim Y \Longrightarrow X' \succsim Y'$

Caractérisation axiomatique de RDU (2/2)



⚠ Remplacer uniquement le ≪Sure Thing Principle≫ ne suffit pas $\Longrightarrow$ il faut ajouter d'autres axiomes

Aversion au risque : rappels (1/2)

Aversion faible pour le risque – Arrow(65), Pratt(64)

Un agent a de l'aversion faible pour le risque si, pour toute loterie L, il préfère l'acte certain d'espérance $E(L)$ à l'acte L :

$$\forall L \in \mathcal{L},\ \langle E(L), 1\rangle \succsim L.$$

L'agent a du goût faible pour le risque si : $\forall L \in \mathcal{L},\ L \succsim \langle E(L), 1\rangle$.

L'agent est neutre vis à vis du risque si : $\forall L \in \mathcal{L},\ L \sim \langle E(L), 1\rangle$.

Aversion au risque : rappels (2/2)

Aversion forte pour le risque – Rothschild & Stiglitz (1970)

- Un agent a de l'aversion forte pour le risque si :
  pour toutes loteries X, Y telles que $Y = \text{MPS}(X)$, $X \succsim Y$
- Un agent a un goût fort pour le risque si :
  pour toutes loteries X, Y telles que $Y = \text{MPS}(X)$, $Y \succsim X$
- Un agent est neutre vis-à-vis du risque si :
  pour toutes loteries X, Y telles que $Y = \text{MPS}(X)$, $Y \sim X$

$\forall X,\ X = \text{MPS}(E(X)) \Longrightarrow$ aversion forte $\Longrightarrow$ aversion faible

Rappel : l'attitude face au risque avec EU

Proposition de Rothschild & Stiglitz (1970-71)

- agent maximisateur d'espérance d'utilité
- les 3 assertions suivantes sont équivalentes :
  1. l'agent a de l'aversion faible pour le risque
  2. l'agent a de l'aversion forte pour le risque
  3. la fonction d'utilité de von Neumann-Morgenstern de l'agent est concave

Problème : Double interprétation de la fonction d'utilité u :

- attitude vis à vis du risque (concavité = aversion)
- préférences dans le certain (concavité = utilité marginale décroissante de la richesse)

Pessimisme et RDU

$$RDU(x) = \sum_{i=1}^{n} \varphi\left(\sum_{k=i}^{n} p(x_k)\right) [u(x_i) - u(x_{i-1})]$$

$$\Longrightarrow \begin{cases} \text{transformation } \varphi \Longrightarrow \text{préférences sur les probas} \\ \text{fonction d'utilité } u \Longrightarrow \text{préférences sur les conséquences} \end{cases}$$

$\varphi(p) \leq p \Longrightarrow$ sous-estime les accroissements d'utilité par rapport à EU

Pessimisme avec RDU

Un agent tel que $\varphi(p) \leq p$ pour tout $p \in [0, 1]$ est dit *pessimiste dans le risque*

Aversion forte pour le risque avec RDU

Proposition de Chew, Karni, Safra (87)

- agent RDU
- les fonctions φ et u sont dérivables
- les 2 assertions suivantes sont équivalentes :
  1. l'agent a de l'aversion forte pour le risque
  2. la fonction d'utilité u est concave ET la fonction de perception des probas φ est convexe

Accroissements de risque

Rappel : accroissement de risque à moyenne constante (MPS)

- X et Y variables aléatoires
- $Y = \text{MPS}(X)$ si $Y = X + Z$, où Z est un bruit blanc

Accroissement monotone de risque à moyenne constante

- X et Y variables aléatoires
- Y accroissement monotone de risque à moyenne constante de X si :
  1. $Y = X + Z$, où Z est un bruit blanc
  2. X et Z sont comonotones

Aversion monotone pour le risque

Un agent a de l'aversion monotone pour le risque s'il n'aime pas l'accroissement monotone pour le risque.

Aversion monotone pour le risque avec RDU

Proposition – Chateauneuf, Cohen, Meilijson (05)

- l'agent a des préférences RDU
- les fonctions u et φ sont dérivables
- indice de pessimisme : $P_\varphi = \inf_{0<p<1} \left[\frac{1-\varphi(p)}{\varphi(p)} / \frac{1-p}{p} \right]$
- indice de non concavité : $G_u = \sup_{y \leq x} \frac{u'(x)}{u'(y)}$
- l'agent a de l'aversion monotone pour le risque si et seulement si $P_\varphi \geq G_u$

Remarques : $G_u \geq 1$; $P_\varphi \geq 1$ si et seulement si $\varphi(p) \leq p$

$\Longrightarrow$ aversion monotone pour le risque $\Longrightarrow \varphi(p) \leq p$

⚠ aversion monotone pour le risque $\not\Longleftarrow$ u concave

Aversion faible pour le risque avec RDU

Pas de caractérisation de l'aversion faible pour le risque

Il y a seulement des conditions suffisantes (e.g., Chateauneuf & Cohen (94))

⚠ la concavité de u n'est forcément pas nécessaire !

Conclusions sur l'aversion au risque avec RDU

- Aversion forte $\Longrightarrow$ aversion monotone $\Longrightarrow$ aversion faible
- En général, les implications inverses sont fausses

4 Modèles non linéaires :

Choquet Expected Utility

La notion de capacité

Définition d'une capacité

- fonction μ des sous-ensembles de $\mathcal{S}$ dans $[0,1]$
- $\mu(\emptyset) = 0, \quad \mu(\mathcal{S}) = 1$
- monotone : $\forall A \subset B \subseteq \mathcal{S}$, $\mu(A) \leq \mu(B)$

Choquet expected utility (1/2)

Définition de l'intégrale de Choquet

- $\mathcal{A}$: sous-ensembles de $\mathcal{S}$
- μ : capacité de $\mathcal{A}$ dans $[0,1]$
- X : fonction (mesurable) de $\mathcal{S}$ dans $\mathbb{R}$
- intégrale de Choquet :

$$\int_{Ch} X d\mu = \int_{-\infty}^{0} [\mu(X > t) - 1] dt + \int_{0}^{\infty} \mu(X > t) dt$$

- μ = proba $P \Longrightarrow \int_{Ch} X dP$ = espérance de X
- X prend un nb fini de valeurs $x_1 \leq x_2 \leq \cdots \leq x_n$:

$$\int_{Ch} X d\mu = x_1 + \sum_{i=2}^{n} (x_i - x_{i-1}) \mu(X \geq x_i)$$

Choquet expected utility (2/2)

Choquet Expected Utility (CEU)

- $\mathcal{X}$: ensemble de conséquences
- $\mathcal{S}$: ensemble des états de la nature
- f : acte : application de $\mathcal{S}$ dans $\mathcal{X}$
- u : application croissante de $\mathcal{X}$ dans $\mathbb{R}$
- $X = u \circ f$: fonction de $\mathcal{S}$ dans $\mathbb{R}$
- espérance d'utilité à la Choquet de l'acte f :

$$
\begin{aligned}
CEU(f) &= \int_{Ch} u(f) d\mu \\
&= \int_{-\infty}^{0} [\mu(u(f) > t) - 1] dt + \int_{0}^{\infty} \mu(u(f) > t) dt
\end{aligned}
$$

Choquet et RDU

Choquet Expected Utility

- μ : capacité sur les sous-ensembles de $\mathcal{S}$
- espérance d'utilité à la Choquet de l'acte f :

$$CEU(f) = \int_{-\infty}^{0} [\mu(u(f) > t) - 1] dt + \int_{0}^{\infty} \mu(u(f) > t) dt$$

Rank Dependent Utility

- $\varphi : [0,1] \mapsto [0,1]$ telle que $\varphi(0) = 0$, $\varphi(1) = 1$, croissante
- X variable aléatoire

$$RDU(X) = \int_{-\infty}^{0} [\varphi(P(u(X) > t)) - 1] dt + \int_{0}^{\infty} \varphi(P(u(X) > t)) dt$$

$\Longrightarrow \mu = \varphi \circ P$

Quelques références (1/2)

Références

1. M. Allais (1953)
"Le Comportement de l'Homme Rationnel devant le Risque : Critique des Postulats et Axiomes de l'Ecole Américaine", *Econometrica*, Vol 21, pp.503-546
2. D. Ellsberg (1961)
"Risk, Ambiguity and the Savage Axioms", *The Quarterly Journal of Economics*, Vol. 75, pp.643–669
3. D. Kahneman & A. Tversky (1972)
"Subjective Probability : A Judgment of Representativeness", Cognitive Psychology, Vol 3, pp. 430-454
4. D. Kahneman & A. Tversky (1979)
"Prospect Theory : an Analysis of Decision under Risk", *Econometrica*, Vol. 47, pp.263–291
5. J. Quiggin (1993)
"Generalized Expected Utility Theory : The Rank-dependent Model", Springer

Quelques références (2/2)

Références

6. G. Choquet (1953)
"Théorie des capacités",
Annales de l'Institut Fourier (Grenoble), vol. V, pp.131–295
7. A. Chateauneuf (1994)
"Modeling attitudes towards uncertainty and risk through the use of Choquet integral", *Annals of Operations Research*, Vol. 52, pp.3–20
8. S. Chew & P.P. Wakker (1996)
"The Comonotonic Sure Thing Principle", *Journal of Risk and Uncertainty*, Vol. 12, pp.5–27
9. S. Chew, E. Karni & Z. Safra (1987)
"Risk aversion in the Theory of Expected Utility with Rank Dependent Preferences", *Journal of Economic Theory*, Vol. 42, pp.370–381

MODE — cours 3:
fonctions de croyance et décisions séquentielles

Christophe Gonzales

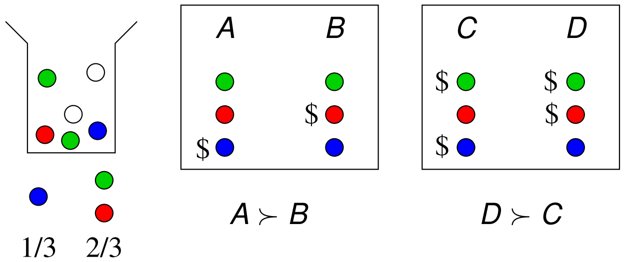
LIP6 – Université Paris 6, France

Plan du cours 3

1. Fonctions de croyance
2. Modèles décisionnels avec fonctions de croyance
3. Décisions séquentielles avec EU
4. Décisions séquentielles avec RDU

1 Les fonctions de croyance

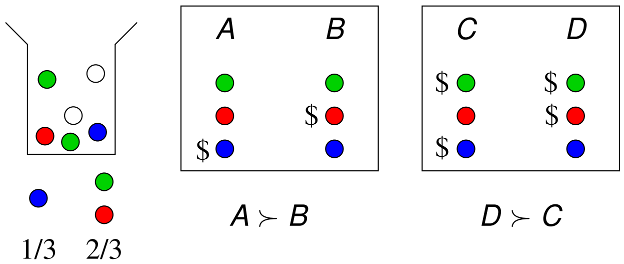
Retour à l'urne d'Ellsberg



Violation du Sure thing principle

Peut-on représenter $\succ$ par RDU ?

L'urne d'Ellsberg et RDU



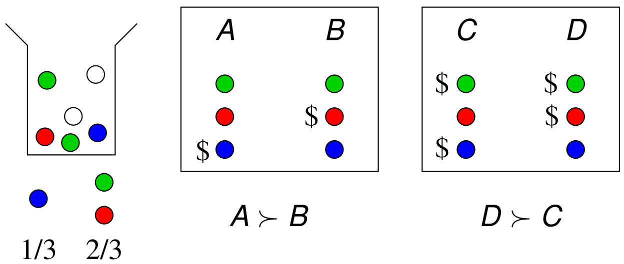
Probas P_r, P_b, P_v

$A \succ B \Longleftrightarrow \varphi(P_b) > \varphi(P_r)$

$D \succ C \Longleftrightarrow \varphi(P_r + P_v) > \varphi(P_b + P_v)$

or φ croissante $\Longrightarrow$ RDU impossible

L'urne d'Ellsberg et les capacités



Problème : $\varphi(P_b) > \varphi(P_r)$ et $\varphi(P_r + P_v) > \varphi(P_b + P_v)$

Définition d'une capacité

- fonction μ des sous-ensembles de $\mathcal{S}$ dans $[0, 1]$
- $\mu(\emptyset) = 0, \quad \mu(\mathcal{S}) = 1$
- monotone : $\forall A \subset B \subseteq \mathcal{S}, \mu(A) \leq \mu(B)$

Les fonctions de croyance

L'urne d'Ellsberg $\Longrightarrow$ capacités à la Choquet

capacités $\stackrel{?}{\Longrightarrow}$ représentation des incertitudes

Théorie des fonctions de croyance

Les fonctions de croyance

Fonction de croyance (Dempster-Shafer)

- $\mathcal{S}$: ensemble des états de la nature
- $\mathcal{A}$: ensemble des sous-ensembles de $\mathcal{S}$
- $\mu : \mathcal{A} \mapsto [0,1]$ fonction de croyance :
  - $\mu(\emptyset) = 0, \qquad \mu(\mathcal{S}) = 1$
  - μ monotone d'ordre ∞ : $\forall n \geq 2$, $\forall A_1, \ldots, A_n \in \mathcal{A}$:

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) \geq \sum_{\emptyset \subset I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \mu\left(\bigcap_{i \in I} A_i\right)$$

Ordre 1 : $\forall A \subseteq B \subseteq \mathcal{S}$, $\mu(A) \leq \mu(B)$

$\Longrightarrow$ capacités à la Choquet

Ordre 2 : $\mu(A \cup B) \geq \mu(A) + \mu(B) - \mu(A \cap B)$

L'inverse de Möbius

Définition de l'inverse de Möbius

Fonction $\phi : \mathcal{A} \mapsto [0,1]$ définie par :

$$\phi(A) = \sum_{B \subseteq A} (-1)^{|A \backslash B|} \mu(B)$$

Expression de la fonction de croyance

$$\mu(A) = \sum_{B \subseteq A} \phi(B)$$

Proposition

μ définie à partir de ϕ est une fonction de croyance ssi :

- $\phi(\emptyset) = 0$
- $\sum_{B \subseteq \mathcal{S}} \phi(B) = 1$
- $\forall B \subseteq \mathcal{S}$, $\phi(B) \geq 0$

Exemple : le camionneur

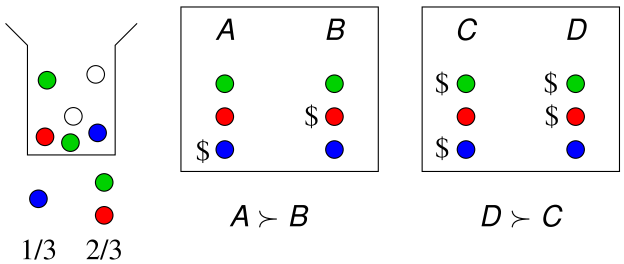
pièces de rechange : $\begin{cases} \text{G(ood)} \\ \text{A(cceptable)} \\ \text{B(ad)} \end{cases}$

Information : au plus la moitié des pièces est d'un type donné

A	$\emptyset$	$\{G\}$	$\{A\}$	$\{B\}$	$\{G,A\}$	$\{G,B\}$	$\{A,B\}$	$\mathcal{S}$
μ	0	0	0	0	$1/2$	$1/2$	$1/2$	1
ϕ	0	0	0	0	$1/2$	$1/2$	$1/2$	$-1/2$

$\Longrightarrow$ μ n'est pas une fonction de croyance

Exemple : l'urne d'Ellsberg



Evt	$\emptyset$	$\{B\}$	$\{R\}$	$\{V\}$	$\{B,R\}$	$\{B,V\}$	$\{R,V\}$	$\mathcal{S}$
μ	0	$1/3$	0	0	$1/3$	$1/3$	$2/3$	1
ϕ	0	$1/3$	0	0	0	0	$2/3$	0

Propriétés et vocabulaire

- $\phi \Longrightarrow$ masses de Möbius
- $\phi(B) > 0$: éléments focaux
- éléments focaux = singletons $\Longleftrightarrow$ μ = proba
- fonction de croyance = enveloppe inférieure de lois de proba

② Modèles décisionnels

avec fonctions de croyance

Décisions avec des fonctions de croyance

- $\mathcal{S}$: ensemble des états de la nature
- $\mathcal{A}$: ensemble des sous-ensembles de $\mathcal{S}$
- $f : \mathcal{A} \mapsto [0, 1]$ fonction de croyance (d'inverse de Möbius ϕ)
- $\mathcal{X}$: ensemble des conséquences
- $\mathcal{Y}$: ensemble des sous-ensembles de $\mathcal{X}$
- décision $\delta \equiv$ acte = fonction $\mathcal{S} \mapsto \mathcal{X}$

$\Longrightarrow \forall Y \in \mathcal{Y},\ g(Y) = f(\delta^{-1}(Y)) =$ incertitude sur l'obtention de Y en prenant la décision δ

g est une fonction de croyance

$\psi : \mathcal{Y} \mapsto [0, 1]$ définie par $\psi(Y) = \sum_{\delta(B)=Y} \phi(B)$ est l'inverse de Möbius de g

Généralisation de EU (1/5)

Généralisation de vNM aux fonctions de croyance

$\mathcal{F} =$ ensemble de toutes les fonctions de croyance sur $\mathcal{Y}$

Mixage de fonctions de croyance

$\forall f, g \in \mathcal{F}, \forall \lambda \in [0, 1], h = \lambda f + (1 - \lambda)g \in \mathcal{F}$:
fonction de croyance

$\forall Y \in \mathcal{Y},\ \ h(Y) = \lambda f(Y) + (1 - \lambda)g(Y)$

$\Longrightarrow$ comme dans von Neumann-Morgenstern, on va exprimer $\succsim$ sur l'espace des mixages de fonctions de croyance $\mathcal{F}$

Généralisation de EU (2/5)

Axiome 1 : Préordre total

$\succsim$ est un préordre large total sur $\mathcal{F}$

Axiome 2 : continuité

$\forall f, g, h \in \mathcal{F}$ tels que $f \succ g \succ h$, $\exists \alpha, \beta \in]0, 1[$ tels que :

$$\alpha f + (1-\alpha)h \succ g \succ \beta f + (1-\beta)h.$$

Axiome 2 : Indépendance

$\forall f, g, h \in \mathcal{F}$, $\forall \alpha \in]0, 1[$:

$$f \succ g \Longrightarrow \alpha f + (1-\alpha)h \succ \alpha g + (1-\alpha)h.$$

Généralisation de EU (3/5)

Théorème (von Neumann-Morgenstern)

Les deux assertions suivantes sont équivalentes :

1. $\succsim$ sur $\mathcal{F}$ vérifie les axiomes 1, 2, 3
2. $\succsim$ est représentable par $U : \mathcal{F} \mapsto \mathbb{R}$ telle que :
   - $\forall f, g \in \mathcal{F}$, $f \succsim g \Longleftrightarrow U(f) \geq U(g)$
   - U est linéaire sur $\mathcal{F}$, i.e., $\forall f, g \in \mathcal{F}$, $\forall \lambda \in [0, 1]$:

$$U(\lambda f + (1-\lambda)g) = \lambda U(f) + (1-\lambda)U(g)$$

De plus, U est unique à une transformation affine strictement positive près

Par récurrence : $\forall \lambda_i \geq 0$ t.q. $\sum_{i=1}^{n} \lambda_i = 1,\ U\left(\sum_{i=1}^{n} \lambda_i f_i\right) = \sum_{i=1}^{n} \lambda_i U(f_i)$

Généralisation de EU (4/5)

Fonction de croyance élémentaire concentrée sur B

fonction de croyance e_B telle que :

$$e_B(A) = 1 \text{ si } A \supseteq B \quad \text{et} \quad e_B(A) = 0 \text{ sinon}$$

$\Longrightarrow$ inverse de Möbius ψ_B telle que $\psi_B(B) = 1$

Ensemble focal

- f fonction de croyance d'inverse de Möbius ψ
- $\mathcal{C}_f$ = ensemble focal de f = $\{B : \psi(B) > 0\}$

$\Longrightarrow$ $\forall$ fonction de croyance f et $\forall$ événement A :

$$f(A) = \sum_{B \subseteq A} \psi(B) = \sum_{B \in \mathcal{C}_f} \psi(B) e_B(A)$$

$$f = \sum_{B \in \mathcal{C}_f} \psi(B) e_B$$

Généralisation de EU (5/5)

Rappel : $U\left(\sum_{i=1}^{n} \lambda_i f_i\right) = \sum_{i=1}^{n} \lambda_i U(f_i)$

Or $f = \sum_{B \in \mathcal{C}_f} \psi(B) e_B \Longrightarrow U(f) = \sum_{B\in\mathcal{C}_f} \psi(B) U(e_B)$

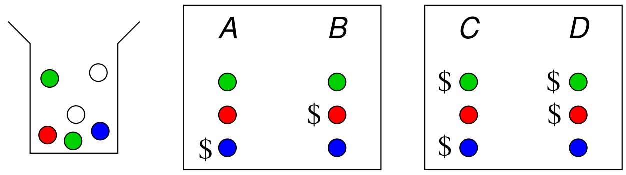
Belief expected utility

- f : fonction de croyance
- ψ : inverse de Möbius
- $BEU(f) = \sum_{B \in \mathcal{C}_f} \psi(B) w(B)$

où $w(B) = U(e_B) =$ utilité de l'ensemble de conséquences B

= utilité d'une croyance e_B dont le seul élément focal est B

BEU et l'urne d'Ellsberg



Evts	0	\$	$\{0,\$\}$
boules	$\{R,V\}$	$\{B\}$	$\mathcal{S}$
μ	2/3	1/3	1
ϕ	2/3	1/3	0

$BEU(A) = 2/3 w(\{0\}) + 1/3 w(\{\$\}) = 1/3$

Evts	0	\$	$\{0,\$\}$
boules	$\{B,V\}$	$\{R\}$	$\mathcal{S}$
μ	1/3	0	1
ϕ	1/3	0	2/3

$BEU(B) = 1/3 w(\{0\}) + 2/3 w(\{0,\$\}) = 2/3\alpha$

Evts	0	\$	$\{0,\$\}$
boules	$\{R\}$	$\{B,V\}$	$\mathcal{S}$
μ	0	1/3	1
ϕ	0	1/3	2/3

$BEU(C) = 1/3 w(\{\$\}) + 2/3 w(\{0,\$\})$

Evts	0	\$	$\{0,\$\}$
boules	$\{B\}$	$\{R,V\}$	$\mathcal{S}$
μ	1/3	2/3	1
ϕ	1/3	2/3	0

$BEU(D) = 1/3 w(\{0\}) + 2/3 w(\{\$\}) = 2/3$

Le modèle de Jaffray (1/3)

Problème de BEU :

$BEU(f) = \sum_{B \in \mathcal{C}_f} \phi(B) w(B) \Longrightarrow w$ doit être défini sur $\mathcal{C}_f$ (voire $\mathcal{Y} = 2^{\mathcal{X}}$)

en comparaison : $EU(P) = \sum_{x \in \mathcal{X}} P(x) u(x) \Longrightarrow u$ défini sur $\mathcal{X}$

$\Longrightarrow$ EU nécessite moins d'élicitation d'utilité

Le modèle de Jaffray (2/3)

Axiome 4 : Dominance

- $\forall B \in \mathcal{Y} : m_B =$ la pire des conséquences dans B
- $\forall B \in \mathcal{Y} : M_B =$ la meilleure des conséquences dans B
- $\forall B \in \mathcal{Y} : e_B$ fonction de croyance élémentaire concentrée en B
- $\forall B, B' \in \mathcal{Y}$, si $m_B \succsim_{\mathcal{X}} m_{B'}$ et $M_B \succsim_{\mathcal{X}} M_{B'}$ alors $e_B \succsim e_{B'}$

Idée :

aucune information sur les sous-événements inclus dans B
$\Longrightarrow$ décider en fonction de la pire et de la meilleure conséquence

Le modèle de Jaffray (3/3)

Théorème — Jaffray (89)

Les deux assertions suivantes sont équivalentes :

1. $\succsim$ sur $\mathcal{F}$ vérifie les axiomes 1, 2, 3, 4
2. $\succsim$ est représentable par $U : \mathcal{F} \mapsto \mathbb{R}$ telle que :
   - $\forall f, g \in \mathcal{F}, f \succsim g \Longleftrightarrow U(f) \geq U(g)$
   - $\exists v$ définie sur $\mathcal{M} = \{(m, M) \in \mathcal{X} \times \mathcal{X} : M \succsim_{\mathcal{X}} m\}$ t.q. :
   $$U(f) = \sum_{B \in \mathcal{C}_f} \psi(B) v(m_B, M_B)$$
   - U et v sont uniques à une transformation affine croissante commune près
   - v est une fonction non décroissante en m et en M
   - $u(x) = v(x, x)$ est une utilité de vNM

Le critère d'Hurwicz

Modèle de Jaffray $\Longrightarrow$ v défini sur $|\mathcal{X}|^2$ éléments

$v(m, M) \Longrightarrow$ attitude $\begin{cases} \text{vis à vis du risque} \\ \text{vis à vis de l'ambiguïté} \end{cases}$

Critère d'Hurwicz

$\forall (m, M) \in \mathcal{M}$, $\alpha(m, M) =$ critère local d'optimisme/pessimisme

$\alpha(m, M) =$ valeur de α pour laquelle le décideur est indifférent entre recevoir :

1. m avec la proba α et M avec la proba $1 - \alpha$
2. au moins m et au plus M, sans autre information

$$v(m, M) = \alpha(m, M) u(m) + [1 - \alpha(m, M)] u(M)$$

pessimisme : $\alpha(m, M) = 1$ optimisme : $\alpha(m, M) = 0$

3 Décisions séquentielles avec EU

Rappel : Utilité espérée

Loteries et espérance d'utilité

- $\mathcal{X}$: ensemble des conséquences
- $\mathcal{S}$: ensemble des états de la nature
- acte : fonction $\mathcal{S} \mapsto \mathcal{X}$
- $\mathcal{V}$: ensemble des applications de $\mathcal{S}$ dans $\mathcal{X}$
- $\succsim_{\mathcal{D}}$: relation de préférence sur $\mathcal{V}$
- $f \in \mathcal{V} \Longrightarrow P_f$ sur $(\mathcal{X}, 2^{\mathcal{X}})$
- $\mathcal{L}$: loteries, ensemble des lois à support fini sur $\mathcal{X}$
  $f \Longrightarrow P_f = \langle c_1, p_1; \ldots; c_n, p_n \rangle$, avec $c_1 \precsim_{\mathcal{X}} c_2 \precsim_{\mathcal{X}} \cdots \precsim_{\mathcal{X}} c_n$
  $\succsim_{\mathcal{D}} \Longrightarrow \succsim$ sur $\mathcal{L}$
- $\succsim$ est représentable par U telle que

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

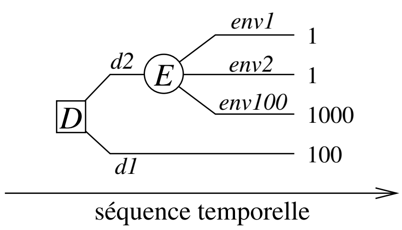
où $u(c_i) = U(\langle c_i, 1 \rangle)$.

Les Arbres de décision

Exemple

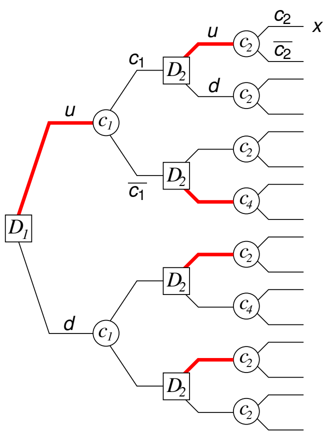
- enveloppe 1 contient 100 €
- enveloppe 2 choisie parmi une pile de 100 enveloppes dont 3 contiennent 1000 € et 97 contiennent 1 €

$\Longrightarrow \begin{cases} \text{enveloppe 1 = 100 €} \\ \text{enveloppe 2 = 3 chances sur 100 d'avoir 1000 € et} \\ \qquad\qquad\quad \text{97 chances sur 100 d'avoir 1 €} \end{cases}$



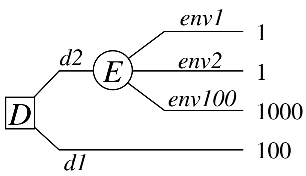
carrés = décisions ronds = nœuds de chance

Décisions séquentielles (1/3)



x = conséquence si $D_1 = u$, $C_1 = c_1$, $D_2 = u$, $C_2 = c_2$

Décisions séquentielles (2/3)



$\Rightarrow$ la décision optimale est celle dont la moyenne des utilités des conséquences est la plus élevée

$d_1 \equiv$ loterie $L_1 = \langle 100, 1 \rangle$
$d_2 \equiv$ loterie $L_2 = \langle 1, 0.97; \quad 1000, 0.03 \rangle$

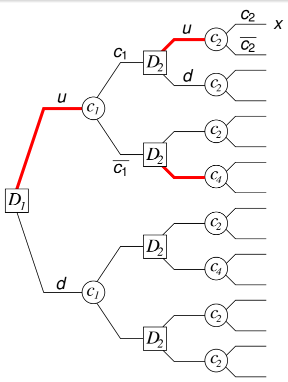
$EU(d_1) = EU(L_1) = 100$
$EU(d_2) = EU(L_2) = 0,97 \times 1 + 0,03 \times 1000 = 30,97$

$\Longrightarrow$ décision optimale selon EU : d_1

Décisions séquentielles (3/3)

Stratégie

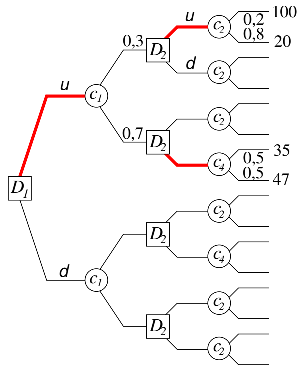
Une stratégie de décision = la sélection en tout sommet de décision D de l'arbre accessible compte tenu des décisions prises précédemment, d'une décision d appartenant à l'ensemble des décisions réalisables de ce sommet.



Décisions séquentielles (3/3)

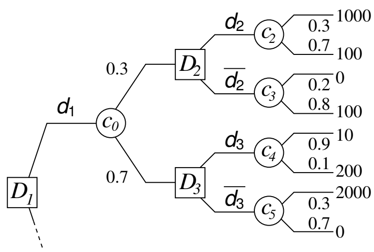
Stratégie et espérance d'utilité

- À toute stratégie correspond une loterie
- critère d'optimalité = espérance max sur les loteries



loterie = $\langle 20,\ 0.24;\ \ 35,\ 0.35;\ \ 47,\ 0.35;\ \ 100,\ 0.06\rangle$

Calculs dans un arbre de décision (1/5)



Stratégie $S_1 = \langle\langle D_1 = d_1, D_2 = d_2, D3 = d_3\rangle\rangle$

$\equiv \langle 10, 0.7 \times 0.9;\ 100, 0.3 \times 0.7,\ 200, 0.7 \times 0.1,\ 1000, 0.3 \times 0.3\rangle$

$E(S_1) = 0.3 \times [0.7 \times 100 + 0.3 \times 1000] + 0.7 \times [0.9 \times 10 + 0.1 \times 200]$

$= 0.3 \times E\left(\langle\langle D_2 = d_2\rangle\rangle\right) + 0.7 \times E\left(\langle\langle D_3 = d_3\rangle\rangle\right)$

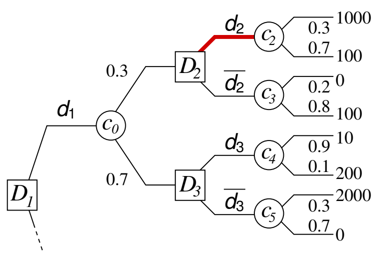
Stratégie $S_2 = \langle\langle D_1 = d_1, D_2 = d_2, D3 = \overline{d_3}\rangle\rangle$

$\equiv \langle 0, 0.7 \times 0.7;\ 100, 0.3 \times 0.7,\ 1000, 0.3 \times 0.3,\ 2000, 0.7 \times 0.3\rangle$

$E(S_2) = 0.3 \times [0.7 \times 100 + 0.3 \times 1000] + 0.7 \times [0.7 \times 0 + 0.3 \times 2000]$

$= 0.3 \times E\left(\langle\langle D_2 = d_2\rangle\rangle\right) + 0.7 \times E\left(\langle\langle D_3 = \overline{d_3}\rangle\rangle\right)$

Calculs dans un arbre de décision (2/5)



Stratégie $S_1 = \langle\langle D_1 = d_1, D_2 = d_2, D3 = d_3\rangle\rangle$

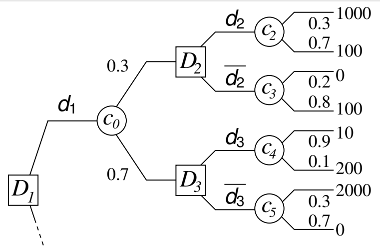
$E(S_1) = 0.3 \times E\left(\langle\langle D_2 = d_2\rangle\rangle\right) + 0.7 \times E\left(\langle\langle D_3 = d_3\rangle\rangle\right)$

Stratégie $S_2 = \langle\langle D_1 = d_1, D_2 = d_2, D3 = \overline{d_3}\rangle\rangle$

$E(S_2) = 0.3 \times E\left(\langle\langle D_2 = d_2\rangle\rangle\right) + 0.7 \times E(\langle\langle D_3 = \overline{d_3}\rangle\rangle)$

$\Longrightarrow$ calculer $E\left(\langle\langle D_2 = d_2\rangle\rangle\right)$ une seule fois, stocker le résultat en D_2 et le réutiliser pour toute stratégie contenant $D_2 = d_2$

Calculs dans un arbre de décision (3/5)



Problème : Doit-on stocker en D_2 les 2 espérances $E(\langle\langle D_2 = d_2\rangle\rangle)$ et $E(\langle\langle D_2 = \overline{d_2}\rangle\rangle)$?

Soit $S_1 = \langle\langle D_1 = d_1, \ldots, D_2 = d_2\rangle\rangle$ et $S_2 = \langle\langle D_1 = d_1, \ldots, D_2 = \overline{d_2}\rangle\rangle$ deux stratégies ne différant que par la décision D_2

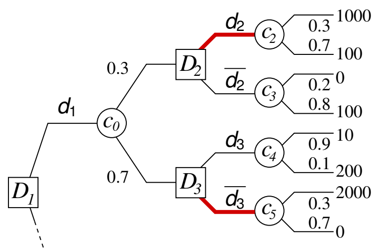
Alors $E(S_1) - E(S_2) = E(\langle\langle D_2 = d_2\rangle\rangle) - E(\langle\langle D_2 = \overline{d_2}\rangle\rangle)$

$\Longrightarrow$ Si $E(\langle\langle D_2 = d_2\rangle\rangle) \geq E(\langle\langle D_2 = \overline{d_2}\rangle\rangle)$ alors $E(S_1) \geq E(S_2)$

$\Longrightarrow$ ne conserver que $E(\langle\langle D_2 = d_2\rangle\rangle)$ dans le nœud D_2

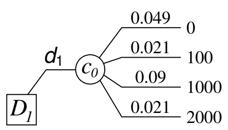
Calculs dans un arbre de décision (4/5)

$\Longrightarrow$ sur les nœuds de décision ≪terminaux≫, ne conserver que les meilleures décisions :



$\Longrightarrow$ Si on choisit $D_1 = d_1$, la sous-stratégie optimale est forcément : $D_2 = d_2$, $D_3 = d_3$ qui correspond à la loterie :

$$\langle 0,\ 0.49;\ \ 100,\ 0.21;\ \ 1000,\ 0.09;\ \ 2000,\ 0.21\rangle$$

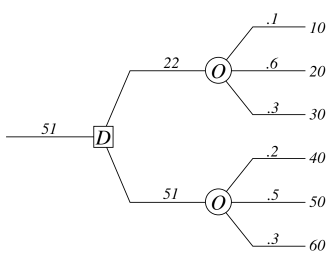


puis réitérer le process...

Calculs dans un arbre de décision (5/5)

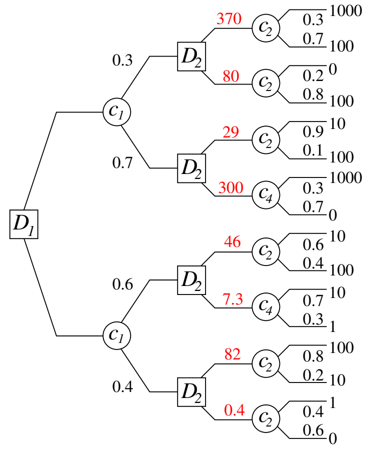
Règle de calcul dans l'arbre de décision

1. si le nœud est un nœud de chance, on calcule une espérance
2. si le nœud est un nœud de décision, on conserve le max



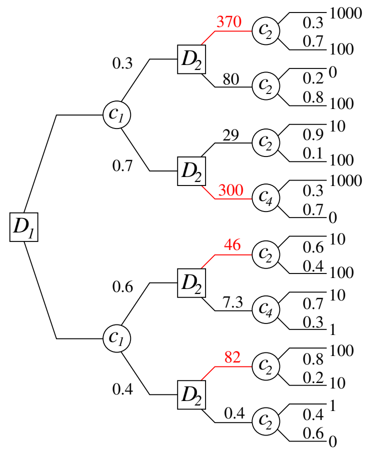
Méthode de calcul = inférence arrière

Exemple d'inférence dans un arbre de décision (1/4)



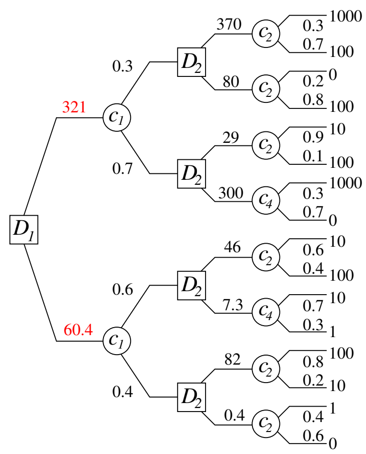
calcul : $EU(C_2) = \sum_{C_2} P(C_2|D_1, C_1, D_2) u(D_1, C_1, D_2, c_2)$

Exemple d'inférence dans un arbre de décision (2/4)



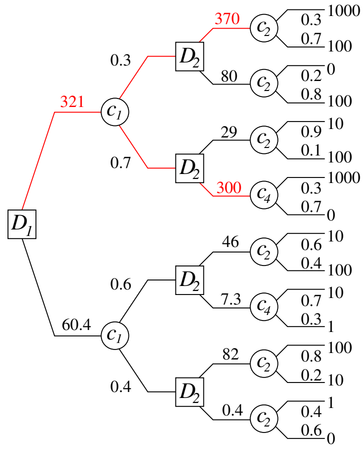
calcul : $EU(D_2) = \max_{D_2} EU(C_2)$

Exemple d'inférence dans un arbre de décision (3/4)



calcul : $EU(C_1) = \sum_{C_1} P(C_1|D_1) EU(D_2)$

Exemple d'inférence dans un arbre de décision (4/4)

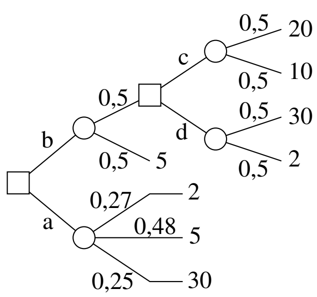


calcul : $EU(D_1) = \max_{D_1} EU(C_1)$

4 Décisions séquentielles avec RDU

RDU : Problème de cohérence dynamique

Supposons que $\varphi(x) = e^{-\sqrt{-ln(x)}}$



$RDU(a) \quad = 2 + (5-2)\varphi(0,73) + (30-5)\varphi(0,25) = 11,41$
$RDU(bc) = 5 + (10-5)\varphi(0,5) + (20-10)\varphi(0,25) = 10,26$
$RDU(bd) = 2 + (5-2)\varphi(0,75) + (30-5)\varphi(0,25) = 11,46$

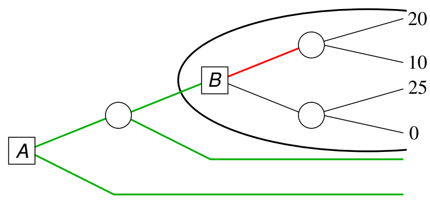
$RDU(c) \quad = 10 + (20-10)\varphi(0,5) = 14,35$
$RDU(d) \quad = 2 + (30-2)\varphi(0,5) = 14,18$

Problème : peut-on échapper au problème de cohérence dynamique ?

Conséquentialisme

Définition

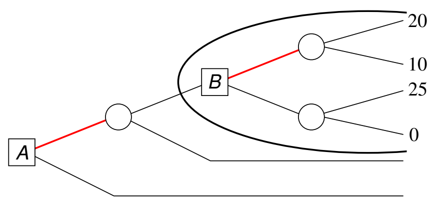
Les préférences dans le sous-arbre de racine B ne dépendent pas du reste de l'arbre



cohérence dynamique

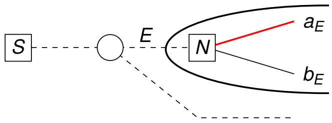
Définition

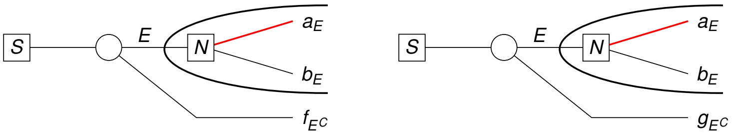
La stratégie préférée en A génère une sous-stratégie de racine B qui est la stratégie préférée en B



Conséquentialisme + cohérence dynamique

conséquentialisme :
pref ne dépend pas du
sous-arbre en pointillé





cohérence dynamique : en S, la stratégie rouge en N est préférée dans les deux arbres

Donc $(a_E, f_{E^c}) \succ (b_E, f_{E^c}) \Longleftrightarrow (a_E, g_{E^c}) \succ (b_E, g_{E^c})$

conséquentialisme + cohérence dynamique
$\Longrightarrow$ sure thing principle

Quelques références

Références

1. A. P. Dempster (1967)
“Upper and Lower Probabilities Induced by a Multivalued Mapping”, Annals of Mathematical Statistics, Vol 38, pp. 325-339
2. G. Shafer (1976)
“Mathematical Theory of Evidence”, Princeton University Press
3. J.-Y. Jaffray (1989)
“Linear Utility Theory For Belief Functions”, Operations Research Letters, Vol 8, pp. 107–112
4. L. Hurwicz (1951)
“Optimality Criteria for Decision Making Under Ignorance”, Cowles Commission discussion paper, Statistics, N° 370