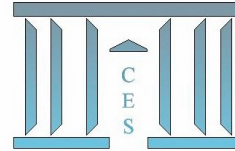




Documents de Travail du Centre d'Economie de la Sorbonne

C
E
S

W
o
r
k
i
n
g
p
a
p
e
r
s



Decision under risk : The classical Expected Utility Model

Alain CHATEAUNEUF, Michèle COHEN, Jean-Marc TALLON

2008.85



CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

Decision under risk : The classical Expected Utility model

Alain Chateauneuf, Michèle Cohen
and Jean-Marc Tallon

Paris School of Economics and
CES -University Paris I Panthéon-Sorbonne

December 3, 2008

Résumé

Ce chapitre d'ouvrage collectif a pour but de présenter les bases de la modélisation de la prise de décision dans un univers risqué. Nous commençons par définir, de manière générale, la notion de risque et d'accroissement du risque et rappelons des définitions et catégorisations (valables en dehors de tout modèle de représentation) de comportements face au risque. Nous exposons ensuite le modèle classique d'espérance d'utilité de von Neumann et Morgenstern et ses principales propriétés. Les problèmes posés par ce modèle sont ensuite discutés et deux modèles généralisant l'espérance d'utilité brièvement présentés.

Mots clé: risque, aversion pour le risque, espérance d'utilité, von Neumann et Morgenstern, Paradoxe d'Allais.

JEL: D81

Abstract

This chapter of a collective book aims at presenting the basics of decision making under risk. We first define notions of risk and increasing risk and recall definitions and classifications (that are valid independently of any representation) of behavior under risk. We then review the classical model of expected utility due to von Neumann and Morgenstern and its main properties. Issues raised by this model are then discussed and two models generalizing the expected utility model are briefly discussed.

Key words: risk, risk aversion, expected utility, von Neumann et Morgenstern, Allais Paradox.

JEL: D81

1 Introduction

We take most of our decisions without knowing for sure their consequences: the outcome of each decision depends on the realization of some uncertain event.

To help an individual facing a decision problem under uncertainty, the axiomatic approach that we adopt in this chapter takes the following steps. 1) Assume that the decision maker has well defined preferences; 2) Propose a set of “rationality” axioms that are sufficiently intuitive; 3) Derive from these axioms a representation of the decision maker’s preferences; 4) Evaluate all possible decisions with this representation and take the one that score best.

Importantly, if the decision maker agrees with the set of axioms, he will also agree with the model derived from these axioms. This model will then help him to take the right decisions.

In this chapter, we deal with a specific uncertain environment, known as “risk”, in which the probability of each event is known. In this setting, we define the main properties of decision under risk, the different possible behavior under risk and their comparison. We study next the standard model of behavior under risk, the expected utility model. We provide axiomatic foundation, study the properties of the model and its behavioral implications. We then discuss the problems raised by this expected utility model, which cannot account for observed behavior. We then expose a few alternative models, although the main generalization of the expected utility model, in which decisions are evaluated by Choquet integrals will be presented in chapter 3.

1.1 Decision under uncertainty

A decision problem under uncertainty is usually described through a set S called the set of states of nature (or states of the world), identifying events with subsets of S .

We will only need to use the sub-family of “relevant” events for the problem at hand and will then use the smallest σ -algebra \mathcal{A} including this sub-family.

Denote \mathcal{C} a set of possible outcomes or consequences, and \mathcal{G} an algebra containing the singletons of \mathcal{C} . A *decision*, or *act* is defined as a measurable mapping from (S, \mathcal{A}) to $(\mathcal{C}, \mathcal{G})$.

Denote \mathfrak{X} the set of all such mappings from S to consequences \mathcal{C} .

We assume that a decision maker has a well defined weak preference relation \succsim on \mathfrak{X} . Strict preference is denoted \succ and indifference is denoted \sim . The preference relation on \mathfrak{X} *induces* (through constant acts) a preference relation on the set \mathcal{C} of consequences. Abusing notation, we also denote this preference relation \succsim on \mathcal{C} .

We aim at representing the decision maker’s preferences (\mathfrak{X}, \succsim) by a real valued utility function, that is, a mapping V from \mathfrak{X} to \mathbb{R} such that: $X \succsim Y \iff V(X) \geq V(Y)$.

This functional will take different forms depending on the set of axioms one imposes.

1.2 Risk vs uncertainty

One can distinguish different forms of uncertainty according to the information the decision maker has on the states of nature.

The two extreme situations are (i) *risk*, in which there exists a *unique probability distribution* P on (S, \mathcal{A}) , and this distribution is objectively known; (ii) *total ignorance*, in which no information whatsoever is available on the events¹.

In between these two extreme cases, one can distinguish among different situations depending on how much information one has on the probability of the various events.

In this chapter, we assume that the decision maker is in a situation of risk. He knows the probability distribution P , which is exogenous, on the set of states of natures.

The set (S, \mathcal{A}) endowed with this probability measure is thus a probability space (S, \mathcal{A}, P) .

Since each decision X induces a probability distribution P_X on $(\mathcal{C}, \mathcal{G})$, and under the rather natural assumption that two decisions with the same probability distribution are equivalent, the preference relation \succsim on \mathfrak{X} , induces a preference relation, denoted \succsim with a slight abuse of notation, defined on the set of probability distributions with support in \mathcal{C} .

We furthermore identify consequence c of \mathcal{C} with the Dirac measure δ_c in \mathcal{L} and will indifferently use the notation $\delta_c \succsim P_X$ or $c \succsim P_X$ or $c \succsim X$.

We therefore work on the set \mathcal{L} of probability distributions on $(\mathcal{C}, \mathcal{G})$ endowed with the relation \succsim . The decision maker must therefore compare probability distributions.

Let $\mathcal{L}_0 \subset \mathcal{L}$ be the set of probability distributions with *finite support* in \mathcal{C} ; in this case, the probability distribution of decision X is denoted $P_X = (x_1, p_1; \dots; x_n, p_n)$ where $x_1 \preceq \dots \preceq x_n$, $p_i \geq 0$ and $\sum_i p_i = 1$. Such distributions with finite supports are called lotteries.

In the following section, we give intrinsic definitions (that is, independent from representation models) of measures of risk and risk aversion.

¹The distinction between risk and uncertainty is due to Knight (1921).

2 Risk and increasing risk: comparison and measures

In this section, we take \mathfrak{X} to be the set of \mathcal{A} -measurable functions² that are real valued and bounded, from (S, \mathcal{A}) to $(\mathbb{R}, \mathcal{B})$, that is, the set of *real, bounded, random variables*.

2.1 Notation and definitions

A decision in \mathfrak{X} is thus a random variable X , whose probability distribution P_X , defined by: for all $B \subset \mathbb{R}$, $P_X(B) = P\{s \in S \mid X(s) \in B\}$. We will restrict our attention to bounded random variables. P_X has a cumulative distribution function $F_X(x) = P(\{s \in S : X(s) \leq x\})$. Its expected value or mean is denoted $E(X)$. Define the function $G_X(x) = P(\{s \in S : X(s) > x\}) = 1 - F_X(x)$ to be the survival function.

We are now about to give definitions for the comparison of various probability distributions. The preference relations so defined will be partial. They are sometimes called stochastic orders.

2.1.1 First order stochastic dominance

Definition 1 *Let X and Y be elements of \mathfrak{X} , X **FSD** Y (X first order stochastically dominates Y) if for all $x \in \mathbb{R}$, $F_X(x) \leq F_Y(x)$.*

We can make a few remarks concerning this definition.

- If X **FSD** Y , the graph of Y is above the graph of X .
- This condition can be expressed through survival functions.
($Pr\{X > x\} \geq Pr\{Y > x\}$) : for all x , the probability of having more than x is always larger for X than for Y .
- This notion implies that $E(X) \geq E(Y)$.
- The relation **FSD** ranks only partially the elements of \mathcal{L} .

Let X and Y be two acts whose consequences are described in each state of nature $s \in S$ in the following table:

	s_1	s_2	s_3	s_4
$Pr(\{s\})$	0.1	0.2	0.3	0.4
X	1	2	3	4
Y	2	1	2	3

²A mapping $X : (S, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$ is measurable if for all $B \in \mathcal{B}$, $X^{-1}(B) \in \mathcal{A}$ where $X^{-1}(B) = \{s \in S : X(s) \in B\}$, and \mathcal{B} is the algebra of Borel sets.

The cumulative distribution functions of X and Y are respectively:

$$F_X(x) = \begin{cases} 0 & \text{if } x < 1 \\ 0.1 & \text{if } 1 \leq x < 2 \\ 0.3 & \text{if } 2 \leq x < 3 \\ 0.6 & \text{if } 3 \leq x < 4 \\ 1 & \text{if } 4 < x \end{cases} \quad F_Y(x) = \begin{cases} 0 & \text{if } x < 1 \\ 0.2 & \text{if } 1 \leq x < 2 \\ 0.6 & \text{if } 2 \leq x < 3 \\ 1 & \text{if } 3 \leq x < 4 \\ 1 & \text{if } 4 < x \end{cases}$$

It is easily seen that for all x in \mathbb{R} , $F_X(x) \leq F_Y(x)$ and hence X **FSD** Y . We can now give a characterization of this notion:

Proposition 1 *X first order stochastically dominates Y (X **FSD** Y) if and only if for all increasing functions u from \mathbb{R} to \mathbb{R} , $\int u(x)dF_X(x) \geq \int u(x)dF_Y(x)$.*

2.1.2 Second order stochastic dominance

One can also compare probability distributions according to their "risks": there exist several possible definitions of what it means for a distribution to be more risky than another one. For each of these notions, there is an associated notion of risk aversion and these different definitions are independent of the decision model that is retained.

The usual notion of increasing risk is the one associated with second order stochastic dominance, denoted **SSD**, introduced in economics by Rothschild and Stiglitz (1970).

Definition 2 *Let X and Y be in \mathfrak{X} , X **SSD** Y (X second order stochastically dominates Y) if for all $T \in \mathbb{R}$,*

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

We have the following implication: $[X \text{ **FSD** } Y] \Rightarrow [X \text{ **SSD** } Y]$. The converse is false.

The relation **SSD** also ranks partially only the elements of \mathcal{L} .

When X and Y have the same mean, we have the following definition:

Definition 3 *For X and Y in \mathfrak{X} , Y **MPS** X (Y is a mean preserving spread of X), if*

- (i) $E(X) = E(Y)$ and (ii) X **SSD** Y .

One also says that Y is more risky with the same mean than X .

Example 1 *Let X and Z be two acts whose decisions in each state of nature $s \in S$ are reported in the following table:*

	s_1	s_2	s_3	s_4
$\Pr(\{s\})$	0,1	0,2	0,3	0,4
X	1	2	3	4
Z	4	1	4	3

Computing the cumulative distribution functions of X and Z respectively, we get:

$$F_X(x) = \begin{cases} 0 & \text{if } x < 1 \\ 0.1 & \text{if } 1 \leq x < 2 \\ 0.3 & \text{if } 2 \leq x < 3 \\ 0.6 & \text{if } 3 \leq x < 4 \\ 1 & \text{if } 4 < x \end{cases} \quad F_Z(x) = \begin{cases} 0 & \text{if } x < 1 \\ 0.2 & \text{if } 1 \leq x < 2 \\ 0.2 & \text{if } 2 \leq x < 3 \\ 0.6 & \text{if } 3 \leq x < 4 \\ 1 & \text{if } 4 < x \end{cases}$$

One can check that $E(X) = E(Y)$ and that the surface between 0 and T of the difference between $F_Z(x)$ and $F_X(x)$ is always positive which implies that X **DS2** Z and hence Z **MPS** X .

The following proposition makes the definition of **MPS** more intuitive:

Proposition 2 (Rothschild and Stiglitz, 1970) *Let X and Y be two random variables with the same mean. The following three assertions are equivalent:*

- (i) Y is a mean preserving spread of X , that is Y **MPS** X .
- (ii) Y has the same distribution as $(X + \theta)$ where θ is a random variable such that $E(\theta|X) = 0$ almost everywhere.
- (iii) For all increasing and concave u from $\mathbb{R} \rightarrow \mathbb{R}$, $\int u(x)dF_X(x) \geq \int u(x)dF_Y(x)$.

The characterization (ii) reveals the intuition behind the definition since it expresses the fact that Y is a mean preserving spread of X when Y can be obtained by adding to X some “noise”. After we define an expected utility decision maker (section 3), we will interpret condition (iii) behaviorally.

Remark 1 *Several other notions of increasing risk can be defined, that we will not develop here although they are sometimes more suited for particular comparative static problems.³ some of these notions will be developed in chapter 3.*

Remark 2 *If Y **MPS** X , then the variance of Y is greater than or equal to that of X . The converse is not always true (see section 3.3.3). The variance could be intuitively used as a measure of increases in risk but we’ll see in section 3.3.3 that it might lead to inconsistencies.*

³Chateauneuf, Cohen et Meilijson (1997) and (2004) contain some of these definitions and their properties as well as applications.

2.2 Behavior under risk

2.2.1 Model-free behavioral definitions

For a decision maker with a preference relation \succsim on \mathfrak{X} , we define some typical behavior under risk.

Let us first recall that comparing $E(X)$ and X amounts to compare $\delta_{E(X)}$ and X .

Risk aversion We can define various notions of risk aversion. We will concentrate on two of them in this chapter.

Definition 4 (*Arrow (1965)-Pratt (1964)*) *An agent is weakly risk averse if for any random variable X in \mathfrak{X} , he prefers its expected value to the random variable itself:*

$$\forall X \in \mathfrak{X}, E(X) \succsim X$$

He is weakly risk seeking if

$$\forall X \in \mathfrak{X}, X \succsim E(X)$$

He is risk neutral if

$$\forall X \in \mathfrak{X}, X \sim E(X)$$

This definition usually is not enough to rank two distributions with the same mean. One can also use another definition:

Definition 5 *a decision maker is strongly risk averse if for any couple of random variables X, Y in \mathfrak{X} with the same expected value and such that Y is more risky than X according to second order stochastic dominance, he prefers the less risky variable X to Y :*

$$\forall X, Y \in \mathfrak{X}, E(X) = E(Y), Y \text{ MPS } X \implies X \succsim Y$$

He is strongly risk seeking if

$$\forall X, Y \in \mathfrak{X}, E(X) = E(Y), Y \text{ MPS } X \implies Y \succsim X$$

He is risk neutral if

$$\forall X, Y \in \mathfrak{X}, E(X) = E(Y), Y \text{ MPS } X \implies Y \sim X$$

If a decision maker does not always rank couple of random variables ($Y \text{ MPS } X$) in the same way, he will not fit into any of these categories.

Remark 3 *It is easy to see that for all X , $X \text{ MPS } E(X)$ (that is, X is a mean preserving spread of $E(X)$); hence, a strongly risk averse decision maker will necessarily be weakly risk averse. The converse does not hold in general.*

2.2.2 Certainty equivalent, risk premium and behavior comparison.

Certainty equivalent For each random variable X in \mathfrak{X} , its *certainty equivalent*, if it exists, is the certain outcome c_X in \mathcal{C} that is indifferent to X : $c_X \sim X$. In this chapter, any random variable in \mathfrak{X} has a unique certainty equivalent.

Risk premium The *risk premium* attached to X is denoted ρ_X . It is the difference between the expected value of X and its certainty equivalent: $\rho_X = E(X) - c_X$.

This premium can be interpreted as the maximum amount the decision maker is willing to pay to exchange the variable X against its mean. This premium is negative whenever the decision maker is weakly risk seeking.

The risk premium captures the intensity of (weak) aversion to risk. It is possible to use it to compare, in a partial manner, different behaviors.

Definition 6 A decision maker 1 is more weakly risk averse than a decision maker 2 if for all $X \in \mathfrak{X}$, the risk premium ρ_X^1 of decision maker 1 is greater than or equal to the risk premium ρ_X^2 of decision maker 2.

3 The expected utility (EU) model of von Neumann-Morgenstern

We now come back to the decision problem faced by a decision maker endowed with a preference relation \succsim on \mathfrak{X} and the associated relation on \mathcal{L} .

We now expose the axiomatic foundation of the classical model of decision under risk: the *expected utility* (EU) model due to von Neumann-Morgenstern (vNM) (1947). We study how notions of risk aversion defined above translate in this model.

We consider here the set \mathcal{L} of probability distributions on $(\mathcal{C}, \mathcal{G})$ where \mathcal{C} is a set endowed with an algebra \mathcal{G} which contains by assumption all the singletons.

Before giving the axioms, we first need to define a mixture operation on the set of probability distributions.

Mixing probability distributions For all $P, Q \in \mathcal{L}$ and $\alpha \in [0, 1]$, we call α -mixture of P and Q , the distribution $R = \alpha P + (1 - \alpha)Q$, such that, for all A in \mathcal{A} , $R(A) = \alpha P(A) + (1 - \alpha)Q(A)$. The mixture of two distributions can be interpreted, when P and Q are discrete, as a two-stage lottery: in the first stage, the distributions P and Q are drawn with probability α and $(1 - \alpha)$ respectively, while in the second stage, a consequence is chosen according to the distribution drawn in the first stage.

\mathcal{L} is then a convex subset of a vector space.

Generalized mixture For any $\alpha_1, \dots, \alpha_n \in [0, 1]$, such that $\sum \alpha_i = 1$ and for all $P_1, \dots, P_n \in \mathcal{L}$, one can define in a similar way the mixture of these n distributions P_i , as the distribution $R = \sum \alpha_i P_i$ such that for all $A \in \mathcal{G}$, $R(A) = \sum \alpha_i P_i(A)$.

Any distribution with finite support can thus be written as a mixture of Dirac distributions: $P = \sum_{i=1}^{i=n} p_i \delta_{x_i}$. This can also be referred to as a "convex combination of consequences".

3.1 Axiomatic foundation of the EU model

The model is based on three fundamental axioms: a weak order axiom, a continuity axiom and an independence axiom, to which an extra axiom is added in certain cases.

One can find intuitive presentations of the axioms as well as elegant proofs of the vNM theorem (both in the finite and infinite case) for instance in Herstein-Milnor (1957), Jaffray (1978, 2000), Jensen (1967), Fishburn (1970, 1982), Kreps (1988).

We follow here the presentation of J.Y. Jaffray (1978, 2000).

Axiom 1 : Weak Order *The preference relation \succsim on \mathcal{L} , is a complete weak order that is non trivial.*

The relation \succsim is thus reflexive, transitive and complete, and furthermore, there exists at least one couple P, Q in \mathcal{L} such that $P \succ Q$.

The completeness assumption, i.e., the fact that the decision maker is assumed to be able to rank all alternatives, is rather strong in some context. It is nevertheless widely accepted in standard economic models.

Adding the continuity axiom to this Weak Order axiom allows the relation \succsim to be *represented by a functional often called utility function* $V : \mathfrak{X} \rightarrow \mathbb{R}$, such that :

$$X \succsim Y \iff V(X) \geq V(Y).$$

The precise formulation of the continuity axiom depends on the topological structure of \mathfrak{X} , on which preferences are defined (see Debreu 1954, Grandmont 1972).

When this set is a set of probability distributions, as it is the case here, the continuity axiom can be expressed as follows :

Axiom 2 : Continuity *The preference relation \succsim on \mathcal{L} is continuous (in the sense of Jensen 1967): for all P, Q, R in \mathcal{L} , such that $P \succ Q \succ R$, $\exists \alpha, \beta \in]0, 1[$ such that $\alpha P + (1 - \alpha)R \succ Q$ and $Q \succ \beta P + (1 - \beta)R$.*

This axiom requires that there does not exist a consequence c_M that is so desirable (actually infinitely desirable) that if a distribution P_0 offered this consequence with the smallest (positive) probability, this distribution would be preferred to any other distribution that would not have c_M as a possible consequence. This axiom also prevents the existence of an infinitely undesirable consequence c_m (one can think of c_M as being paradise and c_m as being hell, as in Pascal's bet).

The independence axiom that we now give is central in the construction of the expected utility model.

Axiom 3 : Independence For all P, Q, R in \mathcal{L} and all $\alpha \in]0, 1]$,

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

This axiom can be interpreted as follows (building on the interpretation of the mixture we gave above). A decision maker that prefers P to Q and who has to choose between the two mixtures $\alpha P + (1 - \alpha)R$ and $\alpha Q + (1 - \alpha)R$ would reason in the following manner. If an event of probability $(1 - \alpha)$ occurs, he gets the distribution R independently of his choice, whereas if the complement event occurs, he faces the choice between P and Q . Since he prefers P , he will prefer the mixture $\alpha P + (1 - \alpha)R$. The logic behind this axiom seems therefore rather intuitive.⁴

However, this axiom has concentrated on itself a lot of criticism: numerous experimental studies (such as the Allais paradox) have shown that most decision makers take decisions that contradict this axiom. We will come back on this issue in section 4.1.

The representation theorem can be decomposed into two important results:

3.1.1 Linear utility theorem

Theorem 1 Let (\mathcal{L}, \succsim) be a convex subset of a vector space on \mathbb{R} endowed with a preference relation \succsim . The following two conditions are equivalent:

- (i) (\mathcal{L}, \succsim) satisfy Weak order, continuity and independence.
- (ii) There exists a linear function U from \mathcal{L} to \mathbb{R} that represents the weak order, that is, (a) such that for all X, Y from \mathcal{L} , $X \succsim Y \iff U(X) \geq U(Y)$; (b) such that for all α in $[0, 1]$, $U(\alpha X + (1 - \alpha)Y) = \alpha U(X) + (1 - \alpha)U(Y)$.

Proof 1 We only give a sketch of the proof. It consists in proving the following points:

- (1) For all P, Q in \mathcal{L} and μ in $]0, 1]$, $P \succ Q \implies P \succ \mu Q + (1 - \mu)P$;
- (2) For all P, Q in \mathcal{L} and λ, μ in $[0, 1]$, $[P \succ Q, \lambda > \mu] \implies \lambda P + (1 - \lambda)Q \succ \mu P + (1 - \mu)Q$;

⁴For an interesting discussion around the independence axiom see Fishburn and Wakker, 1992.

(3) For all P, Q, R in \mathcal{L} such that $P \succ Q \succ R$, there exists α in $]0, 1[$ such that $Q \sim \alpha P + (1 - \alpha)R$;

(4) For all P_1, P_2 in \mathcal{L} such that $P_1 \succ P_2$, there exists a linear utility function on the set \mathcal{L}_{12} defined by: $\mathcal{L}_{12} = \{P \in \mathcal{L}, P_1 \succ P \succ P_2\}$;

(5) This linear utility on \mathcal{L}_{12} is unique up to a positive affine transformation;

(6) Extrapolating; there exists a linear utility function on (\mathcal{L}, \succsim) which is unique up to a positive affine transformation.

This theorem is true as soon as \mathcal{L} is a convex subset of a vector space. In order to be able to express the functional U as an expected utility, it is necessary to explore several cases according to the structure of \mathcal{C} .

- When the distributions do not have finite support, the algebra \mathcal{G} must be rich enough and, furthermore, one need to add a dominance axiom.
- When \mathcal{C} is already ordered (by a complete weak order) –in particular when \mathcal{C} is a subset of \mathbb{R} , it is necessary to impose a stronger axiom (sometimes labeled monotonicity axiom) that ensures the compatibility of the preference relation on \mathfrak{X} and the preference relation on \mathcal{C} .

3.1.2 vNM theorem for distributions with finite support in $(\mathcal{C}, \mathcal{G})$

We restrict our attention to the set \mathcal{L}_0 of distributions with finite support over \mathcal{C} .

Any distribution P can then be written $P = (x_1, p_1; \dots; x_n, p_n)$ (i.e., P yields consequence x_i with probability p_i and $\sum p_i = 1$).

$(\mathcal{L}_0, \succsim)$ is a convex set and the previous axioms are sufficient for the representation theorem of \succsim as an expected utility.

Theorem 2 *Let \mathcal{L}_0 be the set of distributions with finite support in \mathcal{C} endowed with the preference relation \succsim . The following two statements are equivalent:*

- $(\mathcal{L}_0, \succsim)$ satisfies Weak order, continuity and independence.
- There exists a utility function U representing the relation \succsim such that

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

where u is a strictly increasing function from \mathcal{C} to \mathbb{R} defined by $u(x) = U(\delta_x)$ and unique up to a positive affine transformation.⁵

One understands easily in this formulation why the expected utility model is said to be linear: it deals with probabilities in a linear way, since $U(\alpha P + (1 - \alpha)Q) = \alpha U(P) + (1 - \alpha)U(Q)$.

⁵That is, any function $v = au + b$ with $a > 0$ is also admissible.

3.1.3 vNM theorem for distributions with bounded support in $(\mathcal{C}, \mathcal{G})$

This theorem can be generalized to the set \mathcal{L} of probability distributions on $(\mathcal{C}, \mathcal{G})$.

We first assume that $\forall c \in \mathcal{C}$, the sets $\{c' \in \mathcal{C}, c' \succsim c\}$ and $\{c' \in \mathcal{C}, c' \preceq c\}$ belong to the algebra \mathcal{G} . We furthermore impose the following dominance axiom, which will guarantee that the utility function u is bounded and therefore that $\int_{\mathcal{C}} u dP$ is well defined.

Axiom 4 *Dominance:* For all P in \mathcal{L} and c_0 in \mathcal{C} ,

- (i) $P\{c \in \mathcal{C}/c \preceq c_0\} = 1$ implies $P \preceq \delta_{c_0}$
- (ii) $P\{c \in \mathcal{C}/c \succ c_0\} = 1$ implies $P \preceq \delta_{c_0}$

This axiom expresses the fact that if all the outcomes of the distribution P are preferred to the outcome c_0 , then P is preferred to the degenerate distribution on c_0 .

For a distribution P with bounded support in $(\mathcal{C}, \mathcal{G})$, define its cumulative distribution function F as follows: F is a mapping from \mathcal{C} to $[0, 1]$ defined by $F(x) = P\{X \preceq x\}$. One can then state vNM theorem in this setting:

Theorem 3 *Let \mathcal{L} be the set of probability distributions with bounded support in $(\mathcal{C}, \mathcal{G})$ endowed with the preference relation \succsim . The following two statements are equivalent:*

- (i) (\mathcal{L}, \succsim) satisfies Weak order, continuity, independence and dominance.
- (ii) There exists a utility function U representing the relation \succsim with

$$U(P) = Eu(P) = \int_{\mathcal{C}} u(x) dF(x)$$

where u is a strictly increasing, bounded mapping from \mathcal{C} to \mathbb{R} defined by $u(x) = U(\delta_x)$, unique up to a positive affine transformation.

Remark 4 *In the remainder, a decision maker who obeys to the axioms of the vNM model will be called an EU decision maker.*

Remark 5 *The behavior of an EU decision maker is entirely characterized by the function u .*

3.1.4 vNM theorem for distributions with bounded support in $(\mathbb{R}, \mathcal{B})$

When \mathcal{C} is a subset of \mathbb{R} , axioms 1, 2, 3, even combined with axiom 4 are not sufficient to obtain the representation. One needs to express the compatibility between the preference relation \succsim and the order relation \geq that exists on \mathbb{R} ; one then need to replace axiom 4 by a stronger axiom:

Axiom 5 *Monotonicity* :

For all x, y in \mathbb{R} , $x \geq y \iff x \succsim y$

One can state the theorem in this setting:

Theorem 4 *Let \mathcal{L} be the set of probability distributions with bounded support in \mathbb{R} endowed with the preference relation \succsim . The following two statements are equivalent:*

(i) (\mathcal{L}, \succsim) satisfies Weak order, continuity, independence and monotonicity

(ii) There exists a utility function U representing the relation \succsim such that

$$U(P) = Eu(P) = \int_{\mathbb{R}} u(x)dF(x)$$

where u is a strictly increasing function from \mathbb{R} to \mathbb{R} defined by $u(x) = U(\delta_x)$, which is unique up to a positive affine transformation.

3.2 Characterization of risk aversion in the EU model

In the remainder of section 3, we consider the set \mathfrak{X} of random variables with bounded support in $(\mathbb{R}, \mathcal{B})$ and the associated set of distributions with bounded support in $(\mathbb{R}, \mathcal{B})$.

How are the two notions of risk aversion that we introduced characterized in the EU model?

The following proposition shows that they cannot be distinguished in this model.

Proposition 3 *(Rothschild and Stiglitz)* The following three assertions are equivalent for an expected utility decision maker:

(i) The decision maker is weakly risk averse.

(ii) The decision maker is strongly risk averse.

(iii) The decision maker utility function u is concave.⁶

Remark 6 *Proposition 3 thus reveals that a EU decision maker cannot make a distinction between weak and strong risk aversion. In the EU model, one can simply speak of risk aversion without ambiguity.*

3.2.1 Characterization of first and second order dominance in the EU model

One can now interpret propositions 2 and 3 (iii) of the previous section, in the EU model:

⁶The equivalence (i) \iff (iii) is straightforward. The equivalence (ii) \iff (iii) was proved by Rothschild and Stiglitz (1970).

- Proposition 2 : X **first order stochastically dominates** Y if and only if *any EU decision maker prefers X to Y .*
- Proposition 3 (iii) : a random Y is a mean preserving spread of X if and only if $E(X) = E(Y)$ and any weakly risk averse EU decision maker prefers X to Y .

3.3 Coefficient of absolute risk aversion, local value of the risk premium

In the EU model, it is possible to define the intensity of risk aversion for a decision maker, through properties of u .

3.3.1 Coefficient of absolute risk aversion

Definition 7 *When the function u , which characterized the behavior of an EU decision maker is strictly increasing, twice continuously differentiable with strictly positive derivative, we call the coefficient of absolute risk aversion the function R_A from \mathbb{R} to \mathbb{R} defined by*

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

Note that this coefficient is independent from the choice of the function u representing the preferences. Any other function $v = au + b, a > 0$ will have the same coefficient.

3.3.2 Local value of the risk premium

When the distribution X with an expected value x and a variance σ^2 , takes its values in an interval $[x - h, x + h]$ where h is “small” with respect to x , one can show (Arrow, 1965 and Pratt, 1964) that the risk premium is proportional to the coefficient of absolute risk aversion :

$$\rho_X(x) \approx -\frac{\sigma^2}{2} \frac{u''(x)}{u'(x)} = \frac{\sigma^2}{2} R_A(x)$$

This approximation is useful as it serves to distinguish in the expression of the risk premium an objective part which depends solely on the variance of the distribution from a subjective part that is linked to the decision maker’s preferences, $R_A(x)$.

This result can be used to show the following theorem that compares the behavior of two EU decision makers:

Theorem 5 *(Arrow, 1965 and Pratt, 1964) Let 1 and 2 be two EU decision makers, with utility functions u_1 and u_2 respectively, that are assumed to be*

strictly increasing and twice continuously differentiable. The following assertions are equivalent:

- (i) 1 is more risk averse than 2: $\rho_X^1 \geq \rho_X^2$.
- (ii) There exists an increasing and concave function φ from \mathbb{R} to \mathbb{R} such that $u_1 = \varphi(u_2)$.
- (iii) The coefficient of absolute risk aversion for u_1 is greater than or equal to the coefficient of absolute risk aversion for u_2 for all $x \in \mathbb{R}$: $R_A^1(x) \geq R_A^2(x)$.

Remark 7 The notions of risk premium and coefficient of absolute risk aversion can also be defined for decision makers that are not necessarily risk averse. The theorem can be applied to compare such decision makers.

Remark 8 The equivalence between (i) and (ii) remains true even when u_1 and u_2 are not twice continuously differentiable.

3.3.3 Variance and EU model

We saw in remark 6 that the notion of an increase in risk in the sense of second order stochastic dominance had a nice justification in the EU model. The variance is not a good indicator of riskiness as the following example, due to Ingersoll (1987) shows.

Consider two lotteries $P = (0, 1/2; 4, 1/2)$ and $Q = (1, 7/8; 9, 1/8)$. One has $E(P) = E(Q)$ and $Var(Q) > Var(P)$. Take a decision maker that satisfies the axioms of the EU model, who exhibits weak aversion with a concave utility function given by $u(x) = \sqrt{x}$. One can check that for this decision maker $Q \succ P$, that is, the expected utility of the lottery P , which has the smallest variance, is lower than the expected utility of Q . Thus, in the EU model, a risk averse decision maker might prefer a random variable with higher variance (holding means constant).

Note however that when $u(x)$ is a second order polynomial (as in the capital asset pricing model-CAPM), a risk averse decision maker will always prefer among two random variables with the same mean, the one with the lowest variance. One can actually show that with such a utility function, the expected utility of a distribution depends only on its means and its variance.

The EU model is widely used in economics. It has nice properties that make it very tractable. In particular, in a dynamic setting, it is compatible with dynamic programming and backward induction.

However, it has also been criticized on a number of different grounds.

4 Some problems raised by the EU model

We analyze here a few problems raised by the EU model.⁷

⁷For a review see Machina (1987).

4.1 Allais paradox

As early as 1953, Allais built a couple of alternatives for which a majority of subjects, confronted with that choice, decided in contradiction with the independence axiom.

We present here the original Allais experiment. Subjects were confronted with the choice between the following lotteries (say in euros):

L_1 : win 1M with certainty

L_2 : win 1M with probability 0.89, 5M with probability 0.10, and 0 with probability 0.01

and then with the choice among the following two lotteries:

L'_1 : win 1M with probability 0.11 and 0 with probability 0.89

L'_2 : win 5M with probability 0.10 and 0 with probability 0.90

Most subjects chose L_1 over L_2 , and L'_2 over L'_1 .

This choice violates the independence axiom. Indeed, let P be the lottery yielding 1 M with probability 1 and Q the lottery yielding 0 with probability 1/11 and 5 M with probability 10/11. One can check that:

$$L_1 = 0,11P + 0,89\delta_1$$

$$L_2 = 0,11Q + 0,89\delta_1$$

$$L'_1 = 0,11P + 0,89\delta_0$$

$$L'_2 = 0,11Q + 0,89\delta_0$$

where δ_0 is the lottery “win 0 with probability 1” and δ_1 is the lottery “win 1 M with probability 1”. The observed choice are thus in contradiction with the independence axiom.

This experiment has been ran many times, on various population of subjects with similar results: about 66% of the choices are in contradiction with the independence axiom.

These results have led researchers to acknowledge some descriptive deficiencies of the EU model. But most of them remain convinced that the normative aspect of the model should be given more weight.

4.2 Interpreting the utility function

On top of experimental violations of the independence axiom, the expected utility model raises a theoretical issue concerning the interpretation of the utility function u . This function captures at the same time two rather distinct behavioral features: (i) it captures the decision maker’s risk attitude (concavity of u implies that the decision maker is risk averse, (ii) it captures the decision maker’s attitude toward certain outcomes (concavity of u implies a decreasing marginal utility of wealth, say). It is thus impossible to represent in this model a decision maker that would be risk seeking and would have a decreasing marginal utility.

The EU model being so parsimonious cannot separate risk attitudes from attitudes toward wealth under certainty. This distinction is possible in more flexible but less parsimonious models that will be presented in chapter 3.

4.3 Weak and strong risk aversion under expected utility

As stated in remark 6, weak and strong risk aversions cannot be distinguished in the expected utility model. One can indeed interpret proposition 3 as a limit of the expected utility model: *any agent who is weakly risk averse but not strongly risk averse cannot satisfy the model's axioms*. One can show that (Chateauneuf, Cohen and Meilijson, 1997 and 2004) several notions of risk aversion, corresponding to different stochastic orders on random variables are confounded in the expected utility model. It is therefore not surprising that this model lacks flexibility to account for a wide range of different behavior in face of increases in risk in different problems. The more general Rank Dependent expected Utility model that will be presented in chapter 3 does not face this problem.

4.4 Problems with the notion of second order stochastic dominance as an indicator of riskiness in the EU model

Despite the fact that second order stochastic dominance has a nice characterization in the EU model, this notion can lead to some counter-intuitive results. We will focus on two of them. First, in the classical portfolio choice problem where the decision maker has the choice between any mixture of a risky asset and a riskless asset, a natural prediction would be that if the risky asset becomes riskier in the sense of MPS, then any risk averse investor should reduce his position in this asset. Rothschild and Stiglitz (1971) have shown that this is not always the case in the expected utility model. The second issue deals with insurance choices. A decision maker that is more risk averse than another is not necessarily ready to pay at least as much as the second one to get the same risk reduction (in the sense of MPS), see for instance Ross, 1981).

5 Some alternative models

There exist several decision models under risk that, to fit better the observed behavior, weaken one or several of the vNM axioms. We briefly expose a few of them. A review of these models and the experimental evidence around them is in Harless et Camerer (1994).

5.1 Machina's model

In the model in Machina (1982), the independence axiom is dropped while weak order and continuity are retained. Still, the functional used to represent

preferences is smooth (i.e., Fréchet differentiable). As a consequence, this model is locally very similar to the EU model since locally, the functional can be approximated by a linear functional.

5.2 Models with security and potential levels

Several experiments have shown that violations of the EU model disappear when the distributions compared have the same maximum and the same minimum. Jaffray (1988), Gilboa (1988), Cohen (1992) and Essid (1997) have axiomatized a model in which the independence axiom is satisfied only for distributions with the same minimum (Jaffray, 1988, Gilboa 1988) or same minimum and maximum (Cohen, 1992, Essid, 1997). This is a rather weak condition. One thus needs to add, on top of weak order and continuity some other axioms. The representation takes the form of the combination of three criteria: the Minimum, the Maximum and the expected utility. This model is compatible with many experiments that have shown that the iso-utility curves are parallel (as in the EU model) when the distributions have the same extremal points, while they move away from parallel lines when extrema are different.

Other models exist in the literature (see Harless et Camerer (1994)). Probably the most well-known compatible with Allais' experiment is the Rank Dependent Expected Utility Model (Quiggin (1982)), based on the Choquet integral and that will be presented in some detail in chapter 3.

References

- [1] Allais, M. "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine", *Econometrica*, **21**, 503–546, 1953.
- [2] Arrow, K.J. "The theory of risk aversion, in Aspects of the theory of risk bearing". Helsinki : Ynjo Jahnsolin Saatio, 1965.
- [3] Chateauneuf, A., M. D. Cohen and I. Meilijson. "New tools to better model behavior under risk and uncertainty: An overview", *Finance*, 18, 1997.
- [4] Chateauneuf, A., M. D. Cohen and I. Meilijson. "Four notions of mean-preserving increase in risk, risk attitudes and applications to the Rank-Dependent Expected Utility Model", *Journal of Mathematical Economics*, 40, 6, 547-571, 2004.
- [5] Cohen, M.D. "Security level, potential level, expected utility : a three-criteria decision model under risk", *Theory and Decision*, Vol 33, 2, 1-34, 1992.

- [6] Debreu, G. "Representation of a preference ordering by a numerical function", in *Decision Processes*, edit par Thrall, Coombs and Davies, p159-165, Wiley, 1954.
- [7] Essid, S. "Choice under risk with certainty and potential effects: A general axiomatic model", *Mathematical Social Sciences*, vol. 34, 3, 223-247, 1997.
- [8] Fishburn, P. *Utility theory for decision making*. Wiley, 1970.
- [9] Fishburn, P. *The foundations of expected utility*. Reidel, 1982.
- [10] Fishburn, P. and Wakker, P. "The invention of the independence condition for preferences", *Management Science*, 41, 1130-1144, 1995.
- [11] Gilboa, I., "A Combination of Expected Utility and Maxmin Decision Criteria", *Journal of Mathematical Psychology*, 32, 405-420, 1988.
- [12] Grandmont, JM. "Continuity Properties of von Neumann-Morgenstern Utility", *Journal of Economic Theory*, 1972.
- [13] Herstein and Milnor, "An axiomatic approach to measurable utility", *Econometrica*, 21, 291-297, 1953.
- [14] Ingersoll, J. *Theory of financial decision making*. Rowman and Littlefield, Totowa, NJ., 1987.
- [15] Jaffray, JY. "Choice under risk and the security factor. An axiomatic model". *Theory and Decision*, Vol 24, 2, 1988.
- [16] Jaffray, J.Y. Polycopié de l'Université de Paris VI, 1978, repris dans "Cours à l'école d'été de Cortona", 2000.
- [17] Jensen, N. "An introduction to Bernoullian utility theory: (1) Utility functions", *Swedish Journal of Economics*, 69, 163-183, 1967.
- [18] Knight, F. *Risk, Uncertainty and Profit*. Houghton Mifflin, 1921.
- [19] Kreps, D. *Notes on the theory of choice*, *Underground classics in economics*. Westview Press, 1988.
- [20] Machina, M. "Expected utility analysis without the independence axiom", *Econometrica*, 50, 277-323, 1982a.
- [21] Machina, M. "Choice under uncertainty: Problems solved and unsolved", *Journal of Economic Perspectives*, 1, 121-154, 1987.
- [22] Pratt, J. "Risk aversion in the small and in the large", *Econometrica*, 32, 122-136, 1964.

- [23] Quiggin, J. "A theory of anticipated utility", *Journal of Economic Behavior and Organisation*, **3**, 323–343, 1982.
- [24] Ross, S.. "Some stronger measures of risk aversion in the small and in the large with applications", *Econometrica*, 49, 621-638, 1981.
- [25] Rothschild, M. and J. Stiglitz. "Increasing Risk: I. A definition", *Journal of Economic Theory*, **2**, 225–243, 1970.
- [26] Rothschild, M. and J. Stiglitz. "Increasing Risk: II. Its economic consequences", *Journal of Economic Theory*, **3**, 66-84, 1971.
- [27] von Neumann, J. and Morgenstern, O. "Theory of games and economic behavior", Princeton University Press, Princeton, N.J., 1947.