Economy and Finance of Insurance

Gianni Bosi

Contents

pag.

1 Preference and utility without probability **3**
1.1 Introduction 3
1.2 Definitions and preliminaries 3
1.3 Order-preserving functions and utility functions 10
Exercises 15
1.4 Continuous utility functions 16

Bibliography **25**

2 Preferences over money lotteries and expected utility **27**
2.1 Introduction 27
2.2 Uncertainty, monetary lotteries and risk aversion 27
2.3 Certainty equivalents and risk premium 30

3 Axiomatization of expected utility under risk **39**
3.1 Introduction 39
3.2 From linear utility to expected utility for simple probabilities 39
3.3 Expected utility for general probabilities 47
3.4 Risk attitudes and the utility function 48

4 A sketch of probability and random variables **51**
4.1 Probability and random variables 51
4.2 Expectations, conditional expectations and joint distributions 52

5 Expected utility and zero-utility premium **55**
5.1 Utility theory 55
5.2 Zero-utility principle 59

6 Premium principles **63**
6.1 Basic definitions 63
6.2 Equity premium 64
6.3 Premium principles and their properties 66
6.4 Solutions to Allais and Ellsberg paradoxes 71

Solution to Allais paradox 71
Solution to Ellsberg paradox 72
6.5 Ordering of risks 72

7 Reinsurance **79**
7.1 Reinsurance types 79
7.2 A motivation for reinsurance 81
7.3 Stop-loss reinsurance 82
7.4 Optimal reinsurance: the problem 83

8 Expected utility and insurance in a two state model **85**
8.1 Expected utility and insurance in a two state model 85
8.2 Demand for insurance 87
8.3 The Portfolio Problem 88

9 Portfolio Theory **93**
9.1 Notation 93
9.2 Multi-objective optimization and portfolio selection 94
9.3 Markowitz portfolio selection 97
The case of risky assets 97
The case of two risky assets 102
The case of one riskless asset 103

10 Risk Sharing **109**
10.1 Optimal Risk Allocations 109
10.2 The sup-convolution problem and feasible set endowed with a general topology116

11 Domanda e offerta **123**
11.1 Il bene assicurazione 123
11.2 Domanda di assicurazione 125
11.3 Offerta di assicurazione 126

12 Exercises **127**
12.1 Risk aversion 127

CHAPTER 1

Preference and utility without probability

1.1 Introduction

This chapter is devoted to the presentation of the main considerations concerning binary relations on generic sets, without invoking any other tools, like for example algebraic structures or probability structures.

The first section presents the classical axioms concerning binary relations (*preference relations*), the second section is devoted to the illustration of the main notions and results concerning the representability of preferences by means of real-valued functions (namely *order-preserving functions*, or, in particular, *utility functions*), and finally the third section concerns the continuity of a utility function on generic topological space.

While, on one hand, the interdisciplinary nature of this chapter should be noticed, on the other hand the concepts presented are essential in order to understand the notions of *linear utility*, and *certainty equivalence*.

1.2 Definitions and preliminaries

We shall denote by X an arbitrary nonempty set, and by $x, y, z, ...$ its elements.

Definition 1.2.1 (axioms concerning binary relations)**.** A binary relation R on a nonempty set X (i.e. a subset of the cartesian product $X \times X$) is said to be[(a)]

1. *reflexive*, if xRx for all $x \in X$,
2. *irreflexive*, if *not*(xRx) for every $x \in X$,
3. *transitive*, if (xRy) *and* (yRz) imply xRz for all $x, y, z \in X$,
4. *negatively transitive*, if *not*(xRy) *and* *not*(yRz) imply *not*(xRz) for all $x, y, z \in X$,
5. *symmetric*, if xRy implies yRx for all $x, y \in X$,
6. *asymmetric*, if xRy implies *not*(yRx) for all $x, y \in X$,
7. *antisymmetric*, if (xRy) *and* (yRx) imply $x = y$ for all $x, y \in X$,
8. *acyclic*, if $x_0 \, R \, x_1 \, R \, x_2 \, R \, x_3 \, R ... R \, x_{n-1} \, R \, x_n$ imply *not*$(x_n \, R \, x_0)$ for all $n \geq 1$, and for all $x_0, ..., x_n \in X$,
9. *total*, if (xRy) *or* (yRx) for all $x, y \in X$,
10. *complete*, if(xRy) *or* (yRx) for all $x, y \in X$ such $x \neq y$.

The pair (X, R) will be referred to as a *related set*.

[(a)] In what follows, given a binary relation R on a set X, for any two elements $x, y \in X$ we shall write xRy instead of $(x, y) \in R$.

Remark 1.2.2. In the sequel, R will be interpreted, loosely speaking, as a *(weak) preference relation.* Therefore, the scripture xRy has to be read as; *the element x (weakly) goes before the element y*, when the attribute “weak" is naturally associated to possible reflexivity of the binary relation.

The reader is invited to furnish a detailed proof, on his own, of the following simple proposition, in order to become familiar with the axioms concerning binary relations.

Remark 1.2.3. Notice that a binary relation R on a set X is negatively transitive if and only if, for all $x, y, z \in X$, xRz implies (xRy) *or* (yRz).

Proposition 1.2.4. *For any given related set* (X, R)*, the following implications hold true:*

(i) *If* R *is total, then it is reflexive;*

(ii) *If* R *is asymmetric, then it is irreflexive;*

(iii) *If* R *is irreflexive and transitive, then it is acyclic;*

(iv) *If* R *is acyclic, then it is asymmetric and not necessarily transitive.*

Definition 1.2.5 (lower and upper sections)**.** If (X, R) is a a related set, then define, for every $x \in X$,

$$L_R(x) = \{z \in X : zRx\}, \;\; U_R(x) = \{z \in X : xRz\}. \tag{1.2.1}$$

$L_R(x)$ and $U_R(x)$ are said to be the *lower section* and respectively the *upper section of the element* $x \in X$ *according to the binary relation* R. When there is no ambiguity about the binary relation involved, the subscript R will be omitted, and we shall simply write $L(x)$ and $U(x)$. Moreover, define

$$L_R = \{L_R(x) : x \in X\}, \;\; U_R = \{U_R(x) : x \in X\}, \tag{1.2.2}$$

so that L_R (U_R) is the family of all the lower sections (respectively, upper sections) associated to the binary relation R on X.

As usual, reflexive and irreflexive binary relations will be denoted by $\precsim$, and respectively by $\prec$.

Definition 1.2.6 (preorders)**.** A *preorder* $\precsim$ on a nonempty set X is a binary relation on X which is reflexive and transitive. If in addition $\precsim$ is antisymmetric, then we shall refer to $\precsim$ as an *order*. If X is a nonempty set, and $\precsim$ is a preorder (order) on X, then the related set $(X, \precsim)$ will be referred to as a *preordered set* (respectively, an *ordered set*).

Definition 1.2.7 (total preorders)**.** A preordered set $(X, \precsim)$ is said to be *totally preordered* if the preorder $\precsim$ on X is *total* (see point 9 in Derfinition 1.2.1) .

Definition 1.2.8 (indifference, strict preference and incomparability)**.** Given a preorder $\precsim$ on a set X, define, for every $x, y \in X$, the binary relations $\sim$ (*indifference relation* , or *symmetric part*), $\prec$ (*strict preference relation*, or *asymmetric part*) and $\bowtie$ (*incomparability relation*):

$$x \sim y \Leftrightarrow (x \precsim y) \text{ and } (y \precsim x), \tag{1.2.3}$$

$$x \prec y \Leftrightarrow (x \precsim y) \text{ and } \text{ not } (y \precsim x), \tag{1.2.4}$$

$$x \bowtie y \Leftrightarrow not(x \precsim y) \text{ and } \text{not}(y \precsim x). \tag{1.2.5}$$

Remark 1.2.9. Clearly, the indifference relation $\sim$ associated to any preorder $\precsim$ on a set X is an *equivalence* on X (i.e. $\sim$ is reflexive, transitive and symmetric). The strict part $\prec$ of any preorder $\precsim$ on a set X is *acyclic*, i.e. it satisfies the following property for all elements $x_0, ..., x_n \in X$ and every positive integer $n > 1$:

$$(x_0 \prec x_1) \text{ } and \text{ } (x_2 \prec x_3) \text{ } and \text{ } ... \text{ } and \text{ } (x_{n-1} \prec x_n) \Rightarrow not(x_n \, R \, x_0).$$

Definition 1.2.10 (jumps in a preordered set)**.** Given an ordered set $(X, \precsim)$, a pair (x, y) of elements of X is said to be a *jump* in $(X, \precsim)$ if $x \prec y$, and for no $z \in X$ it happens that $x \prec z \prec y$.

Therefore, a jump (x, y) in an ordered set $(X, \precsim)$ is such that $x \prec y$ and there is no point $z \in X$ which lays strictly between x and y.

Definition 1.2.11 (partial orders)**.** A *strict partial order* $\prec$ on a nonempty set X is a binary relation on X which is irreflexive and transitive. In this case, the related set $(X, \prec)$ will be referred to as a *strictly partially ordered set.*

Definition 1.2.12 (associated order)**.** Given any partial order $\prec$ on a set X, define a binary relation $\preceq$ on X by $x \preceq y$ if and only if either $x \prec y$ or $x = y$ $(x, y \in X)$. Then $\preceq$ is an order on X according to definition 1.2.6.

Definition 1.2.13 (weak orders)**.** A *weak order* $\prec$ on a nonempty set X is a binary relation on X which is asymmetric and negatively transitive. In this case, the pair $(X, \prec)$ will be referred to as a *weakly ordered set.*

Remark 1.2.14. Notice that the negative transitivity property (see Definition 1.2.1, point 4) can be written as follows:

$$\text{For all } x, y, z \in X, \;\; x \prec z \Rightarrow (x \prec y) \;\; or \;\; (y \prec z).$$

Remark 1.2.15. It is easily seen that a binary relation $\prec$ on a set X is a weak order in case that there exists a real-valued function u on X such that, for all $x, y \in X$,

$$x \prec y \Leftrightarrow u(x) < u(y).$$

Proposition 1.2.16 (weak orders are transitive)**.** *If $\prec$ is a weak order on a set X, then $\prec$ is transitive.*

Proof. By contraposition, assume that $\prec$ is an asymmetric and nontransitive binary relation on a set X. Then there exists three elements $x, y, z \in X$ such that $x \prec y \prec z$ and $not(x \prec z)$. Since $\prec$ is asymmetric, we have that $x \prec y$ implies $not(y \prec x)$. Therefore, it happens that $not(y \prec x)$ and $not(x \prec z)$, but at the same time $y \prec z$. Hence, $\prec$ is not negatively transitive. This observation completes the proof. □

Definition 1.2.17 (linear orders)**.** A *linear order* $\prec$ on a nonempty set X is a complete partial order (weak order) on X. In this case, the pair $(X, \prec)$ will be referred to as a *linearly ordered set.*

If $(X, \precsim)$ is any preordered set, then the associated strict preference relation $\prec$ is a partial order on X.

Definition 1.2.18 (incomparability relation)**.** Given a partial order $\prec$ on a set X, define, for every $x, y \in X$,

$$x \precsim y \Leftrightarrow not(y \prec x), \tag{1.2.6}$$

$$x \sim y \Leftrightarrow \neg(x \prec y) \text{ and } \neg(y \prec x). \tag{1.2.7}$$

The binary relations $\precsim$ and $\sim$ defined above will be called the *preference-indifference relation* and the *incomparability relation* associated to the partial order $\prec$.

Definition 1.2.19 (quotient order)**.** If $(X, \precsim)$ is any (totally) preordered set, denote by $X_{|\sim}$ the *quotient set modulo the equivalence relation* $\sim$, and define a binary relation $\precsim_{|\sim}$ on $X_{|\sim}$ in the following way:

$$[x] \precsim_{|\sim} [y] \Leftrightarrow x \precsim y.$$

Then it is easily seen that $(X_{|\sim}, \prec_{|\sim})$ is a (totally) ordered set.

Proposition 1.2.20 (weak orders and total preorders)**.** *Let $(X, \precsim)$ be a totally preordered set. Then the asymmetric part $\prec$ of $\precsim$ is a weak order on X. Conversely, if $(X, \prec)$ is a weakly ordered set, then the preference-indifference relation $\precsim$ associated to $\prec$ is a total preorder on X.*

Proof. Let $(X, \precsim)$ be a totally preordered set, and consider the strict preference $\prec$ defined in (1.2.4). Since it is clear that $\prec$ is asymmetric, let us show that $\prec$ is negatively transitive. Consider $x, y, z \in X$ such that $\neg(x \prec y)$ *and* $not(y \prec z)$. Then, using the fact that $\precsim$ is total, we obtain $(y \precsim x)$ *and* $(z \precsim y)$, which in turn implies $z \precsim x$ since $\precsim$ is transitive. Therefore $x \prec z$ is contradictory.
Conversely, let $(X, \prec)$ be a weakly ordered set, and consider the preference-indifference relation $\precsim$ defined in (1.2.6). Since $\prec$ is irreflexive, it is clear that $\precsim$ is reflexive. Observe that transitivity of $\precsim$ is equivalent to negative transitivity of $\prec$. Finally, let us show that $\precsim$ is total. Assume that there

exist two elements $x, y \in X$ such that $\neg(x \precsim y)$ *and* $\neg(y \precsim x)$. Then we have that $(y \prec x)$ *and* $(x \prec y)$, and this is contradictory since $\prec$ is transitive and asymmetric (therefore, irreflexive). □

Definition 1.2.21 (decreasing subsets of a preordered set)**.** A subset E of a preordered set $(X, \precsim)$ is said to be *decreasing* (*increasing*) if $b \in E$, $a \precsim b$ imply $a \in E$ ($a \in E$, $a \precsim b$ imply $b \in E$).

Given any preordered set $(X, \precsim)$, and a decreasing (increasing) subset E of X, it is easy to check that the set $X \setminus E$ is increasing (decreasing). For example, assume that $E \subset X$ is decreasing, and consider two elements $x, y \in X$ such that $x \in X \setminus E$, $x \precsim y$. Then it must be $y \in X \setminus E$. Otherwise we have $x \in E$ since E is decreasing.

Definition 1.2.22 (decreasing set generated by a subset)**.** Given a preordered set $(X, \precsim)$, and a set $E \subset X$, denote by $D(E)$ ($I(E)$) the intersection of all the decreasing (increasing) subsets of X containing E (i.e., $D(E)$ ($I(E)$) is the smallest decreasing (increasing) subset of X containing E). Then, by definition 1.2, it is $D(\{x\}) = L(x)$, $I(\{x\}) = U(x)$.

Proposition 1.2.23 (lower sections in a totally preordered set)**.** *Let* $(X, \precsim)$ *be a preordered set. Then the preorder* $\precsim$ *on* X *is a total if and only if* $L_{\precsim}$ *is totally ordered by set inclusion (i.e., for every pair* $(x, y) \in X \times X$*, either* $L_{\precsim}(x) \subset L_{\precsim}(y)$*, or* $L_{\precsim}(y) \subset L_{\precsim}(x)$*.*

Proof. Consider any preordered set $(X, \precsim)$. If $\precsim$ is total, then for two elements $x, y \in X$ either $x \precsim y$ or $y \precsim x$. Hence, by transitivity of $\precsim$, either $L_{\precsim}(x) \subset L_{\precsim}(y)$ or $L_{\precsim}(y) \subset L_{\precsim}(x)$, and therefore $L_{\precsim}$ is totally ordered by set inclusion. In order to show that if $L_{\precsim}$ is totally ordered by set inclusion then the preorder $\precsim$ on X is total, assume by contraposition that $\precsim$ is not total. Then there exist two elements $x, y \in X$ such that neither $x \precsim y$ nor $y \precsim x$. Hence we have $x \notin L_{\precsim}(y)$ and $y \notin L_{\precsim}(x)$. Since it is clear that $x \in L_{\precsim}(x)$ and $y \in L_{\precsim}(y)$ by reflexivity of $\precsim$, $L_{\precsim}$ is not totally ordered by set inclusion, since neither $L_{\precsim}(x) \subset L_{\precsim}(y)$, nor $L_{\precsim}(y) \subset L_{\precsim}(x)$. This consideration completes the proof. □

1.3 Order-preserving functions and utility functions

In this section we shall consider (totally) preordered sets $(X, \precsim)$.

Definition 1.3.1 (increasing function)**.** Given a preordered set $(X, \precsim)$, a function $u : (X, \precsim) \to (\mathbb{R}, \leq)$ is said to be a *(real-valued) increasing function* on $(X, \precsim)$ if, for all points $x, y \in X$,

$$x \precsim y \Rightarrow u(x) \leq u(y).$$

The existence of a real-valued increasing function u on a preordered set $(X, \precsim)$ does not give enough information on the preorder $\precsim$. Indeed, given any constant real-valued function u on an arbitrary set X, for every preorder $\precsim$ on X we have that u is increasing on the preordered set $(X, \precsim)$.

Definition 1.3.2 (order-preserving function)**.** Given a preordered set $(X, \precsim)$, a function $u : (X, \precsim) \to (\mathbb{R}, \leq)$ is said to be a *(real-valued) order-preserving function* on $(X, \precsim)$ if it is increasing on $(X, \precsim)$ and, for all points $x, y \in X$,

$$x \prec y \Rightarrow u(x) < u(y).$$

Definition 1.3.3 (utility function)**.** Given a preordered set $(X, \precsim)$, a function $u : (X, \precsim) \to (\mathbb{R}, \leq)$ is said to be a *utility function* on $(X, \precsim)$ if, for all points $x, y \in X$,

$$x \precsim y \Leftrightarrow u(x) \leq u(y). \qquad (1.3.1)$$

Remark 1.3.4. Clearly, if a function $u : (X, \precsim) \to (\mathbb{R}, \leq)$ is a utility function on $(X, \precsim)$, then $\precsim$ is a total preorder on X.

The simple proof of the following propositions are left to the reader.

Proposition 1.3.5. *Let $(X, \precsim)$ be a totally preordered set. Then the following conditions are equivalent on a function $u : (X, \precsim) \to (\mathbb{R}, \leq)$:*

1. *u is an order-preserving function on $(X, \precsim)$;*
2. *u is a utility function on $(X, \precsim)$*

Proposition 1.3.6. *Let $(X, \precsim)$ be a totally preordered set. Then a function $u : (X, \precsim) \rightarrow (\mathbb{R}, \leq)$ is a utility function on $(X, \precsim)$ is and only if the following conditions are verified for all points $x, y \in X$:*

1. $x \sim y \Rightarrow u(x) = u(y)$;
2. $x \prec y \Rightarrow u(x) < u(y)$.

Let us present an example of a nontotal preorder admitting a real representation by means of an order-preserving function.

Example 1.3.7. Let X be the real interval $[0, 1]$ and consider the nontotal preorder $\precsim$ on X defined as follows:

$$x \precsim y \Leftrightarrow \begin{cases} x \leq y \text{ and } x, y \in \mathbb{Q} \cap [0, 1] \\ \qquad \text{or} \\ x \leq y \text{ and } x, y \in [0, 1] \setminus \mathbb{Q} \end{cases},$$

where $\mathbb{Q}$ is the set of all the rational numbers. It is immediate to check that the identity function $u = i_X$ is an order-preserving function for $\precsim$ on X. On the other hand, we have that $x \bowtie y$ for all pairs

$$(x, y) \in (\mathbb{Q} \cap [0, 1]) \times ([0, 1] \setminus \mathbb{Q}),$$

and for all pairs

$$(x, y) \in ([0, 1] \setminus \mathbb{Q}) \times (\mathbb{Q} \cap [0, 1]).$$

Actually, the existence on order-preserving function $u : (X, \precsim) \rightarrow (\mathbb{R}, \leq)$ does not require $\precsim$ to be transitive (this is clearly the case of the binary relation considered in the previous example). Therefore, the following examples presents a reflexive and nontransitive binary relation admitting an order-preserving function.

Example 1.3.8. Consider the binary relation $\precsim$ on the real interval $X = [0,1]$ defined as follows:

$$x \precsim y \Leftrightarrow \begin{cases} (x \leq y) \text{ and } (x, y \in [0, \frac{1}{2}]) \\ \text{or} \\ (x \leq y) \text{ and } (x, y \in [\frac{1}{2}, 1]) \\ \text{or} \\ (x \leq \frac{1}{4}) \text{ and } (y \geq \frac{3}{4}) \end{cases}$$

Clearly, $\precsim$ is reflexive. Notice that $x \bowtie y$ for all $x, y \in [0,1]$ such that $\frac{1}{4} < x \leq \frac{1}{2} < y < \frac{3}{4}$. We have that $\precsim$ is not transitive since, for example, $\frac{1}{3} \precsim \frac{1}{2} \precsim \frac{3}{5}$ but $\frac{1}{3} \bowtie \frac{3}{5}$. On the other hand, it is easily seen that the identity function $id_{[0,1]}$ on $[0,1]$ is a order-preserving function for $\precsim$.

Given a real-valued order-preserving function u on a preordered set $(X, \precsim)$, it is clear that the composition $u' = \phi \circ u$ of u with any strictly increasing (i.e., order-preserving) function $\phi : (u(X), \leq) \to (\mathbb{R}, \leq)$ is also a real-valued order-preserving function on $(X, \precsim)$.

Recall that a set $\mathcal{D}$ is countable if there exists a bijection $f : \mathcal{D} \to \mathbb{N}'$ for some set $\mathbb{N}' \subset \mathbb{N}$. Clearly, $\mathbb{N}$ stands for the set of the *natural numbers* ($\mathbb{N}^+$ stands for the set of the positive integers).

Proposition 1.3.9 (jumps and order-preserving functions)**.** *Given a totally ordered set $(X, \precsim)$, if there exists a real-valued order-preserving function u on $(X, \precsim)$, then there are only countably many jumps in $(X, \precsim)$.*

Proof. Denote by $\mathcal{J}$ the subset of $X \times X$ consisting of all the jumps in $(X, \precsim)$. Clearly, for every $(x,y), (x',y') \in \mathcal{J}$ such that $(x,y) \neq (x',y')$, we have that the nonempty real intervals $]u(x), u(y)[$ and $]u(x'), u(y')[$ are disjoint. Therefore, we may associate to each $(x,y) \in \mathcal{J}$ a rational number p such that $u(x) < p < u(y)$. Hence, $\mathcal{J}$ is countable, since the set $\mathbb{Q}$ of all the rational numbers is countable. □

Definition 1.3.10 (order-separability)**.** A preorder $\precsim$ on a set X is said to be

(i) *order-separable* if there exists a countable set $\mathcal{D} \subset X$ such that for every $x, y \in X$ with $x \prec y$ there exists $d \in \mathcal{D}$ such that $x \prec d \prec y$;

(ii) *weakly order-separable* if there exists a countable set $\mathcal{D} \subset X$ such that for every $x, y \in X$ with $x \prec y$ there exist $d_1, d_2 \in \mathcal{D}$ such that $x \precsim d_1 \prec d_2 \precsim y$.

$\mathcal{D}$ is said to be a (*weakly*) *order-dense* subset of X.

Clearly, if a preorder $\precsim$ on a set X is order-separable, then it is also weakly order-separable, while the converse is not true.

Example 1.3.11 (weak order-separability, not order-separability)**.** Define $X = [0,1] \cup [2,3]$ $(\subset \mathbb{R})$, and endow X with the induced natural total preorder $\leq$. Then the total preorder $\leq$ on X is weakly order-separable, and $\mathcal{D} = ([0,1] \cup [2,3]) \cap \mathbb{Q}$ is a countable weakly order-dense subset of X. On the other hand, $\precsim$ is not order-separable, since $1 \prec 2$ and for no $x \in X$ it is $1 \prec x \prec 2$.

Proposition 1.3.12 (condition for the existence of an order-preserving function)**.** *If a preorder $\precsim$ on set X is weakly order-separable, then there exists a real-valued order-preserving function u on $(X, \precsim)$ (with values in $[0,1]$).*

Proof. Consider a weakly order-separable preorder $\precsim$ on a set X, and let $\mathcal{D} = \{d_n : n \in \mathbb{N}^+\}$ be a weakly order-dense subset of X. Then define a real valued function u on X by letting

$$u(x) = \begin{cases} \sum\limits_{\{n \in \mathbb{N}^+ : d_n \precsim x\}} 2^{-n} & if \;\; d_n \precsim x \;\; for \;\; some \;\; n \in \mathbb{N}^+ \\ 0 & otherwise \end{cases} . \tag{1.3.2}$$

We claim that u is an order-preserving function on $(X, \precsim)$. Clearly, u is increasing, since for every $x, y \in X$ with $x \precsim y$, $d_n \precsim x$ implies $d_n \precsim y$.

Now consider $x, y \in X$ with $x \prec y$. From the definition of weak separability, for every $x, y \in X$ with $x \prec y$ there exists $d \in \mathcal{D}$ such that $x \prec d \precsim y$. Therefore, it must be $u(x) < u(y)$ from the definition of u. So the proof is complete. □

We are now ready to present a characterization of the existence of a utility function on a totally preordered set.

Theorem 1.3.13 (utility on a totally preordered set)**.** *Let $(X, \precsim)$ be a totally preordered set. Then the following conditions are equivalent:*

(i) *There exists a utility function u on $(X, \precsim)$ with values in $[0, 1]$;*

(ii) *The total preorder $\precsim$ on X is weakly order-separable.*

Proof. By Proposition 1.3.12, it suffices to show that condition (i) implies condition (ii). Assume that there exists a utility function u on a totally preordered space $(X \precsim)$. Then $(X_{|\sim}, \precsim_{|\sim})$ is a totally ordered set with only countably many jumps (see Definition 1.2.10 and Proposition 1.3.9). For each jump $([x], [y])$ in $(X_{|\sim}, \precsim_{|\sim})$, consider the elements $x, y \in X$, and let $\mathcal{A}$ be the countable subset of X which is the union of all such points. Further, for each pair (p_i, q_i) of rational numbers in $[0, 1]$, such that $u^{-1}(]p_i, q_i[) \neq \emptyset$, consider an element $x \in u^{-1}(]p_i, q_i[)$, and define the countable set $\mathcal{B}$ of X as the union of all such elements. It is easily seen that $\mathcal{D} = \mathcal{A} \cup \mathcal{B}$ is a countable weakly order-dense subset of X. This consideration finishes the proof. □

We finish this section with the classical example of the *lexicographic order*, which is not representable by a utility function.

Example 1.3.14 (lexicographic order)**.** Consider the binary relation $\precsim$ on $X = \mathbb{R}_+ \times \mathbb{R}_+$ defined as follows, for all pairs (x, z), $(y, w) \in \mathbb{R}_+ \times \mathbb{R}_+$:

$$(x, z) \precsim (y, w) \Leftrightarrow \begin{cases} (x \leq y) \\ \text{or} \\ (x = y) \text{ and } (z \leq w) \end{cases} .$$

We have that $\precsim = \precsim_{lex}$ is a total (pre)order, the so called *lexicographic order* on $\mathbb{R}^2_+ = \mathbb{R}_+ \times \mathbb{R}_+$. Let us show that there is no utility function u on $(X, \precsim)$. Suppose, to the contrary, that there exists some utility function $u : \mathbb{R}^2_+ \to \mathbb{R}$ representing $\precsim = \precsim_{lex}$.

We thus have $u(x, 0) < u(x, 1)$, as $(x, 0) \prec (x, 1)$. We construct the real interval $I(x) = [u(x, 0), u(x, 1)]$. Consider that, for two distinct points $x, y \in \mathbb{R}_+$, we have that $I(x) \cap I(y) = \emptyset$, as, for example, $x < y$ implies that $(x, 1) \prec (y, 0)$.

Define $\mathbb{I} = \{I(x) : x \in \mathbb{R}_+\}$. Clearly, $\mathbb{I}$ is an uncountable set. We have that $\precsim$ cannot be weakly order-separable. Otherwise, if there exists a countable weakly order-dense subset $\mathcal{D}$ of X, we can associate an element $d(x) \in \mathcal{D}$ to every interval $I(x)$, so that $x \neq y \Leftrightarrow I(x) \neq I(y)$ implies $d(x) \neq d(y)$, a contradiction since $\mathbb{I}$ is an uncountable set. Hence, $\precsim$ does not admit a utility function by Theorem 1.3.13, and the proof is complete.□

.

Exercises

1. Show that the binary relation $\precsim$ defined as follows, for all $x, y \in \mathbb{R}$,

$$x \precsim y \Leftrightarrow x \leq y + 1$$

is total and nontransitive, while its *strict part* $\prec$ defined to be

$$x \prec y \Leftrightarrow not(y \prec x)$$

is nontotal and transitive (for example, $2 \precsim 1 \precsim 0$ and $2 \not\precsim 0$);

2. The same question as regards the binary relation $\precsim$ defined as follows, for all $x, y \in [0,1]$,
$$x \precsim y \Leftrightarrow x^2 \leq y$$
(for example, $\frac{1}{\sqrt{3}} \precsim \frac{1}{3} \precsim \frac{1}{9}$ and $\frac{1}{\sqrt{3}} \not\precsim \frac{1}{9}$);

3. Show that, in general, given any pair (u, v) of continuous real-valued functions on X such that $u \leq v$, the binary relation $\precsim$ on X defined to be
$$x \precsim y \Leftrightarrow u(x) \leq v(y)$$
is such that $L_{\precsim}(x)$ and $U_{\precsim}(x)$ are open subsets of X for every $x \in X$.

1.4 Continuous utility functions

We first recall the basic definitions and some preliminary results concerning generic topological spaces, and in particular topological related spaces.

Definition 1.4.1 (topological space)**.** A family τ of subsets of a nonempty set X is a *topology* on X if the following conditions are verified:

(i) $X, \emptyset \in \tau$;

(ii) τ is closed under arbitrary unions;

(iii) τ is closed under finite intersections.

The pair (X, τ) is said to be a *topological space*.

Definition 1.4.2 (open sets)**.** An element of τ is a (τ-)*open set*, and τ is the family of open subsets of X.

A classical type of a topological space is represented by a *metric space* (X, d). This is the case when the topology τ on X is induced by a *metric*, in the sense that there exists a metric $d : X \times X \to \mathbb{R}_+$ (i.e., d takes nonnegative real values on $X \times X$, and it satisfies the following conditions for all $x, y, z \in X$: 1. $d(x, y) = 0$ if and only if $x = y$; 2. $d(x, y) = d(y, x)$; 3. $d(x, y) + d(y, z) \geq d(x, z)$), and a set $O \subset X$ is declared to be open if and only if, for every $x \in O$ there exists a real number $r > 0$ such that
$$B_r(x) = \{z \in X : d(x, z) < r\} \subset O. \tag{1.4.1}$$

We say that $B_r(x)$ is the *open ball* centered at $x \in X$ with radius $r > 0$. In this case, $d(x, y)$ is said to be the *distance between x and y*. A classical example is represented by the *interval topology* τ_{int} on then real line $\mathbb{R}$, when

$$d(x, y) = \mid x - y \mid,$$

with $\mid \cdot \mid$ the *absolute value*. In other words, τ_{int} is the set of open real intervals along with their arbitrary unions and finite intersections.

Definition 1.4.3 (convergence of a sequence in a metric space)**.** Let (X, d) be a metric space. A sequence $\{x_n\}_{n \in \mathbb{N}} \subset X$ *converges* to a point $x \in X$ ($x_n \to x$) if

$$\lim_{n \to \infty} d(x_n, x) = 0.$$

Definition 1.4.4 (neighborhoods and closed sets)**.** If $x \in X$ belongs to a subset U of X, and there exists a set $V \in \tau$ such that $x \in V \subset U$, then U is a *neighborhood* of x. A subset U of X is (τ-)*closed* if its complement $X \setminus U$ is open.

Clearly, given any topological space (X, τ), the family of closed subsets of X is closed under finite unions and arbitrary intersections, and contains X and $\emptyset$.

Definition 1.4.5 (closure and dense sets)**.** Given a topological space (X, τ), the (*topological*) *closure* $\overline{U}$ of any subset U of X is the intersection of all the closed subsets of X containing U. A set $\mathcal{S} \subset X$ is said to be *dense* (in X) if $\overline{\mathcal{S}} = X$.

Definition 1.4.6 (continuous real-valued function)**.** A real-valued function u on an arbitrary topological space (X, τ) is said to be *continuous* if

$$u^{-1}(]-\infty, \alpha[) = \{x \in X : u(x) < \alpha\}, \;\; u^{-1}(]\alpha, +\infty[) = \{x \in X : \alpha < u(x)\}$$

are both open sets for every $\alpha \in \mathbb{R}$.

We state without proof the following proposition concerning the continuity of a real-valued function u on a metric space (X, d).

Proposition 1.4.7 (continuous real-valued function on a metric space). *A real-valued function u on a metric space (X, d) is continuous if and only if the following condition holds:*
For every sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ and every point $x \in X$

$$x_n \to x \Rightarrow u(x_n) \to u(x). \tag{1.4.2}$$

Definition 1.4.8 (continuity of a total preorder). A total preorder $\precsim$ on a topological space (X, τ) is said to be *continuous* if $L_{\prec}(x)$ and $U_{\prec}(x)$ are open subsets of X for every $x \in X$ (or, equivalently, $L_{\precsim}(x)$ and $U_{\precsim}(x)$ are closed subsets of X for every $x \in X$).

Remark 1.4.9 (continuity of a total preordey on a metric space). In the particular case when we consider a total preorder $\precsim$ on a metric space (X, τ), we have that $\precsim$ is continuous if and only if the following conditions hold for every sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ and all points $x,\ \bar{x} \in X$:

$$(x_n \to x)\ and (x_n \precsim \bar{x} \ \text{ for every } n \in \mathbb{N}) \Rightarrow x \precsim \bar{x}; \tag{1.4.3}$$

$$(x_n \to x)\ and (\bar{x} \precsim x_n \ \text{ for every } n \in \mathbb{N}) \Rightarrow \bar{x} \precsim x. \tag{1.4.4}$$

Lemma 1.4.10 (normality of a total preorder). *Let $(X, \tau, \precsim)$ be a topological totally preordered space, and assume that $\precsim$ is continuous. Then $\precsim$ satisfies the following property*[(a)]*:*

(*) Given a closed decreasing set F_0 and a closed increasing set F_1 with $F_0 \cap F_1 = \emptyset$, there exist an open decreasing set A_0, and an open increasing set A_1 such that $A_0 \supset F_0$, $A_1 \supset F_1$, $A_0 \cap A_1 = \emptyset$ (this is equivalent to require that, given a closed decreasing set F_0, and an open decreasing set A_0 such that $F_0 \subset A_0$, there exist an open decreasing set A_1 and a closed decreasing set F_1, such that $F_0 \subset A_1 \subset F_1 \subset A_0$).

[(a)] A preorder on a topological space satisfying property (*) is said to be *normal*

Proof. Given a topological totally preordered space $(X, \tau, \precsim)$ with $\precsim$ continuous, consider a closed decreasing set F_0 and a closed increasing set F_1 with $F_0 \cap F_1 = \emptyset$. If $F_0 \cup F_1 = X$, then it is clear that $A_0 = F_0$ is an open decreasing set, $A_1 = F_1$ is an open increasing set, and $A_0 \cap A_1 = \emptyset$. Now assume that there exists a point $x \in X \setminus (F_0 \cup F_1)$. Since $\precsim$ is total, it must be $z \prec x$ for every $z \in F_0$ and $x \prec z$ for every $z \in F_1$. Hence, $A_0 = L_{\prec}(x)$ is an open decreasing set containing F_0, $A_1 = U_{\prec}(x)$ is an open increasing set containing F_1 and clearly $A_0 \cap A_1 = \emptyset$. So the proof is complete. □

Definition 1.4.11 (decreasing scale)**.** Given a topological preordered space $(X, \tau, \precsim)$, a family $\mathcal{A} = \{A_r : r \in \mathcal{S}\}$ of open decreasing subsets of X is said to be a *decreasing scale* in $(X, \tau, \precsim)$ if the following conditions are satisfied:

(i) $\mathcal{S}$ is a dense subset of $[0, 1]$ such that $1 \in \mathcal{S}$ and $A_1 = X$;

(ii) For every $r_1, r_2 \in \mathcal{S}$ with $r_1 < r_2$, it is $\overline{A_{r_1}} \subset A_{r_2}$.

Proposition 1.4.12 (characterization of continuity of a total preorder)**.** *Let* $(X, \tau, \precsim)$ *be a totally preordered topological space. Then the following conditions are equivalent:*

(i) *For every* $x, y \in X$ *such that* $x \prec y$ *there exists a real-valued continuous increasing function* $u_{x,y}$ *on* $(X, \tau, \precsim)$ *with values in* $[0, 1]$ *such that* $u_{x,y}(x) = 0$ *and* $u_{x,y}(y) = 1$*;*

(ii) $\precsim$ *is continuous on* (X, τ).

Proof. (i) ⇒ (ii). Let $(X, \tau, \precsim)$ be any totally preordered topological space, and assume that condition (i) is verified. In order to show that $L_{\prec}(x)$ is open for every $x \in X$, consider any point $z \in L_{\prec}(x)$. By condition (i), there exists a continuous increasing function $u_{z,x}$ on $(X, \tau, \precsim)$ with values in $[0, 1]$ such that $u_{z,x}(z) = 0$ and $u_{z,x}(x) = 1$. Then $u_{z,x}^{-1}([0, u_{z,x}(x)[)$ is an open subset of $L_{\prec}(x)$ containing z, and therefore $L_{\prec}(x)$ is an open set. Analogously it can be shown that $U_{\prec}(x)$ is an open set for every $x \in X$. □

(ii) ⇒ (i). Let $\precsim$ be a continuous total preorder on (X, τ). Let us show that $\precsim$ satisfies the above condition (i). Consider any pair $(x, y) \in X \times X$ such that $x \prec y$. Recall that $L_{\precsim}(x)$ and $U_{\precsim}(y)$ are both closed sets, since $\precsim$ is continuous on (X, τ). In addition, it is clear that $L_{\precsim}(x) \cap U_{\precsim}(x) = \emptyset$.

From now on consider that property (*) of Lemma 1.4.10 holds true. Set $F_0 = L_{\precsim}(x)$, and notice that $X \setminus U_{\precsim}(y) = L_{\prec}(y)$ contains F_0. Since F_0 is a closed decreasing set, and $L_{\prec}(y)$ is an open decreasing set, by normality of the topological preordered space $(X, \tau, \precsim)$ there exist an open decreasing set $A_{\frac{1}{2}}$, and a closed decreasing set $F_{\frac{1}{2}}$ such that $F_0 \subset A_{\frac{1}{2}} \subset F_{\frac{1}{2}} \subset L_{\prec}(y)$. Similar considerations lead to the existence of open decreasing sets $A_{\frac{1}{4}}$, $A_{\frac{3}{4}}$, and closed decreasing sets $F_{\frac{1}{4}}$, $F_{\frac{3}{4}}$ such that $F_0 \subset A_{\frac{1}{4}} \subset F_{\frac{1}{4}} \subset A_{\frac{1}{2}} \subset F_{\frac{1}{2}} \subset A_{\frac{3}{4}} \subset F_{\frac{3}{4}} \subset L_{\prec}(y)$. So one obtains two families $\{F_s\}$ and $\{A_s\}$, where $s = \frac{p}{2^q}$ is any dyadic rational number with $p = 1, 2, ..., 2^q - 1$, $q = 1, 2, ...$ Set $A_1 = X$, $\mathcal{A} = \{A_s\} \cup A_1$. It is clear that the collection of all such dyadic rationals is dense in $[0, 1]$ (i.e., $0 \leq x < y \leq 1 \Rightarrow \exists s = \frac{p}{2^q}$ such that $x < \frac{p}{2^q} < y$). Further, since F_s is closed for all s, we have that $\overline{A_{s_1}} \subset A_{s_2}$ for all dyadic rationals s_1, s_2 with $s_1 < s_2$. Hence, $\mathcal{A}$ is a decreasing scale in $(X, \tau, \precsim)$ (see Definition 1.4.11), such that $F_0 \subset A_s$, $U_{\precsim}(y) \subset X \setminus A_s$ for all dyadic rationals previously defined.

Define $\mathcal{S}$ the set of all this dyadic rationals, plus 1, and then define the function $u_{x,y} : X \to [0, 1]$ as follows, for every $z \in X$:

$$u_{x,y}(z) = \inf\{r \in \mathcal{S} : z \in A_r\}$$

Let us show that $u_{x,y}$ is an increasing and continuous function on $(X, \tau, \precsim)$. In order to prove that $u_{x,y}$ is an increasing function on the preordered set $(X, \precsim)$, consider any $x, y \in X$ such that $x \precsim y$. Then, since A_r is a decreasing set for every $r \in \mathcal{S}$, we have $\{r \in \mathcal{S} : y \in A_r\} \subset \{r \in \mathcal{S} : x \in A_r\}$, and therefore $u_{x,y}(x) \leq u_{x,y}(y)$ from the definition of $u_{x,y}$.

In order to prove that $u_{x,y}$ is continuous on (X, τ), let us first show that actually, for every $z \in X$:

$$u_{x,y}(z) = inf\{r \in \mathcal{S} : z \in \overline{A_r}\}.$$

It is clear that, for every $z \in X$, $\inf\{r \in \mathcal{S} : z \in \overline{G_r}\} \leq \inf\{r \in \mathcal{S} : z \in G_r\}$, since $\{r \in \mathcal{S} : z \in A_r\} \subset \{r \in \mathcal{S} : z \in \overline{A_r}\}$. Assume that there exists $z \in X$ such that $\inf\{r \in \mathcal{S} : z \in \overline{A_r}\} < inf\{r \in \mathcal{S} : z \in A_r\}$. Now consider $r_1, r_2 \in \mathcal{S}$ such that $\inf\{r \in \mathcal{S} : z \in \overline{A_r}\} < r_1 < r_2 < \inf\{r \in \mathcal{S} : z \in A_r\}$. Then we have $z \in \overline{A_{r_1}}$ and $z \notin A_{r_2}$, and this is contradictory since $\mathcal{A} = \{A_r : r \in \mathcal{S}\}$ is a decreasing scale in $(X, \tau, \precsim)$ and therefore $\overline{A_{r_1}} \subset A_{r_2}$.

Now, consider any element $x \in X$, and any real number $\alpha \leq 1$, such that $u_{x,y}(x) < \alpha$. From the definition of $u_{x,y}$, there exists $\bar{r} \in \mathcal{S}$ such that $u_{x,y}(x) < \bar{r} < \alpha$, $x \in A_{\bar{r}}$. Then, $A_{\bar{r}}$ is an open subset of X such that $u_{x,y}(z) < \alpha$ for every $z \in A_{\bar{r}}$, since $z \in A_{\bar{r}}$ entails $u_{x,y}(z) \leq \bar{r}$.

Now, consider any element $z \in X$, and any real number $\alpha \geq 0$, such that $\alpha < u_{x,y}(z)$. Let $r \in \mathcal{S}$ be such that $\alpha < r < u_{x,y}(z)$. Then, from considerations above, $u_{x,y}(z') \geq r$ for every $z' \in X \setminus \overline{A_r}$, since $u_{x,y}(z') < r$ entails $z' \in \overline{A_r}$. Hence, $X \setminus \overline{A_r}$ is an open set such that $z \in X \setminus \overline{A_r}$ and $u_{x,y}(z') > \alpha$ for every $z' \in X \setminus \overline{A_r}$.

Finally, we have that $u_{x,y}(x) = 0$, and $u_{x,y}(y) = 1$ from the definition of $u_{x,y}$, since $L_{\precsim}(x) \subset A_r$ for every $r \in \mathcal{S}$, and $A_r \subset L_{\prec}(x) = X \setminus U_{\precsim}(y)$ for every $r \in \mathcal{S} \setminus \{1\}$. So the proof is complete. □

Theorem 1.4.13 (continuous utility for a total preorder)**.** *Let $(X, \tau, \precsim)$ be a topological totally preordered space. Then the following conditions are equivalent:*

(i) *There exists a continuous utility function u on $(X, \tau, \precsim)$ with values in $[0, 1]$;*

(ii) *The total preorder $\precsim$ on (X, τ) is weakly order-separable and continuous.*

Proof. (i) ⇒ (ii). Assume that there exists a real-valued continuous order-preserving function u on $(X, \tau, \precsim)$ with values in $[0, 1]$. Then the total preorder $\precsim$ on X is weakly order-separable by Theorem 1.3.13. Further, $\precsim$ is continuous, since $L_{\prec}(x) = \{y \in X : y \prec x\} = u^{-1}([0, u(x)[)$ and $U_{\prec}(x) = \{y \in X : x \prec y\} = u^{-1}(]u(x), 1])$ are open sets for every $x \in X$.
(ii) ⇒ (i). Consider a weakly order-separable and continuous total preorder $\precsim$ on a topological space (X, τ). Let $\mathcal{D} = \{d_n : n \in \mathbb{N}^+\}$ be a countable weakly order-dense subset in $(X, \precsim)$. Then, by Proposition 1.4.12, for every $d_m, d_n \in \mathcal{D}$ such that $d_m \prec d_n$ there exists a continuous increasing function u_{d_m,d_n} with values in $[0, 1]$ such that $u_{d_m,d_n}(d_m) = 0$ and $u_{d_m,d_n}(d_n) = 1$. Since $\mathcal{D}$ is countable, it is clear that there are at most countably many pairs $(d_m, d_n) \in \mathcal{D} \times \mathcal{D}$ such that $d_m \prec d_n$. Denote by $Z = \{z_n : n \in \mathbb{N}^+\}$ the set whose elements are all such pairs, and denote by u_n the continuous increasing function with values in $[0, 1]$ which separates the corresponding elements of $\mathcal{D}$. Hence, $u = \sum_{n \in \mathbb{N}^+} 2^{-n} u_n$ is a real-valued continuous utility function for $\precsim$ with values in $[0, 1]$. Indeed, for every $x, y \in X$ such that $x \prec y$ there exists $m, n \in \mathbb{N}^+$ such that $x \precsim d_m \prec d_n \precsim y$, and therefore there exists $n \in \mathbb{N}^+$ such that $u_n(x) = 0$, and $u_n(y) = 1$. □

Let us finish this section by presenting the two most popular theorems on the existence of continuous utility representations for continuous total preorders, i.e. the *Eilenberg theorem* and the *Debreu theorem*. Some simple and widely used definitions are needed.

Definition 1.4.14 (connected topology)**.** A topology τ on X is said to be *connected* if X cannot be partitioned into two (nonempty) subsets which are both closed (or equivalently, if X and $\emptyset$ are the only subsets of X which are at the same time open and closed).

Definition 1.4.15 (separable topology)**.** A topology τ on X is said to be *separable* if there exists a countable *dense subset* $\mathcal{D}$ of X (i.e., a countable set $\mathcal{D} \subset X$ such that $\mathcal{D} \cap O \neq \emptyset$ for every $O \in \tau$).

Theorem 1.4.16 (Eilenberg theorem)**.** *Every continuous total preorder on a connected and separable topological space* (X, τ) *is representable by a continuous utility function* u*.*

Proof. By Theorem 1.4.13, we only need to prove that the total preorder $\precsim$ on X is (weakly) order-separable. Since the topology τ on X is separable, consider a countable dense subset $\mathcal{D}$ of X. Consider any pair $x \prec y$. Then $L_{\precsim}(x)$ and $U_{\precsim}(y)$ are closed and disjoint, and therefore the order interval $X \setminus (L_{\precsim}(x) \cup U_{\precsim}(y)) = U_{\prec}(x) \cap L_{\prec}(y)$ is nonempty, due to the fact that τ is a connected topology. Therefore, there exists $d \in \mathcal{D}$ such that $d \in U_{\prec}(x) \cap L_{\prec}(y) \Leftrightarrow x \prec d \prec y$. These arguments shows that $\mathcal{D}$ is a countable order-dense subset of $(X, \precsim)$, and therefore that $\precsim$ is order-separable on X. This consideration completes the proof. □

Definition 1.4.17 (basis of a topological space)**.** If τ is a *topology* on X, then a family $\mathcal{B} \subset \tau$ is said to be a *basis* of τ if every set $O \in \tau$ is the union of some sets of $\mathcal{B}$.

Definition 1.4.18 (second countable topology)**.** A topology τ on X is said to be *second countable* if there is a countable basis $\mathcal{B} = \{B_n : n \in \mathbb{N}^+\}$ of τ.

A relevant example of a second countable topology is furnished by a separable metric space, as the following proposition shows.

Proposition 1.4.19. *A separable metric space* (X, d) *is second countable.*

Proof. Just consider that, if $\mathcal{D}$ is a countable dense subset in (X, d), then $\mathcal{B} = \{B_r(d) : d \in \mathcal{D},\ r \in \mathbb{Q}\}$ is a countable basis of (X, d) (see Definition 1.4.1). □

Theorem 1.4.20 (Debreu theorem)**.** *Every continuous total preorder on a second countable topological space (X, τ) is representable by a continuous utility function u.*

Proof. We only need to show that there exists a utility function u' on $(X, \precsim)$. Then, we have that the total $\precsim$ on X is weakly order-separable by Theorem 1.3.13, and the thesis follows from Theorem 1.4.13. Let $\mathcal{B} = \{B_n : n \in \mathbb{N}^+\}$ be a countable basis of τ. Then define a real valued function u on X by letting

$$u'(x) = \begin{cases} \sum_{\{n \in \mathbb{N}^+ : B_n \subset L_{\prec}(x)\}} 2^{-n} & if \;\; L_{\prec}(x) \neq \emptyset \\ 0 & otherwise \end{cases} . \tag{1.4.5}$$

We claim that u' is an order-preserving function on $(X, \precsim)$. Clearly, u' is increasing, since for every $x, y \in X$ with $x \precsim y$, it happens that $L_{\prec}(x) \subset L_{\prec}(y)$. If we now consider $x, y \in X$ with $x \prec y$, we have that $L_{\prec}(x) \subsetneqq L_{\prec}(y)$, and therefore there exists $B \in \mathcal{B}$ such that $B \subset L_{\prec}(y)$, $not(B \subset L_{\prec}(x))$. This clearly implies that $u'(x) < u'(y)$ from the definition of u'. So the proof is complete. □

Bibliography

[1] Bridges, D.S. and Mehta, G.B., *Representations of Preference Orderings*, Springer, 1995.

[2] Debreu, G., Representation of a preference ordering by a numerical function, in *Decision Processes*, R. Thrall, C. Coombs and R. Davis, New York, Wiley (1954), 159-166.

[3] Debreu, G., Continuity properties of paretian utility, *International Economic Review* **5** (1964), 285-293.

[4] Eilenberg, S., 1941. Ordered topological spaces, American Journal of Mathematics 63, 39-45.

[5] Herden, G., On the existence of utility functions, *Mathematical Social Sciences* **17** (1989), 297-313.

[6] Mehta, G.B., Preference and utility, in: *Handbook of Utility Theory*, eds. S. Barberá, P.J. Hammond, C. Seidl, 1998, Kluwer Academic Publishers, pp. 1-47.

CHAPTER 2

Preferences over money lotteries and expected utility

2.1 Introduction

The present chapter aims to present the main concepts concerning preferences over lotteries, risk aversion, certainty equivalence and expected utility.
The second section, indeed, introduces the notions of a *money lottery*, *expected value of a money lottery*, *strictly increasing preference* and the different specifications of *risk aversion*.
The third section defines the *certainty equivalent* of a total preorder, and the corresponding *risk premium*. Results can be deduced, concerning the existence of continuous certainty equivalents in the case of strictly increasing preferences.
The fourth and last section presents a possible axiomatization, including a complete proof, of expected utility for total preorders over money lotteries.

2.2 Uncertainty, monetary lotteries and risk aversion

We introduce the basic concepts concerning money lotteries and attitudes to risk of an agent.

Definition 2.2.1 (monetary lottery)**.** A *monetary lottery* is a probability distribution over a (finite) list of outcomes, consisting of sums of money (expressed, for example, in €). Thus, it is an object of the form

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix},$$

where n is a positive integer, $0 \leq p_i \leq 1$ for all $i \in \{1, \dots, n\}$, and $p_1 + p_2 + \dots + p_n = 1$.

We shall assume, in this chapter, that actually, X is a closed bounded interval of the real line, $X = [0, M]$, for some real number $M > 0$. This is not a very restrictive assumption, since the choice of M is arbitrary.

Definition 2.2.2 (expected value of a monetary lottery)**.** The *expected value* (or *expectation*) of a money lottery

$$L = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

is defined to be

$$\mathbb{E}[L] = \sum_{i=1}^{n} x_i p_i.$$

In the sequel, we shall denote by $\mathcal{L}(X)$ a set of (finite) money lotteries on X, which in addition contains all the *degenerate* (i.e., concentrated) money lotteries

$$\begin{pmatrix} x \\ 1 \end{pmatrix} = \delta(x).$$

Therefore, we shall denote by $\delta(x)$ the *certain lottery* which amounts to $x \in [0, M]$ with probability one.
We shall consider an individual with a preference relation (preorder) $\precsim$ on $\mathcal{L}(X)$. Needless to say, there is a bijective correspondence between sums of money $x \in X$ and degenerate money lotteries $\delta(x)$.

Definition 2.2.3 (strictly increasing preference)**.** We say that a preorder $\precsim$ on $\mathcal{L}(X)$ (or, equivalently, the individual with preference relation $\precsim$ on $\mathcal{L}(X)$) is *strictly increasing* if, for every $x, y \in X$,

$$x < y \Rightarrow \delta(x) \prec \delta(y).$$

It is easy to verify that, if $\precsim$ is a strictly increasing preorder on $\mathcal{L}(X)$, then actually, for every $x, y \in X$,

$$x < y \Leftrightarrow \delta(x) \prec \delta(y).$$

Definition 2.2.4 (risk aversion and such)**.** Consider the preordered set $(\mathcal{L}(X), \precsim)$. We say that $\precsim$ (or, equivalently, the individual with preorder $\precsim$ on $\mathcal{L}(X)$) is

1. *(strictly) risk averse* if $L \precsim \delta(\mathbb{E}[L])$ for every $L \in \mathcal{L}(X)$ (respectively, $L \prec \delta(\mathbb{E}[L])$ for every $L \in \mathcal{L}(X)$ such that $L \neq \mathbb{E}[L]$);
2. *risk neutral* if $\delta(\mathbb{E}[L]) \sim L$ for every $L \in \mathcal{L}(X)$;
3. *(strictly) risk seeking* (or *risk loving*) if $\delta(\mathbb{E}[L]) \precsim L$ for every $L \in \mathcal{L}(X)$ (respectively, $\delta(\mathbb{E}[L]) \prec L$ for every $L \in \mathcal{L}(X)$ such that $L \neq \mathbb{E}[L]$).

Proposition 2.2.5 (risk neutrality implies a total preorder)**.** *If $\precsim$ is a strictly increasing and risk neutral preorder on $\mathcal{L}(X)$, then $\precsim$ is total and $\mathbb{E} : \mathcal{L}(X) \to \mathbb{R}_+$ is a utility functional on $(\mathcal{L}(X), \precsim)$.*

Proof. Consider any two money lotteries $L_1, L_2 \in \mathcal{L}(X)$. Since $\precsim$ is risk neutral, we have that $L_1 \sim \delta(\mathbb{E}[L_1])$ and $L_2 \sim \delta(\mathbb{E}[L_2])$. Assume, without loss of generality, that $\mathbb{E}[L_1] < \mathbb{E}[L_2]$. Since $\precsim$ is strictly increasing, this implies that $\delta(\mathbb{E}[L_1]) \prec \delta(\mathbb{E}[L_2])$. Then we get

$$L_1 \sim \delta(\mathbb{E}[L_1]) \prec \delta(\mathbb{E}[L_2]) \sim L_2,$$

and this implies that $L_1 \prec L_2$ by transitivity of $\precsim$. Hence, $\precsim$ is a total preorder. The proof that $\mathbb{E} : \mathcal{L}(X) \to \mathbb{R}_+$ is a utility functional on $(\mathcal{L}(X), \precsim)$ is now immediate. □

Example 2.2.6. Consider a strictly increasing and risk neutral individual with a preorder $\precsim$ on $\mathcal{L}(X)$. With respect to the following two lotteries

$$L_1 = \begin{pmatrix} 5 & 100 \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix}, \quad L_2 = \begin{pmatrix} 30 & 45 & 90 \\ \frac{1}{3} & \frac{5}{9} & \frac{1}{9} \end{pmatrix},$$

one gets $\mathbb{E}[L_1] = 43$ and $\mathbb{E}[L_2] = 45$, so that it must be $L_1 \sim \delta(43)$ and $L_2 \sim \delta(45)$ by risk neutrality, and $\delta(43) \prec \delta(45)$ by using the fact that the preorder is strictly increasing. Therefore $L_1 \sim \delta(43) \prec \delta(45) \sim L_2$ implies that $L_1 \prec L_2$ by transitivity.

2.3 Certainty equivalents and risk premium

Definition 2.3.1 (certainty equivalence functional)**.** Let $\precsim$ be a total preorder on $\mathcal{L}(X)$. Then a functional $\mathcal{C} : \mathcal{L}(X) \to \mathbb{R}_+$ is said to be the *certainty equivalence functional* on $(\mathcal{L}(X), \precsim)$ if $\mathcal{C}$ satisfies the following two conditions:

1. $\mathcal{C}$ is a utility functional on $(\mathcal{L}(X), \precsim)$ (i.e., $L_1 \precsim L_2$ is equivalent to $\mathcal{C}(L_1) \leq \mathcal{C}(L_2)$ for every $L_1, L_2 \in \mathcal{L}(X)$);
2. $L \sim \delta(\mathcal{C}(L))$ for every $L \in \mathcal{L}(X)$.

Therefore, for every $L \in \mathcal{L}(X)$, the value $\mathcal{C}(L)$ represents the numerical level such that the agent is indifferent between receiving it with certainty and receiving the money lottery L. .

Proposition 2.3.2. *Given a strictly increasing total preorder $\precsim$ on $\mathcal{L}(X)$, if there exists a certainty equivalence functional $\mathcal{C}$ on $(\mathcal{L}(X), \precsim)$, then it is unique.*

Proof. By contraposition, assume that there exist two certainty equivalence functionals $\mathcal{C}_1 \neq \mathcal{C}_2$ on $(\mathcal{L}(X), \precsim)$. Then there exists $L \in \mathcal{L}(X)$ such that $\mathcal{C}_1(L) \neq \mathcal{C}_2(L)$. Without loss of generality, assume that $\mathcal{C}_1(L) < \mathcal{C}_2(L)$. Then, from condition 2 in Definition 2.3.1, using the fact that $\precsim$ is strictly increasing, we have that $L \sim \delta(\mathcal{C}_1(L)) \prec \delta(\mathcal{C}_2(L)) \sim L$, and we arrive at the contradiction $L \prec L$ by using transitivity of $\precsim$. Hence, the proof is complete. □

In the following proposition, we characterize the utility functionals which are also certainty equivalence functionals.

Proposition 2.3.3. *Let $\precsim$ be a strictly increasing total preorder on $\mathcal{L}(X)$, and assume that there exists a utility functional $\mathcal{C}$ on $(\mathcal{L}(X), \precsim)$. Then the following conditions are equivalent:*

(i) *$\mathcal{C}$ is the certainty equivalence functional on $(\mathcal{L}(X), \precsim)$;*

(ii) *$\mathcal{C}(\delta(x)) = x$ for every $x \in X$.*

Proof. (i) $\Rightarrow$ (ii). By contraposition, assume that $\mathcal{C}(\delta(x)) \neq x$ for some $x \in X$, for example $\mathcal{C}(\delta(x)) < x$. Then we have that $\delta(\mathcal{C}(\delta(x))) \prec \delta(x)$ by using the fact that $\precsim$ is strictly increasing, and therefore $\mathcal{C}$ cannot be the certainty equivalence functional on $(\mathcal{L}(X), \precsim)$.
(ii) $\Rightarrow$ (i). $\mathcal{C}(\delta(x)) = x$ for every $x \in X$ implies that $\mathcal{C}(L) = \mathcal{C}(\delta(\mathcal{C}(L)))$, which in turn implies that $L \sim \delta(\mathcal{C}(L))$ as a consequence of the fact that $\mathcal{C}$ is a utility functional on $(\mathcal{L}(X), \precsim)$. This consideration completes the proof. □

In the following proposition we consider the case of a neutral strictly increasing agent.

Proposition 2.3.4. *If $\precsim$ is a strictly increasing and risk neutral total preorder on $\mathcal{L}(X)$, then $\mathbb{E} : \mathcal{L}(X) \to \mathbb{R}_+$ is the certainty equivalence functional on $(\mathcal{L}(X), \precsim)$.*

Proof. By Proposition 2.2.5, we have that $\mathbb{E} : \mathcal{L}(X) \to \mathbb{R}_+$ is a utility functional on $(\mathcal{L}(X), \precsim)$. The fact that $L \sim \delta(\mathbb{E}(L))$ for every $L \in \mathcal{L}(X)$ is immediately implied by the definition of risk neutrality. So the proof is complete. □

We incorporate some arguments concerning continuity in the following theorem. Indeed, we present a characterization of the existence of a continuous certainty equivalence functional.

Theorem 2.3.5. *Let $\precsim$ be a strictly increasing total preorder on $\mathcal{L}(X)$, and let τ_L be any topology on $\mathcal{L}(X)$. Then the following conditions are equivalent:*

(i) *There exists a (necessarily unique) nonnegative continuous certainty equivalence functional $\mathcal{C}$ on $(\mathcal{L}(X), \tau_L, \precsim)$;*

(ii) *The following conditions are verified:*

(a) *$\precsim$ is order-separable;*

(b) *$\precsim$ is continuous on the topological space $(\mathcal{L}(X), \tau_L)$;*

(c) *$\mathcal{C}(\delta(x)) = x$ for every $x \in X$.*

Proof. (i) $\Rightarrow$ (ii). We have that $\mathcal{C}(\delta(x)) = x$ for every $x \in X$ by Proposition 2.3.3. Therefore, $\mathcal{D} = \{\delta(q) : q \in [0, M] \cap \mathbb{Q}\}$ is a countable order-dense subset of $(\mathcal{L}(X), \precsim)$. Indeed, consider any pair $L_1, L_2 \in \mathcal{L}(X)$ such that $L_1 \prec L_2$. Then, we have that $\mathcal{C}(L_1) < \mathcal{C}(L_2)$. If we consider any rational number q such that $\mathcal{C}(L_1) < q < \mathcal{C}(L_2)$, it must be $L_1 \prec \delta(q) \prec L_2$, since $q = \mathcal{C}(\delta(q))$. Further, $\precsim$ is continuous on $(\mathcal{L}(X), \tau_L, \precsim)$ by Theorem 1.4.13.
(ii) $\Rightarrow$ (i). Conditions (a) and (b) imply the existence of a continuous utility functional $\mathcal{C}$ on $(\mathcal{L}(X), \tau_L, \precsim)$. Actually, Proposition 2.3.3 implies that $\mathcal{C}$ is a (necessarily unique) nonnegative certainty equivalence functional, since condition (c) holds. This consideration completes the proof. □

Definition 2.3.6 (risk premium)**.** Assume that there exists a certainty equivalence functional $\mathcal{C}$ on $(\mathcal{L}(X), \precsim)$. Then the *risk premium* Π is defined to be as the functional $\Pi : \mathcal{L}(X) \to \mathbb{R}_+$ such that, for all $L \in (\mathcal{L}(X)$,

$$L \sim \delta(\mathbb{E}[L] - \Pi(L)) \Leftrightarrow \mathcal{C}(L) = \mathbb{E}[L] - \Pi(L) \Leftrightarrow \Pi(L) = \mathbb{E}[L] - \mathcal{C}(L). \quad (2.3.1)$$

Remark 2.3.7. From Definition 2.3.1, we have that $\Pi \equiv 0$ if $\precsim$ is risk neutral, $\Pi \geq 0$ if $\precsim$ is risk averse, and $\Pi \leq 0$ if $\precsim$ is risk loving.

The risk premium $\Pi(L)$ can be interpreted as the price (relative to the expected value) that the individual is willing to pay in order not to face lottery L.

Definition 2.3.8 (von Neumann-Morgenstern utility representation)**.** A total preorder $\precsim$ on $\mathcal{L}(X)$ is said to have a *von Neumann-Morgenstern utility representation* if the functional

$$\mathbb{E}_u = U : \mathcal{L}(X) \to \mathbb{R}, \ \ \mathbb{E}_u[L] = U(L) = \sum_{i=1}^{n} u(x_i)p_i, \ \ L = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix} \tag{2.3.2}$$

is a utility functional for $\precsim$ for some function $u : X \to \mathbb{R}$.

We recall that a function $u : (\mathbb{R} \supset I) \longrightarrow \mathbb{R}$ defined on a real interval I is said to be *(strictly) concave* if, for all $x, y \in I$ and for every real number $0 < \alpha < 1$,

$$u(\alpha x + (1-\alpha)y) \geq \ \ (>) \ \ \alpha u(x) + (1-\alpha)u(y)$$

.

Convexity is defined in a perfectly symmetric way.

Theorem 2.3.9 (Jensen inequality)**.** *For any concave (convex) function* $u : (\mathbb{R} \supset) I \longrightarrow \mathbb{R}$*, the following inequality is verified:*

$$\mathbb{E}[u(X)] \leq u(\mathbb{E}[X]) \quad (\textit{respectively } \mathbb{E}[u(X)] \geq u(\mathbb{E}[X])). \tag{2.3.3}$$

.

Example 2.3.10 (two outcomes random variable)**.** *Given a random variable X with exactly two outcomes:*

$$\begin{aligned} &x_1 \quad \textit{with probability} \quad p, \\ &x_2 \quad \textit{with probability} \quad 1-p, \end{aligned}$$

then by applying property (2.3.3) to a concave utility function u we get

$$u(px_1 + (1-p)x_2) \geq p\ u(x_1) + (1-p)\ u(x_2),$$

which expresses the concavity of the function u.

Theorem 2.3.11. *Assume that a total preorder $\precsim$ on $\mathcal{L}(X)$ has a von Neumann-Morgenstern utility representation $U = \mathbb{E}_u$. Then the following statements hold:*

(i) *$\precsim$ is strictly increasing if and only if u is increasing on X;*

(ii) *$\precsim$ is strictly risk averse if and only if u is strictly concave on X.*

Proof. (i). Consider that $\precsim$ on $\mathcal{L}(X)$ is strictly increasing if and only if

$$x < y \Rightarrow \delta(x) \prec \delta(y) \Leftrightarrow \mathbb{E}_u[\delta(x)] = u(x) < u(y) = \mathbb{E}_u[\delta(y)].$$

(ii). Let us first show that, if $\precsim$ is strictly risk averse, then u is strictly concave on X. We have that for all distinct $x, y \in X$, and $\alpha \in [0,1]$,

$$\alpha\delta(x)+(1-\alpha)\delta(y) \prec \delta(\alpha x+(1-\alpha)y) \Leftrightarrow \alpha u(x)+(1-\alpha)u(y) < u(\alpha x+(1-\alpha)y),$$

so that u is strictly concave. Conversely, if u is strictly concave, Jensen's inequality 2.3.3 implies strict risk aversion, since, for every $L \in \mathcal{L}(X)$,

$$\mathbb{E}_u[L] \leq u(\mathbb{E}[L]) = \mathbb{E}_u[\delta(\mathbb{E}[L])] \Leftrightarrow L \precsim \delta(\mathbb{E}[L]),$$

with equality if and only if $L = \delta(\mathbb{E}[L])$. □

Definition 2.3.12 (expected utility representation)**.** We say that a function $u : X = [0, M] \to \mathbb{R}$ is utility function in a von Neumann-Morgenstern utility representation of a total preorder $\precsim$ on $\mathcal{L}(X)$ if $\mathbb{E}_u$ is a utility functional for $\precsim$ with u strictly increasing and concave on X. In this case we say that $\mathbb{E}_u$ is an *expected utility representation* of $\precsim$.

Remark 2.3.13 (positive linear transformations of utilities)**.** It is known that, if a total preorder $\precsim$ on $\mathcal{L}(X)$ has an expected utility representation $\mathbb{E}_u$, then the utility function u is defined *up to positive linear transformations*. This immediate fact means that, if u is a utility function, then also u' is a utility function, provided that $u' = au + b$, with $a, b \in \mathbb{R}$ and $a > 0$. Indeed, we have that, for every L_1, $L_2 \in \mathcal{L}(X)$,

$$L_1 \precsim L_2 \Leftrightarrow \mathbb{E}_u[L_1] \leq \mathbb{E}_u[L_2] \Leftrightarrow \mathbb{E}_{u'}[L_1] \leq \mathbb{E}_{u'}[L_2].$$

For this reason, if u is differentiable at 0, it is not restrictive to assume that $u(0) = 0$ and $u'(0) = 1$. If, further, u is twice differentiable at 0, by using the second degree Taylor's polynomial, we have that

$$u(x) \approx x + \frac{u''(0)}{2}x^2.$$

The utility functions are also called Bernoulli utility functions because Daniel Bernoulli (1700-1782) introduced them to answer the famous *St. Petersburg Paradox*.

Example 2.3.14 (St. Petersburg paradox). *Consider the following game. A fair coin is tossed until a head appears. The player receives an amount 2^n if head appears for the first time at the n-th toss. Therefore, the expected gain from the game (i.e., the price of the game) is*

$$\mathbb{E}[X] = \sum_{n=1}^{+\infty} 2^n \frac{1}{2^n} = +\infty.$$

This is in contrast with the intuition that many people would like to enter the game. On the other hand, under utility theory, the price (expected utility) $\mathbb{E}[u(X)]$ is finite as soon as, for example, $u(x) = log\, x$ (in this case we have that $\mathbb{E}[u(X)] = log\, 2\ \sum_{n=1}^{\infty} \frac{n}{2^n} < \infty$). With a utility function $u(x) = \sqrt{x}$, we get

$$\mathbb{E}[u(X)] = \sum_{n=1}^{+\infty} \frac{\sqrt{2^n}}{2^n} = \sum_{n=1}^{+\infty} \frac{1}{\sqrt{2^n}}$$

which is a little over 2.4.

The first major challenge to this principle of expected return maximization appeared in 1738 at the hand of Daniel Bernoulli, a member of the Swiss family of distinguished mathematicians. Bernoulli proposed two theses.

His first thesis does not involve risk or probability. It says that a person's subjective value $u(w)$ of wealth w does not increase linearly in w but rather increases at a decreasing rate, a proposition known later in economics as the principle of diminishing marginal utility of wealth. Bernoulli argued further that the rate of increase in $u(w)$ is inversely proportional to w and, hence, that u is a logarithmic function of wealth.

Bernoulli's second thesis, set forth in opposition to maximization of expected return or expected wealth, says that a risky prospect p' on levels of wealth ought to be evaluated by its expected subjective value

$$\sum u(w)p(w).$$

Alternatively, if w_0 is present wealth, then the expected subjective value of p is

$$\mathbb{E}(u, p) = \sum_{x \in X} u(w_0 + x)p(x),$$

with p less desirable than q when $\mathbb{E}(u, p)) < \mathbb{E}(u, q)$.

Remark 2.3.15 (Examples of utility functions)**.** The following classes of utility function can be considered:

1. *linear utility*: $u(x) = x$;
2. *quadratic utility*: $u(x) = -(\alpha - x)^2$ $(x \leq \alpha)$;
3. *logarithmic utility*: $u(x) = log(\alpha + x)$ $(x > -\alpha)$;
4. *exponential utility*: $u(x) = -\alpha e^{-\alpha x}$ $(\alpha > 0)$;
5. *power utility*: $u(w) = x^c$ $(x > 0, 0 < c \leq 1)$.

Needless to say, it is natural to consider the *certainty equivalent under expected utility* (when available).

Definition 2.3.16 (certainty equivalent under expected utility)**.** The *certainty equivalent under expected utility* $\mathcal{C}_u$, corresponding to an expected utility representation $\mathbb{E}_u$ of a total preorder $\precsim$ on $\mathcal{L}(X)$, is defined to be, for every $L \in \mathcal{L}(X)$,

$$\mathcal{C}_u(L) = u^{-1}\left(\mathbb{E}_u[L]\right). \tag{2.3.4}$$

Remark 2.3.17. Notice that, in order to include the possibility of losses, the consequence space can contain negative reals, so that, for example, we can set $X = [-M, M]$, for some real number $M > 0$.

CHAPTER 3

Axiomatization of expected utility under risk

3.1 Introduction

The present chapter aims to present a possible axiomatization, including a complete proof, of expected utility for total preorders. In particular, the second section is devoted to the case *simple probability distributions*, the third section contains some arguments about the general case of probability distributions, including some references to continuity issues, and the fouth section presents a short review of some classical arguments concerning risk attitudes reflected by the utility function.

3.2 From linear utility to expected utility for simple probabilities

A theory of choice among risky decisions is said to be a *Bernoullian expected utility theory under risk* when it consists of the following elements:

1. A set X of *outcomes* (*consequences*) and a set $\mathcal{P}$ of probability measures (distributions) on X;
2. A *utility function* u on X, usually presumed unique up to positive linear transformations;
3. The *principle of choice*, which says that the most desirable distributions, or their corresponding risky alternatives, are those that maximize

expected utility

$$\mathbb{E}_u^p = \sum u(x)p(x).$$

In this section we are going to present the classical axiomatization of Bernoullian expected utility theory under risk based on the famous *von Neumann and Morgenstern theorem.*

Definition 3.2.1 (Convex set)**.** A set $\mathcal{P}$ is said to be *convex* if, for every real number $\lambda \in [0,1]$, and for all pairs $(p,q) \in \mathcal{P} \times \mathcal{P}$,

$$\lambda p + (1-\lambda)q \in \mathcal{P}.$$

Definition 3.2.2 (Linear functional)**.** A functional U on a convex set $\mathcal{P}$ is said to be *linear* if, for every real number $\lambda \in [0,1]$, and for all pairs $(p,q) \in \mathcal{P} \times \mathcal{P}$,

$$U(\lambda p + (1-\lambda)q) = \lambda U(p) + (1-\lambda)U(q).$$

Definition 3.2.3 (Independence and continuity)**.** Let $\precsim$ be a preorder on a convex set $\mathcal{P}$. Then $\precsim$ is said to satisfy the

1. *Independence axiom*, if, for all $p,q,r \in \mathcal{P}$, and for all $0 < \lambda < 1$,

$$p \precsim q \Leftrightarrow \lambda p + (1-\lambda)r \precsim \lambda q + (1-\lambda)r;$$

2. *Continuity axiom*, if, for all $p,q,r \in \mathcal{P}$,

$$p \prec r \prec q \Rightarrow \text{ there exists } \lambda \in]0,1[\text{ such that } r \sim \lambda p + (1-\lambda)q.$$

In the following lemma, we present some consequences of the fundamental axioms contained in Definition 3.2.3.

Lemma 3.2.4. *Let $\precsim$ be a total preorder on a convex set $\mathcal{P}$, and assume that $\precsim$ satisfies both the Independence axiom and the Continuity axiom. The following statements are true for all $p,q,r \in \mathcal{P}$, and for all $\lambda, \mu \in [0,1]$:*

1. $(p \prec q)$ *and* $(0 < \lambda < 1) \Rightarrow \lambda p + (1-\lambda)r \prec \lambda q + (1-\lambda)r$*;*

2. $(p \prec q)$ *and* $(0 < \lambda < 1) \Rightarrow p \prec \lambda p + (1-\lambda)q \prec q$;

3. $(p \prec q)$ *and* $(\lambda > \mu) \Rightarrow \lambda p + (1-\lambda)q \prec \mu p + (1-\mu)q$;

4. $p \prec r \prec q \Rightarrow$ *there exists a unique* $\lambda \in]0,1[$ *with* $r \sim \lambda p + (1-\lambda)q$;

5. $p \sim q \Rightarrow \lambda p + (1-\lambda)r \sim \lambda q + (1-\lambda)r$.

Proof. In order to prove statement 1, consider that, by the Independence axiom, since $\precsim$ is a total preorder,

$$\lambda q + (1-\lambda)r \precsim \lambda p + (1-\lambda)r \Rightarrow q \precsim p \Rightarrow not(p \prec q).$$

In order to prove statement 2, consider that, by statement 1, for all $p, q \in \mathcal{P}$ such that $p \prec q$, and for all $0 < \lambda < 1$,

$$p = (1-\lambda)p + \lambda p \prec (1-\lambda)q + \lambda p = \lambda p + (1-\lambda)q \prec \lambda q + (1-\lambda)q = q.$$

Observe that statement 3 implies statement 4. Indeed, for $\lambda > \mu$, $r \sim \lambda p + (1-\lambda)q \prec \mu p + (1-\mu)q$ implies that $r \prec \mu p + (1-\mu)q$, so that it cannot happen $r \sim \lambda p + (1-\lambda)q$ for more than one number $\lambda \in]0,1[$. Therefore, we are done if we prove statement 3. To this aim, consider $p, q \in \mathcal{P}$ such that $p \prec q$, and two real numbers $\lambda > \mu$, with λ, $\mu \in]0,1[$. If $\lambda = 1$, or else $\mu = 0$. then condition (a) reduces to $p \prec \mu p + (1-\mu)q$, and respectively to $\lambda p + (1-\lambda)q \prec q$, and both statements are true by the above statement 2. Hence, let us assume that $1 > \lambda > \mu > 0$. We have that

$$\lambda p + (1-\lambda)q = \frac{\lambda-\mu}{1-\mu}p + \frac{1-\lambda}{1-\mu}(\mu p + (1-\mu)q) \prec$$
$$\frac{\lambda-\mu}{1-\mu}(\mu p + (1-\mu)q) + \frac{1-\lambda}{1-\mu}(\mu p + (1-\mu)q) = (\mu p + (1-\mu)q),$$

due to the Independence axiom and statement 1, according to which $p \prec \mu p + (1-\mu)q$ (see again the above statement 2).
Statement 5 is an immediate consequence of the Independence axiom. □

Recall that a *Boolean algebra* $\mathcal{B}$ of subsets of a set X is a family of subsets of X which is closed under finite unions and the complement, and which contains X and $\emptyset$.

Theorem 3.2.5 (von Neumann and Morgenstern, 1944). *Let $\mathcal{P}$ be a convex set of probability measures defined on a Boolean algebra $\mathcal{B}$ of subsets of a set X of consequences, and let $\precsim$ be a reflexive binary relation on $\mathcal{P}$. Then the following conditions are equivalent:*

1. *There exists a linear utility functional U for $\precsim$ on $\mathcal{P}$;*

2. *The binary relation $\precsim$ on $\mathcal{P}$ satisfies the following conditions:*

   (a) *$\precsim$ is a total preorder on $\mathcal{P}$;*

   (b) *$\precsim$ satisfies the Independence axiom:*

   (c) *$\precsim$ satisfies the Continuity axiom.*

Moreover, U is unique up to a positive linear transformation.

Proof. $1 \Rightarrow 2$. This part of the proof is very simple and it is left to the reader.
$2 \Rightarrow 1$. Part I. If $p \sim q$ for every pair $(p, q) \in \mathcal{P} \times \mathcal{P}$, then we have that any constant $U \equiv c$ $(c \in \mathbb{R})$ is a linear utility functional for $\precsim$. Therefore, consider any pair $(p, q) \in \mathcal{P} \times \mathcal{P}$ such that $p \prec q$, and denote by $[p, q]$ the "closed interval" defined by

$$[p, q] = \{r \in \mathcal{P} : p \precsim r \precsim q\}.$$

Then, from Lemma 3.2.4, 4, actually there exists a unique $\lambda \in [0, 1]$ such that

$$r \sim \lambda p + (1 - \lambda) q.$$

Therefore, for every $r \in [p, q]$ there exists a unique real number $f(r) \in [0, 1]$ such that

$$r \sim f(r)p + (1 - f(r))q, \quad f(p) = 1, \quad f(q) = 0.$$

Let us show that $U = 1 - f$ is a utility functional for $\precsim$ on $[p, q]$. Consider any pair $(r, s) \in [p, q] \times [p, q]$. If $U(r) < U(s) \Leftrightarrow f(r) > f(s)$, then from the definition of f, and Lemma 3.2.4, 3, we have that

$$(1 - U(r))p + U(r)q \prec (1 - U(s))p + U(s)q.$$

Indeed, we know that

$$(p \prec q) \textit{ and } (\lambda > \mu) \Rightarrow \lambda p + (1 - \lambda) q \prec \mu p + (1 - \mu) q.$$

Hence, from the definition of f, we get

$$r \sim f(r)p + (1 - f(r))q \prec f(s)p + (1 - f(s))q \sim s,$$

and transitivity of $\prec$ guarantees that $r \prec s$. On the other hand, if $U(r) = U(s) \Leftrightarrow f(r) = f(s)$, then

$$r \sim f(r)p + (1 - f(r))q = f(s)p + (1 - f(s))q \sim s,$$

so that $r \sim s$. So, U is a utility functional for $\precsim$ on $[p,q]$.
It remains to show that U is linear on $[p,q]$. To this aim, consider any pair $(r,s) \in [p,q] \times [p,q]$, and any real number $0 \leq \lambda \leq 1$. Then, by convexity, we have that $\lambda r + (1-\lambda)s \in [p,q]$, and, from the definition of f, we have that

$$\lambda r + (1-\lambda)s \sim f(\lambda r + (1-\lambda)s)p + [1 - f(\lambda r + (1-\lambda)s)]q.$$

Since, by statement 5 in Lemma 3.2.4,

$$p \sim q \Rightarrow \lambda p + (1-\lambda)r \sim \lambda q + (1-\lambda)r,$$

we arrive at

$$\lambda r + (1-\lambda)s \sim \lambda[f(r)p + (1-f(r))q] + (1-\lambda)[f(s)p + (1-f(s))q],$$

that is

$$\lambda r + (1-\lambda)s \sim [\lambda f(r) + (1-\lambda)f(s)]p + \{1 - [\lambda f(r) + (1-\lambda)f(s)]\}q.$$

Therefore,

$$\begin{aligned} f(\lambda r + (1-\lambda)s)p \;&+\; [1 - f(\lambda r + (1-\lambda)s)]q \sim [\lambda f(r) + (1-\lambda)f(s)]p \\ &+\; \{1 - [\lambda f(r) + (1-\lambda)f(s)]\}q, \end{aligned}$$

and this implies that

$$f(\lambda r + (1-\lambda)s) = \lambda f(r) + (1-\lambda)f(s),$$

i.e. f is linear, and equivalently U is linear. The fact that U is defined up to positive linear transformations is immediate.
Part II. It remains to show that such linear utility functional U for the total preorder $\precsim$ on the fixed preference interval $[p,q]$ actually serves as a linear utility functional for $\precsim$ on all of $\mathcal{P}$. To this aim, consider a closed preference interval $[p,q]$, and two other preference intervals $[p_1,q_1]$ and $[p_2,q_2]$ containing $[p,q]$. Consider two linear utility functionals U_1 and U_2 for $\precsim$ on $[p_1,q_1]$ and $[p_2,q_2]$, respectively, whose existence is guaranteed by the above Part I of the proof. Assume that U_1 and U_2 are scaled by positive linear transformations, in such a way that $U_1(p) = U_2(p) = 0$ and $U_1(q) = U_2(q) = 1$. Let us show that, for every $r \in \mathcal{P}$,

$$r \in [p_1,q_1] \cap [p_2,q_2] \Rightarrow U_1(r) = U_2(r).$$

Consider any $r \in [p_1,q_1] \cap [p_2,q_2]$. Then, one of the following three cases obtains:

1. $p \prec q \prec r$: In this case, by Lemma 3.2.4, 4, there exists a unique real number $\lambda \in]0,1[$ such that $q \sim \lambda p + (1-\lambda)r$;

2. $p \precsim r \precsim q$: In this case, by Lemma 3.2.4, 4, there exists a unique real number $\mu \in [0,1]$ such that $r \sim \mu p + (1-\mu)q$ (actually, $\mu = 1$ if it happens that $r \sim p$, and $\mu = 0$ if it happens that $r \sim q$);

3. $r \prec p \prec q$: In this case, by Lemma 3.2.4, 4, there exists a unique real number $\nu \in]0,1[$ such that $p \sim \nu r + (1-\nu)q$.

Therefore, by using scaling and linearity of each U_i ($i \in \{1,2\}$), we get

$$\begin{cases} 1 = (1-\lambda)U_i(r) \quad (i \in \{1,2\}) \\ U_i(r) = 1-\mu \quad (i \in \{1,2\}) \\ 0 = \nu U_i(r) + 1 - \nu \quad (i \in \{1,2\}) \end{cases},$$

so that $U_1(r) = U_2(r)$ in each case.
Since every pair of measures in $\mathcal{P}$ is in at least one preference interval $[p_i, q_i]$ that includes $[p,q]$, it follows that actually U is a linear utility functional for $\precsim$ on $\mathcal{P}$.
The proof that U is unique up to a positive linear transformation (i.e., if U, U' are two linear functionals representing the same total preorder $\precsim$ on $\mathcal{P}$, then there are scalars $a, b \in \mathbb{R}$ with $a > 0$, such that $U' = aU + b$), is omitted for the sake of brevity. □

Definition 3.2.6 (Simple probability measure)**.** A probability measure p on a Boolean algebra $\mathcal{B}$ on an arbitrary nonempty set X is said to be *simple* if there exists a finite subset $\{x_1, ..., x_n\}$ of X such that $p(X \setminus \{x_1, ..., x_n\}) = 0$. A simple probability measure p on X will be denoted as a *finite lottery*

$$\{x_i, p_i\}_{i=1}^n = \{x_1, p_1; x_2, p_2; ...; x_n, p_n\} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix},$$

where $p(x_i) = p_i$ for $i = 1, ..., n$, $p_i > 0$ for all $i \in \{1, ..., n\}$, and $p(x) = 0$ for every $x \notin \{x_1, ..., x_n\}$. In particular, the symbol $\delta(x) = p_x$ will stand for the *degenerate probability* (*degenerate lottery*) valued 1 at $x \in X$.

Theorem 3.2.7 (Expected utility for simple probability measures)**.** *Let $\mathcal{P}$ be a convex set of probability measures defined on a Boolean algebra $\mathcal{B}$ of subsets of a set X of consequences, assume that $\mathcal{P}$ contains all the degenerate probabilities p_x ($x \in X$), and let $\precsim$ be a total preorder on $\mathcal{P}$. Assume that there exists a linear utility functional U for $\precsim$ on $\mathcal{P}$. If we define a real-valued function u on X by*

$$u(x) = U(p_x) \quad (x \in X), \tag{3.2.1}$$

then we have that, for all simple probability measures $p = \{x_i, p_i\}_{i=1}^n$ *and* $q = \{y_j, q_j\}_{j=1}^m$,

$$p = \{x_i, p_i\}_{i=1}^n \precsim \{y_j, q_j\}_{j=1}^m = q \Leftrightarrow \mathbb{E}_u^p = \sum_{i=1}^n u(x_i)p_i \leq \sum_{j=1}^m u(y_j)q_j = \mathbb{E}_u^q.$$

Proof. We only have to show that, for every simple probability measure $p \in \mathcal{P}$,

$$U(p) = \sum_{x \in X} u(x)p(x).$$

□

Let n be the number of points of X at which a simple probability measure $p \in \mathcal{P}$ is positive. Then the above fact is clear by the above definition 3.2.1 if $n = 1$, by linearity of U if $n = 2$, and then by induction on n when $n \geq 3$, □

From Theorem 3.2.5 and Theorem 3.2.7, we get the following corollary, presenting a sufficient condition for expected utility for simple probability measures.

Corollary 3.2.8 (Condition for expected utility for simple probability measures)**.** *Let* $\mathcal{P}$ *be a convex set of probability measures defined on a Boolean algebra* $\mathcal{B}$ *of subsets of a set* X *of consequences, assume that* $\mathcal{P}$ *contains all the degenerate probabilities* p_x $(x \in X)$*, and let* $\precsim$ *be a total preorder on* $\mathcal{P}$ *satisfying both the Independence axiom and the Continuity axiom. Then there exists a real-valued function* u *on* X *such that, for all simple probability measures* $p = \{x_i, p_i\}_{i=1}^n$ *and* $q = \{y_j, q_j\}_{j=1}^m$,

$$p = \{x_i, p_i\}_{i=1}^n \precsim \{y_j, q_j\}_{j=1}^m = q \Leftrightarrow \mathbb{E}_u^p = \sum_{i=1}^n u(x_i)p_i \leq \sum_{j=1}^m u(y_j)q_j = \mathbb{E}_u^q.$$

An alternative axiomatization of linear utility is due to Herstein and Milnor (1953).

Theorem 3.2.9 (Herstein and Milnor characterization)**.** *Let $\mathcal{P}$ be a convex set of probability measures defined on a Boolean algebra $\mathcal{B}$ of subsets of a set X of consequences, and let $\precsim$ be a reflexive binary relation on $\mathcal{P}$. Then the following conditions are equivalent:*

1. *There exists a linear utility functional U for $\precsim$ on $\mathcal{P}$;*
2. *The binary relation $\precsim$ on $\mathcal{P}$ satisfies the following conditions:*
   (a) *$\precsim$ is a total preorder on $\mathcal{P}$;*
   (b) *For all $p, q, r \in \mathcal{P}$, $p \sim q \Rightarrow \frac{1}{2}p + \frac{1}{2}r \sim \frac{1}{2}q + \frac{1}{2}r$;*
   (c) *For all $p, q, r \in \mathcal{P}$, $\{\alpha \in [0,1] : \alpha p + (1-\alpha)r \precsim q\}$ and $\{\beta \in [0,1] : q \precsim \beta p + (1-\beta)r\}$ are closed subsets of $[0,1]$.*

Let us present an example of a total preorder for which the Independence axiom is not satisfied.

Example 3.2.10. A binary relation $\precsim_m^M$ over simple probability measures on $X \subset \mathbb{R}$ is said to be a *maximin preference relation* if, for any two simple probability measures p, q, the following condition is verified:

$$p \precsim_m^M q \Leftrightarrow \forall x' \in X \text{ such that } q(x') > 0 \ \exists x \in X \text{ such that } (p(x) > 0) \textit{ and } (x \leq x').$$

It is easy to show that $\precsim_m^M$ is a total preorder which violates the Independence axiom.

Indeed, given any two outcomes $x_1 < x_2$, we have that

$$p = \delta(x_1) \prec_m^M \delta(x_2) = q,$$

while

$$\frac{1}{2}p + \frac{1}{2}p \sim_m^M \frac{1}{2}q + \frac{1}{2}p.$$

3.3 Expected utility for general probabilities

We present a general theorem on the existence of a bounded continuous utility function in a von Neumann and Morgenstern representation of total preorder on general probability measures. Some topological definitions are needed.

Definition 3.3.1 (coarser topology)**.** For two topologies τ, τ' on a set X, we say that τ' is coarser that τ if $\tau' \subset \tau$.

Let $X = [0, M]$ endowed with the usual interval topology, and denote by $\mathcal{M}$ the set of all probability measures on $(X, \mathcal{B})$, where $\mathcal{B}$ denotes the Borel σ-algebra on X ((i.e., $\mathcal{B}$ is the smallest σ-algebra of subsets of X containing the open subsets of X). Assume that $\mathcal{M}$ is endowed with the topology of *weak convergence* $\tau^w_{\mathcal{M}}$.

Definition 3.3.2 (topology of weak convergence)**.** The *topology of weak convergence* $\tau^w_{\mathcal{M}}$ on the set $\mathcal{M}$ of all probability measures on $(X, \mathcal{B})$ is the coarsest topology such that, for every continuous and bounded real-valued function f on X, the map

$$p \to \int_X f(x)dp(x)$$

is continuous. Therefore, a sequence $\{p_n\} \subset \mathcal{M}$ converges to $p \in \mathcal{M}$ according to the topology of weak convergence if, for every continuous and bounded real-valued function f on X,

$$\lim_{n \to +\infty} \int_X f(x)dp_n(x) = \int_X f(x)dp(x).$$

Then, the following general theorem holds true.

Theorem 3.3.3. *The following conditions are equivalent on a binary relation* $\precsim$ *on* $\mathcal{M}$*:*

1. *There exists a real-valued, bounded and continuous function* u *on* X *such that the following functional* U *on* $\mathcal{M}$ *is a utility functional for* $\precsim$*:*

$$U(p) = \int_X u(x)dp(x) \quad (p \in \mathcal{M}); \tag{3.3.1}$$

2. *The binary relation* $\precsim$ *on* $\mathcal{M}$ *satisfies the following conditions:*

   (a) $\precsim$ *is a total preorder on* $\mathcal{M}$*;*

   (b) $\precsim$ *satisfies the Independence axiom:*

   (c) $\precsim$ *is a continuous preorder on the topological space* $(\mathcal{M}, \tau^w_{\mathcal{M}})$.

   Moreover, u *is unique up to a positive linear transformation.*

3.4 Risk attitudes and the utility function

The theory of risk attitudes developed by Pratt (1964) and Arrow (1974) is concerned with curvature properties of the real-valued function u on X as defined in the expected utility representation 3.3.1, when X is an interval of monetary amounts interpreted either as wealth levels or gains and losses around a given present wealth. Its purpose is to interpret various types of economic behavior in risky situations in terms of curvature and perhaps other properties of u on X within the von Neumann-Morgenstern framework of maximizing expected utility.

Definition 3.4.1 (risk aversion)**.** Assume that u on X is twice differentiable and strictly increasing in x, so $u'(x) > 0$ for every x. Following Pratt and Arrow, we say that u, defined on an interval of X, is

1. *risk averse*, if $u''(x) < 0$ for every x;
2. *risk seeking* (or *risk loving*), if $u''(x) > 0$ for every x;
3. *risk neutral*, if $u'(x) = 0$ for every x.

Definition 3.4.2 (risk aversion measurement)**.** Risk-averse utility functions u, which increase in x at a decreasing rate, are further characterized by their

1. *index of absolute risk aversion*

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

2. *index of relative risk aversion*

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

These indices, which can also be used when u'' is not negative, are invariant to positive linear transformations of u.
Stochastic dominance also involves the shape of u on X. It is concerned with comparative aspects of measures p and q.
Let p^1 and p^2 denote the first two first cumulative distributions of the simple measure p on X, namely:

$$p^1(x) = \sum_{z \le x} p(z) \quad (z \in X),$$

$$p^2(x) = \int_{-\infty}^{x} p^1(z)dz \quad (z \in X).$$

Then, denote by $\precsim^1$ and $\precsim^2$ the *first order stochastic dominance relation*, and respectively the *second order stochastic dominance relation*:

$$p \precsim^1 q \Leftrightarrow p^1(x) \ge q^1(x) \quad \text{for all } x \in X;$$

$$p \precsim^2 q \Leftrightarrow p^2(x) \ge q^2(x) \quad \text{for all } x \in X.$$

Clearly, $\precsim^1$ and $\precsim^2$ are both nontotal preorders, and $\prec^1 \subset \prec^2$.
Let $\mathcal{U}^1$ and $\mathcal{U}^2$ be the class of all strictly increasing real-valued functions u on X, and respectively the class of all strictly increasing and strictly concave real-valued functions u on X. Then, with $\mathbb{E}_u^p = \sum u(x)p(x)$, it is not hard to show that

$$p \prec^1 q \Leftrightarrow \mathbb{E}_u^p < \mathbb{E}_u^q \quad \text{for all } u \in \mathcal{U}^1;$$

$$p \prec^2 q \Leftrightarrow \mathbb{E}_u^p < \mathbb{E}_u^q \quad \text{for all } u \in \mathcal{U}^2.$$

e

CHAPTER 4

A sketch of probability and random variables

4.1 Probability and random variables

Denote by Ω an abstract space of *elementary events* and by $\mathcal{F}$ a Boolean σ-algebra of subsets of Ω. A probability *Prob* on the *measurable space* $(\Omega, \mathcal{F})$ is a function $Prob : \mathcal{F} \longrightarrow [0,1]$ satisfying the following conditions:

1. $Prob(\emptyset) = 0$, $Prob(\Omega) = 1$;

2. $Prob(\cup F_n) = \sum_{n=1}^{\infty} Prob(F_n)$ for every countable sequence $\{F_n\}_{n\in\mathbb{N}}$ of pairwise disjoint elements of $\mathcal{F}$ (that is, $F_m \cap F_n = \emptyset$ for all $m \neq n$, $m, n \in \mathbb{N}$).

Denote by $\mathcal{B}$ the *Borel σ-algebra* on the real line $\mathbb{R}$ (i.e., $\mathcal{B}$ is the smallest σ-algebra of subsets of the real line containing the open sets). A real random variable X on the measurable space $(\Omega, \mathcal{F})$ is a *measurable real valued function on* $(\Omega, \mathcal{F})$ (in the sense that $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$ for every $B \in \mathcal{B}$).

A random variable X on some *probability space* $(\Omega, \mathcal{F}, Prob)$ is "known" when we know its *cumulative distribution function*

$$F_X(x) = Prob(X \leq x) \quad x \in \mathbb{R}. \tag{4.1.1}$$

In the sequel we shall deal with nonnegative random variables.
For every nonnegative random variable X we have that

1. $\lim_{x \to +\infty} F_X(x) = 1$;

2. F_X is non-decreasing;
3. F_X is right-continuous.

Essentially two cases may occur:

1. F_X is a step function. Then F_X has at most countably many points of discontinuity. In this case the random variable X is said to be *discrete* and if $\bar{x}$ is a point of discontinuity of F_X then
$$F_X(\bar{x}) - F_X(\bar{x} - 0) = F_X(\bar{x}) - \lim_{x \to \bar{x}^-} F_X(x) = Prob(X = \bar{x}).$$
Therefore $Prob(X = \bar{x})$ is precisely the *jump* of F_X at $\bar{x}$;
2. F_X is *absolutely continuous*. Then F_X is differentiable and $f(x) = F_X'(x)$ is its *probability density function*. Then
$$F_X(x) = \int_0^x f(t)dt \quad (f(x) \geq 0 \text{ and } \int_0^\infty f(t)dt = 1).$$

In the sequel, we shall simply $F(x)$ insted of $F_X(x)$ when it is clear that we refer to some particular random variable X. write F_X

4.2 Expectations, conditional expectations and joint distributions

The *expectation* $\mathbb{E}[X]$ of a random variable X is

1. $\mathbb{E}[X] = \sum_{n=0}^\infty x_n Prob(X = x_n)$ if X is discrete with outcomes x_n $(n \in \mathbb{N})$;
2. $\mathbb{E}[X] = \int_0^\infty x dF(x) = \int_0^\infty x f(x) dx$ if X is absolutely continuous with density $f(x)$.

If X is a random variable, then its *decumulative distribution function* $S(x)$ is defined to be
$$S(x) = Prob(X > x) = 1 - F(x).$$

The following expression of $\mathbb{E}[X]$ may thought of as useful:
$$\mathbb{E}[X] = \int_0^\infty S(t)dt = \int_0^\infty Prob(X > t)dt. \tag{4.2.1}$$

In order to illustrate the validity of the previous formula, let us observe that
$$\int_0^\infty x dF(x) = \int_0^\infty dF(x) \int_0^x dt = \int_0^\infty dt \int_t^\infty dF(x) = \int_0^\infty S(t)dt.$$

The expectation is *linear*, in the sense that, for all $a, b \in \mathbb{R}$,

$$\mathbb{E}[a + bX] = a + b\mathbb{E}[X].$$

The *variance* of a random variable X is

$$Var[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - E^2[X].$$

We have that, for all $a, b \in \mathbb{R}$,

$$Var[a + bX] = b^2 Var[X].$$

Now consider two random variables X and N, with X absolutely continuous and N discrete with outcomes $n = 0, 1, 2, ...$ Then the *conditional distribution of X given $N = n$* is defined to be

$$Prob(X \leq x \mid N = n) = \frac{Prob((X \leq x) and (N = n))}{Prob(N = n)}, \tag{4.2.2}$$

when $Prob(N = n) > 0$.
We say that X and N are *independent* if, for all $x \geq 0$ and $n \in \mathbb{N}$,

$$Prob(X \leq x \mid N = n) = Prob(X \leq x). \tag{4.2.3}$$

The *expectation of X conditional to $n \in N$* (for some fixed $n \in N$) is defined to be

$$\mathbb{E}[X \mid N = n] = \int_0^\infty x d\, Prob(X \leq x \mid N = n), \tag{4.2.4}$$

while $\mathbb{E}[X \mid N]$ is the discrete random variable taking value $\mathbb{E}[X \mid N = n]$ with probability $p_n = Prob(N = n)$.
In general, given any two random variables X and Y, the *joint cumulative distribution function of X and Y* is defined to be

$$F(x, y) = Prob((X \leq x) \;\; and \;\; (Y \leq y)).$$

A nonnegative function $f(x, y)$ is said to be the *joint density* of the random variables X and Y if

$$F(x, y) = \int_0^x \int_0^y f(u, v) du dv.$$

If X and Y are both discrete, then it is clear that the pair (X, Y) is known when the probabilities $p_{mn} = Prob((X = m) \;\; and \;\; (Y = n))$ are assigned for all the possible outcomes m and n of X and Y, respectively.
The random variables X and Y are said to be *independent* if $F(x, y) = F_X(x) F_Y(y)$. Such a condition is reflected by $p_{mn} = p_m p_n$ and $f(x, y) = f_X(x) f_Y(y)$, respectively.

Given two random variables X and Y with joint density $f(x,y)$, the *covariance of X and Y* is defined to be

$$\begin{aligned} cov(X,Y) &= \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \\ &= \int_0^\infty \int_0^\infty (x-\mathbb{E}[X])(y-\mathbb{E}[Y])f(x,y)dxdy = \\ &= \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]. \end{aligned} \tag{4.2.5}$$

If $cov(X,Y)=0$, then we say that X and Y are *uncorrelated*. It is clear that X and Y are uncorrelated whenever they are independent.
For any two random variables X and Y, it holds that

$$Var[X+Y] = Var[X] + 2cov[X,Y] + Var[Y].$$

A relevant property of the covariance is the following one:

$$| \, cov(X,Y) \, | \leq \sqrt{Var[X]}\sqrt{Var[Y]} = \sigma_X \sigma_Y, \tag{4.2.6}$$

where σ_X is called the *standard deviation* (or *mean square deviation*) of the random variable X.
Therefore, the *linear correlation coefficient*

$$\rho_{XY} = \frac{cov[X,Y]}{\sigma_X \sigma_Y}$$

is such that, for all random variables X, Y,

$$| \, \rho_{XY} \, | \leq 1.$$

CHAPTER 5

Expected utility and zero-utility premium

5.1 Utility theory

A lottery is represented by a random variable X, or equivalently by its cumulative distribution functions F. When necessary, we shall assume that a cumulative distribution function F has a bounded support $[a, b]$, $F(a) = 0$ and $F(b) = 1$.
When *utility theory* applies (Daniel Bernoulli (1700-1782)), the decision maker attaches a value $u(x)$ to his wealth x instead of just x, where

$$u : (\mathbb{R} \supset)X \longrightarrow \mathbb{R}$$

is called his or her utility function. In this way, all decisions related to random losses/gains are done by comparing the expected changes in utility. Although it is impossible to determine a person's utility function exactly, we can give some plausible properties of it. For example, more wealth would imply a higher utility level, so the utility function u should be an increasing function. It is also logical that "reasonable" decision makers are *risk averse*, which means that they prefer a fixed gain over a random gain with the same expected value. In particular, a risk averse agent (weakly) prefers a "sure" gain $c > 0$ to a "random" gain X paying $c - \epsilon$ and $c + \epsilon$ ($\epsilon > 0$, $c - \epsilon \geq 0$) both with probability $\frac{1}{2}$, or, more explicitly,

$$\mathbb{E}[u(c)] = u(c) \geq \frac{u(c-\epsilon) + u(c+\epsilon))}{2} = \mathbb{E}[u(X)].$$

Defining $x_1 = c - \epsilon$, $x_2 = c + \epsilon$, one gets

$$u\left(\frac{x_1 + x_2}{2}\right) \geq \frac{u(x_1) + u(x_2)}{2},$$

that is the condition of *mid-point concavity*, which is clearly implied by the concavity condition and equivalent to it in the case of a continuous function u.
In conclusion, a twice differentiable utility function u is assumed to be *increasing* ($u'(x) > 0$ for all x) and *concave*, i.e. the relative value of money will decrease while x increases ($u''(x) < 0$ for all x).

If expected utility applies corresponding to a choice of some utility function u, we introduce on the set of all possible random gains a *total preorder* $\precsim$ (i.e., a reflexive, transitive and total binary relation) defined as follows for all X, Y:

$$X \precsim Y \Leftrightarrow \mathbb{E}[u(X)] \leq \mathbb{E}[u(Y)]. \tag{5.1.1}$$

Unless the choice of a criterion of this kind appears reasonable according to the above considerations, there are violations of the expected utility principle, as the famous example below shows.

Example 5.1.1 (Allais paradox (1953))**.** Consider the following random gains:

$$X = 1.000.000 \text{ with probability } 1$$

$$Y = \begin{cases} 5.000.000 & \text{with probability } 0,10 \\ 1.000.000 & \text{with probability } 0,89 \\ 0 & \text{with probability } 0,01 \end{cases},$$

$$V = \begin{cases} 1.000.000 & \text{with probability } 0,11 \\ 0 & \text{with probability } 0,89 \end{cases},$$

$$W = \begin{cases} 5.000.000 & \text{with probability } 0,10 \\ 0 & \text{with probability } 0,90 \end{cases}.$$

It was observed that for many individuals X is preferred to Y, $\mathbb{E}[u(X)] > \mathbb{E}[u(Y)]$, and at the same time W is preferred to V, $\mathbb{E}[u(W)] > \mathbb{E}[u(V)]$. This leads to a contradiction, since it is immediate to check that one has

$$\begin{aligned} u(1.000.000) &> 0,10u(5.000.000) + 0,89u(1.000.000) > \\ &> 0,11u(1.000.000) + 0,89u(1.000.000) = u(1.000.000). \end{aligned}$$

Example 5.1.2 (Ellsberg paradox (1961)). Another well-known paradox is the *Ellsberg* paradox. Suppose that you are told that an urn contains 30 red balls and 60 more balls that are either blue or yellow. You don't know how many blue or how many yellow balls there are, but the number of blue balls plus the number of yellow ball equals 60 (they could be all blue or all yellow or any combination of the two). The balls are well mixed so that each individual ball is as likely to be drawn as any other. You are given a choice between bets A and B, where

$X_r = A =$ you get 100 if you pick a red ball and nothing otherwise,

$X_b = B =$ you get 100 if you pick a blue ball and nothing otherwise.

Many subjects in experiments state a strict preference for A over B: $B \prec A$.

Consider now the following bets:

$X_{r,y} = C =$ you get 100 if you pick a red or yellow ball and nothing otherwise,

$X_{b,y} = D =$ you get 100 if you pick a blue or yellow ball and nothing otherwise.

Do the axioms of expected utility constrain your ranking of C and D? Many subjects in experiments state the following ranking $B \prec A$ and $C \precsim D$. All such people violate the axioms of expected utility.

The fraction of red balls in the urn is $\frac{30}{90} = \frac{1}{3}$. Let p_2 be the fraction of blue balls and p_3 the fraction of yellow balls (either of these can be zero: all we know is that $p_2 + p_3 = \frac{60}{90} = \frac{2}{3}$). Then A, B, C and D can be viewed as the following lotteries:

$$A = \begin{pmatrix} 100 & 0 \\ \frac{1}{3} & p_2 + p_3 \end{pmatrix}, \quad B = \begin{pmatrix} 100 & 0 \\ p_2 & \frac{1}{3} + p_3 \end{pmatrix},$$

$$C = \begin{pmatrix} 100 & 0 \\ \frac{1}{3} + p_3 & p_2 \end{pmatrix}, \quad D = \begin{pmatrix} 100 & 0 \\ p_2 + p_3 = \frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

Let u be the normalized von Neumann-Morgenstern utility function that represents the individual's ranking, in such a way that then $u(100) = 1$ and $u(0) = 0$.
Thus, $\mathbb{E}[u(A)] = \frac{1}{3}$, $\mathbb{E}[u(B)] = p_2$, $\mathbb{E}[u(C)] = \frac{1}{3}+p_3$, and $\mathbb{E}[u(D)] = p_2+p_3 = \frac{2}{3}$. Hence, $B \prec A$ if and only if $p_2 < \frac{1}{3}$, which implies that $p_3 > \frac{1}{3}$, so that $\mathbb{E}[u(C)] = \frac{1}{3} + p_3 > \frac{2}{3} = \mathbb{E}[u(D)]$ and thus $D \prec C$.

5.2 Zero-utility principle

We begin this section with the following remark concerning a utility function $u(\cdot)$.

Remark 5.2.1. Without loss of generality, we can assume that the utility function of any individual is such that $u(0) = 0$. Indeed, any linear transformation $au(\cdot) + b$ with $a > 0$ is *equivalent* to $u(\cdot)$ (in the sense that the decision maker arrives at the same conclusion by using the expected utility principle). Further, it is not restrictive to assume that $u'(0) = 1$, due to the fact that, whenever $u'(0) > 0$, the utility function $v(\cdot)$ defined by

$$v(x) = \frac{u(x) - u(0)}{u'(0)}$$

is such that $v(0) = 0$ and $v'(0) = 1$.

From the insurer's viewpoint facing a random loss X, under expected utility corresponding to a utility function u such that $u(0) = 0$, an *acceptable premium* $\mathbb{P}(X)$ must satisfy the following inequality:

$$\mathbb{E}[u(\mathbb{P}(X) - X)] \geq 0 = u(0), \tag{5.2.1}$$

according to which the expected utility of the stochastic situation deriving from the acceptance of the contract is greater or equal to the sure situation of non-acceptance.

Generally speaking, a *premium* $\mathbb{P}$ is a mapping

$$\mathbb{P} : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\},$$

defined on a space $\mathcal{X}$ of *risks* (i.e., random payments or losses).

Definition 5.2.2. A premium $\mathbb{P}_u$ is said to be a *zero-utility* principle if it satisfies the condition

$$\mathbb{E}[u(\mathbb{P}_u(X) - X)] = 0 \quad \text{for every loss } X. \tag{5.2.2}$$

Therefore, the zero-utility principle $\mathbb{P} = \mathbb{P}_u$ is the minimum premium accepted by the insurer with utility function u.

Proposition 5.2.3. *The zero-utility principle $\mathbb{P}_u$ corresponding to a strictly increasing and concave utility function u is greater or equal than the expectation, i.e. for every loss X we have that $\mathbb{P}_u(X) \geq \mathbb{E}[X]$.*

Proof. Assume that $\mathbb{P}_u$ satisfies condition (5.2.2) above. From Jensen inequality (2.3.3), we have that

$$\mathbb{E}[u(\mathbb{P}_u(X) - X)] = 0 \leq u(\mathbb{P}_u(X) - \mathbb{E}[X]) = u(\mathbb{E}[\mathbb{P}_u(X) - X]).$$

Since $u(0) = 0$ and $u(\cdot)$ is increasing, we must have that $\mathbb{P}_u(X) - \mathbb{E}[X] \geq 0$. $\square$

Let us now consider the particular case of a quadratic utility u, that we shall write in the following form for the sake of convenience:

$$u(x) = x - \frac{1}{2B}x^2 \qquad \text{with } (x \leq B) \text{ and } (B > 0). \tag{5.2.3}$$

Let us first notice that the above condition $x \leq B$ is a consequence of the positivity off the derivative $u'(x) = 1 - \frac{x}{B}$. The second order *Taylor - Mac Laurin* polynomial of a generic utility function u of a risk averse agent is of the previous form, that is

$$u(x) \approx u(0) + u'(0)x + \frac{u''(0)}{2}x^2,$$

provided that the typical *normalization conditions* $u(0) = 0$, $u'(0) = 1$ hold true.
Our aim is now to approximate the zero-utility premium in the case when the utility function of the insurer is of the form (5.2.3). For the sake of brevity, let us simply write $\mathbb{P}_u(X) = \mathbb{P}_u$.

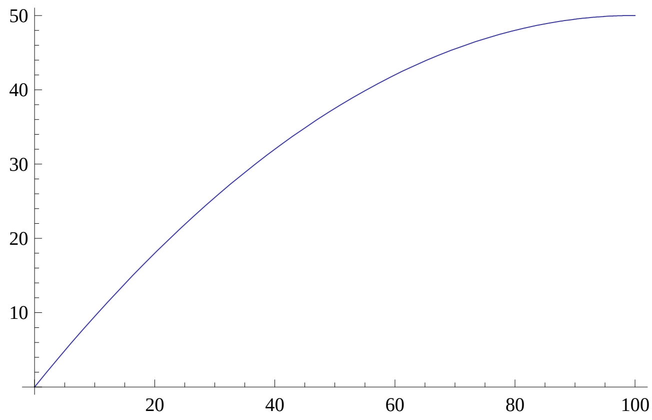


Figure 5.1 Graph of $u(x) = x - \frac{1}{200}x^2$ with $x \leq 100$

The condition

$$\mathbb{E}[u(\mathbb{P}_u - X)] = 0$$

is verified if and only if

$$\mathbb{E}[(\mathbb{P}_u - X) - \frac{1}{2B}(\mathbb{P}_u - X)^2] = 0.$$

Form the well known property

$$Var[Y] = \mathbb{E}[Y - \mathbb{E}[Y]]^2 = \mathbb{E}[Y^2] - [\mathbb{E}[Y]]^2$$

we arrive at

$$\mathbb{P}_u = \mathbb{E}[X] + \frac{1}{2B}[[\mathbb{E}[\mathbb{P}_u - X]]^2 + Var[\mathbb{P}_u - X]].$$

Solving the previous quadratic equation

$$\mathbb{P}_u^2 - 2\mathbb{P}_u(\mathbb{E}[X] + B) + Var[X] + \mathbb{E}^2[X] + 2B\mathbb{E}[X] = 0$$

with respect to $\mathbb{P}_u$, since we must have that $\mathbb{P}_u - X \leq B \Rightarrow \mathbb{P}_u \leq \mathbb{E}[X] + B$, we arrive at the unique solution

$$\begin{aligned} \mathbb{P}_u &= \mathbb{E}[X] + B - \sqrt{B^2 - Var[X]} \\ &= \mathbb{E}[X] + B\left(1 - \sqrt{1 - \frac{Var[X]}{B^2}}\right). \end{aligned}$$

From the Taylor - Mac Laurin approximation of the function

$$\sqrt{1+t} = (1+t)^{1/2},$$

since

$$(1+t)^{1/2} \approx 1 + \frac{t}{2},$$

we finally arrive at the following approximation:

$$\mathbb{P}_u \approx \mathbb{E}[X] + \frac{Var[X]}{2B}. \tag{5.2.4}$$

Therefore the zero-utility premium $\mathbb{P}_u$ is approximately expressed as the sum of the *equity premium* $\mathbb{E}[X]$ and a *loading* that is proportional to the variance of X.

CHAPTER 6

Premium principles

6.1 Basic definitions

Almost every human activity is related to some risk. The basis of *insurance* (and *reinsurance*) is the transfer of risks from the *policyholder* to the *insurer* (the *insurance company*). In particular, nonlife insurance provides compensations for losses. The insurance industry exists because people are willing to pay price for being insured (the *premium*).
Let X be the actual loss and let $c(X)$ denote the calculated *compensation*. In the sequel. for the sake of simplicity we shall assume that the calculated compensation $c(X)$ corresponding to a (single) loss X is $c(X) = X$. Therefore, the case of *full insurance* applies. Observe that the calculated compensation is not yet the compensation that the policyholder actually receives, it is usually reduced by a *deductible*.
There are several reasons for introducing deductibles:

a) loss prevention to lower the probability of claim occurrence;

b) loss reduction to lower the claim amount in case of a loss event;

c) avoidance of small claims (as administration cost are dominant when handling small claims);

d) premium reduction (the first three properties clearly simplify the insurer's risk management, in return the insurer can decrease the insurance premium).

Let $I(X) = h(X)$ denote the *actual compensation* paid to the claiming policyholder. In the sequel it will be referred to as the *indemnity*.
There are 3 main principles of deductibles that can be applied.

1. *Fixed amount deductible* d:

$$h_1(X) = (X - d)_+ = \max\{0; X - d\}; \qquad (6.1.1)$$

2. *Proportional deductible* β:

$$h_2(X) = (1 - \beta)X; \qquad (6.1.2)$$

3. *Franchise deductible* b:

$$h_3(X) = \chi_{\{X \geq b\}} X, \qquad (6.1.3)$$

   where $\chi_A : \mathbb{R} \to \mathbb{R}$ is the *indicator function* of any set $A \subset \mathbb{R}$, $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ if $x \notin A$.

In many case a *maximum coverage* M is fixed. For example, this is the case of *liability insurance*. Therefore, in the case of a fixed amount deductible d and maximal M, the indemnity is

$$h_1(X) = \begin{cases} 0 & \text{if } X \leq d \\ X - d & \text{if } d \leq X < M \\ M - d & \text{if } X \geq M \end{cases}, \qquad (6.1.4)$$

while in the case of a Franchise deductible b and maximal M, the indemnity is

$$h_3(X) = \begin{cases} 0 & \text{if } X \leq b \\ X & \text{if } b \leq X < M \\ M & \text{if } X \geq M \end{cases}. \qquad (6.1.5)$$

6.2 Equity premium

If X is a nonnegative random variable, then its *decumulative distribution function* $S(x)$ is defined to be

$$S(x) = Prob(X > x) = 1 - F(x) = 1 - Prob(X \leq x).$$

The following expression of $\mathbb{E}[X] = \int_0^{+\infty} x dF(x)$ may thought of as useful:

$$\mathbb{E}[X] = \int_0^\infty S(t)dt = \int_0^\infty Prob(X > t)dt. \qquad (6.2.1)$$

Figure 6.1 Indemnity in the case of fixed amount deductible $d = 10$ e maximal $M = 100$ $(c(X) = X)$

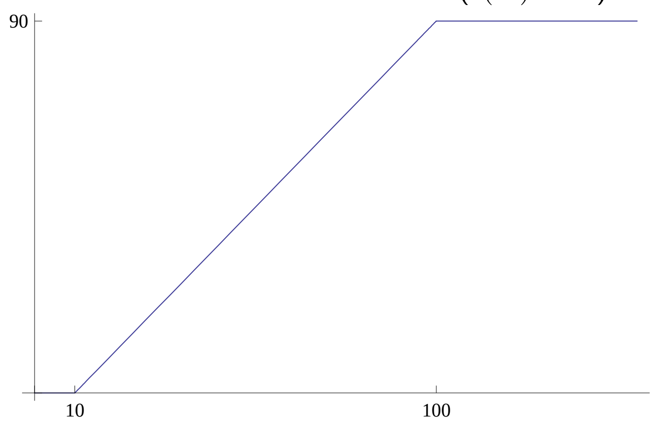
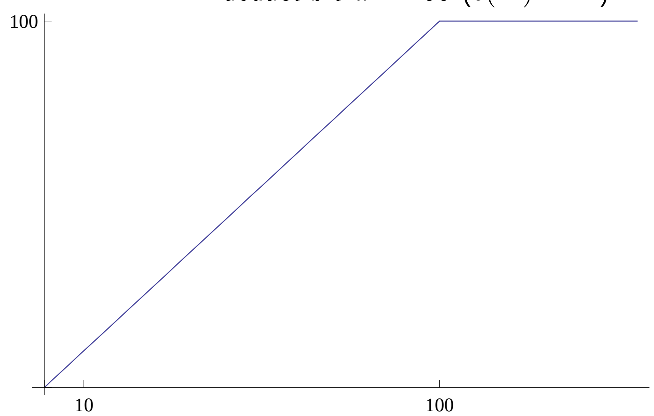


Figure 6.2 Loss charged to the policyholder in the case of fixed amount deductible $d = 100$ $(c(X) = X)$



In order to illustrate the validity of the previous formula, let us observe that

$$\int_0^\infty x dF(x) = \int_0^\infty dF(x) \int_0^x dt = \int_0^\infty dt \int_t^\infty dF(x) = \int_0^\infty S(t)dt.$$

Analogous considerations can be used in order to determine the expectation of the single loss X in the case of *stop-loss* (re)insurance. This is the case when, for a fixed amount deductible $d > 0$ (see definition (6.1.1)), the insurer only pays the difference $X - d$ (for example, in *vehicle insurance*). We have already seen that in this case, the *indemnity* $I(X)$ associated to a loss X (i.e., the amount that the insurer actually pays in the presence of a deductible d) is $(X-d)_+ = \max(X-d, 0)$ and, by using considerations analogous to those supporting formula (6.2.1), it can be shown that

$$\mathbb{E}[(X - d)_+] = \int_d^\infty (x - d)dF(x) = \int_d^\infty S(t)dt. \tag{6.2.2}$$

If in addition there is a *maximum coverage* M (for example, in *liability insurance* the insurer only pays $M-d$ if the loss X is greater than M), then the expectation of the random payment of the insurer is

$$\mathbb{E}[I(X)] = \int_d^M (x-d)dF(x) + (M-d)\int_M^\infty dF(x) = \int_d^M S(t)dt. \quad (6.2.3)$$

Indeed, we have that

$$\begin{aligned}\int_d^M (x-d)dF(x) + (M-d)\int_M^\infty dF(x) &= \int_d^M dF(x)\int_d^x dt \\ + \int_M^\infty dF(x)\int_d^M dt = \int_d^M dt\int_t^M dF(x) &+ \int_d^M dt\int_M^\infty dF(x) \\ &= \int_d^M dt\int_t^\infty dF(x).\end{aligned}$$

On the other hand, in the case of a Franchise deductible b and maximal M, we have that

$$\mathbb{E}[I(X)] = \int_b^M x dF(x) + M\int_M^\infty dF(x) = bS(b) + \int_b^M S(t)dt. \quad (6.2.4)$$

Indeed, in this latter case, we can observe that the expectation of the loss is obtained by the expression relative to the fixed amount deductible where $d=b$ by adding b with probability $S(b) = Prob(X > b)$.

6.3 Premium principles and their properties

The real premium $\mathbb{P}(X)$ is usually the sum of the equity premium $\mathbb{E}[X]$ and a *safety loading*. In the case when the loss X is a *bounded random variable* (so that its expectation and variance are both finite), a possible choice for the premium $\mathbb{P}(X)$ is the *variance principle*

$$\mathbb{P}(X) = \mathbb{E}[X] + \alpha Var[X], \quad (6.3.1)$$

where α is a positive coefficient. As we have seen before, this is an approximation of the zero-utility principle (when we adopt a quadratic utility function). So, the previous premium takes into account the *riskiness* of X by using the variance (which is a dispersion measure).
It should be noted that the variance principle is *additive for independent risks* since $\mathbb{P}(X+Y) = \mathbb{P}(X) + \mathbb{P}(Y)$ for independent risks X and Y.

Definition 6.3.1 (premium principles). A premium principle $\mathbb{P}$ is said to be:

1. the *Expected Value Principle*:
$$\mathbb{P}(X) = (1+\alpha)\mathbb{E}[X],$$
where the loading is $\alpha\mathbb{E}[X]$ with $\alpha > 0$;

2. the *Standard Deviation Principle*:
$$\mathbb{P}(X) = \mathbb{E}[X] + \alpha\sigma(X),$$
where $\sigma(X)$ is the *standard deviation* (i.e., $\sigma(X) = \sqrt{Var(X)}$);

3. the *Mean Value Principle*
$$\mathbb{P}(X) = v^{-1}\mathbb{E}[v(X)],$$
where v is a concave and increasing evaluation function;

4. the *Percentile Principle*:
$$\mathbb{P}(X) = \min\{t : F(t) \geq 1-\epsilon\} = \min\{t : S(t) \leq \epsilon\},$$
where the probability of a loss on contract is at most ϵ, $0 \leq \epsilon \leq 1$;

5. the *Choquet Premium Principle*:
$$\mathbb{P}(X) = \mathbb{E}^{g \circ Prob}[X] = \int_0^\infty g(S(t))dt,$$
where $g : [0,1] \to [0,1]$ is a concave and increasing function such that $g(0) = 0$ and $g(1) = 1$, which is referred to as a *probability distortion*[(a)];

6. the *Exponential Premium Principle*:
$$\mathbb{P}(X) = \frac{1}{\alpha} log\mathbb{E}[e^{\alpha X}],$$
for some $\alpha > 0$.

[(a)] The *Choquet integral* of a random variable X with respect to the distorted probability $\mu = g \circ Prob$ is defined to be $\int X dg \circ Prob = \int_0^{+\infty} \mu(X > t)dt = \int_0^{+\infty} g(S(t))dt$.

Remark 6.3.2 (exponential premium is zero-utility premium)**.** It should be noted that the exponential premium principle is the zero-utility principle (see the above condition (5.2.2)) when the utility function

$$u(x) = 1 - \frac{1}{\alpha}(1 - e^{-\alpha x})$$

is adopted. Indeed, we have that

$$\mathbb{E}[u(\mathbb{P}(X) - X)] = 1 = u(0) \quad \Leftrightarrow \quad \mathbb{E}[\frac{1}{\alpha}(1 - e^{-\alpha(\mathbb{P}(X)-X)})] = 0 \Leftrightarrow$$
$$\Leftrightarrow \quad e^{\alpha\mathbb{P}(X)} = \mathbb{E}[e^{\alpha X}] \Leftrightarrow \alpha\mathbb{P}(X) = log\mathbb{E}[e^{\alpha X}].$$

Further, the exponential premium is a particular case of the mean value principle if $v(x) = e^{\alpha x}$.

Definition 6.3.3 (properties of premium principles)**.** A premium principle $\mathbb{P}$ is said to satisfy

1. *Independence* (I) if $\mathbb{P}(X)$ depends only on the (de)cumulative distribution function of X (i.e., the premium depends only on the monetary loss of the insurable event and the probability that a given monetary loss occurs, not the cause of the monetary loss);

2. *Risk loading* (Rl) if $\mathbb{P}(X) \geq \mathbb{E}[X]$ for all X;

3. *No unjustified risk loading* (Nurl) if $\mathbb{P}(X) = c$ whenever a risk X is identically equal to a constant $c \geq 0$ (almost everywhere);

4. *Maximal loss* (Ml) (or *no rip-off*) if $\mathbb{P}(X) \leq ess \sup(X)$ for all X[a];

[a] Recall that, for a random variable X, the *essential supremum of X is defined to be*

$$ess \sup(X) = \inf\{\alpha \in \mathbb{R} : Prob\,(X^{-1}(]\alpha, \infty[)) = 0\}.$$

5. *Translation invariance* (Ti) if $\mathbb{P}(X + c) = \mathbb{P}(X) + c$ for all X and $c \in R_+$;
6. *Scale invariance* (Si) if $\mathbb{P}(bX) = b\mathbb{P}(X)$ for all X and all $b \geq 0$;
7. *Sublinearity* (Sl) if $\mathbb{P}$ is scale invariant and *subadditive*, i.e., for all X, Y,
$$\mathbb{P}(X + Y) \leq \mathbb{P}(X) + \mathbb{P}(Y);$$
8. *Additivity for independent risks* (Air) if $\mathbb{P}(X + Y) = \mathbb{P}(X) + \mathbb{P}(Y)$ for all X, Y such that X and Y are independent;
9. *Comonotone Additivity* (Ca) if $\mathbb{P}(X + Y) = \mathbb{P}(X) + \mathbb{P}(Y)$ for all comonotone X, Y (i.e., for all X, Y such that $(X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \geq 0$ for $\omega_1, \omega_2 \in \Omega$;
10. *Iterativity* (It) if $\mathbb{P}(\mathbb{P}(X \mid Y) = \mathbb{P}(X)$ for all X, Y.

Remark 6.3.4 (on the properties satisfied by the premium principles)**.** All the premium principles we have presented satisfy independence (I).

Additivity for independent risks is satisfied by the variance principle, the expected value principle and the exponential principle.

The maximal loss principle is verified by the exponential principle, the mean value principle and the percentile principle.

Further, all the premium principles except for the percentile principle satisfy the property of risk loading.

We have that the variance and standard deviation principle are both translation invariant.

The Choquet premium principle is scale invariant and translation invariant.

The iterativity property is satisfied by $\mathbb{P}(X) = \mathbb{E}[X]$ and by the mean value principle.

Proposition 6.3.5. *The standard deviation principle is sublinear.*

Proof. We have that

$$\mathbb{P}(X+Y) = \mathbb{E}[X+Y] + \alpha\sqrt{Var[X+Y]} \leq \mathbb{E}[X] + \alpha\sigma_X + \mathbb{E}[Y] + \alpha\sigma_Y$$

since

$$Var[X+Y] = Var[X] + 2cov(X,Y) + Var[Y] \leq Var[X] + 2\sigma_X\sigma_Y + Var[Y],$$

as a consequence of the well known property

$$\mid cov(X,Y) \mid \leq \sigma_X\sigma_Y.$$

□

Proposition 6.3.6 (iterativity of exponential premium)**.** *The exponential premium principle satisfies the iterativity property.*

Proof. Consider that

$$\mathbb{P}(\mathbb{P}(X \mid Y)) = \frac{1}{\alpha}log\mathbb{E}[e^{\alpha\mathbb{P}(X|Y)}] = \frac{1}{\alpha}log\mathbb{E}[exp(\alpha\frac{1}{\alpha}log\mathbb{E}[e^{\alpha X} \mid Y])] =$$
$$= \frac{1}{\alpha}log\mathbb{E}[\mathbb{E}[e^{\alpha X} \mid Y]] = \frac{1}{\alpha}log\mathbb{E}[e^{\alpha X}] = \mathbb{P}(X).$$

Here, we have used the well-known property of the *conditional expectation*, according to which

$$\mathbb{E}(\mathbb{E}[X \mid Y]) = \mathbb{E}[X],$$

and more generally, for any real-valued function ψ,

$$\mathbb{E}(\mathbb{E}[\psi(X) \mid Y]) = \mathbb{E}[\psi(X)].$$

□

We recall that a *capacity* μ on $\mathcal{A}$ is a function from $\mathcal{A}$ into $[0,1]$ such that $\mu(\emptyset) = 0$, $\mu(\Omega) = 1$ and $\mu(A) \leq \mu(B)$ for all $A \subseteq B$, $A, B \in \mathcal{A}$. If g is a *probability distortion* (i.e., $g : [0,1] \longrightarrow [0,1]$ is non-decreasing and $g(0) = 0$, $g(1) = 1$), then it is clear that $\mu = g \circ \mathcal{P}$ is a capacity (*distorted probability*) on $\mathcal{A}$.
We shall denote by $\int_\Omega X d\mu$ the *Choquet integral* of $X \in \mathbb{X}^+(\Omega, \mathcal{A}, \mathcal{P})$ with respect to a capacity μ on $\mathcal{A}$, namely

$$\int_\Omega X d\mu = \int_0^\infty \mu(X > t)dt.$$

Proposition 6.3.7 (comonotone additivity of Choquet premium)**.** *The Choquet premium principle satisfies comonotone additivity.*

Proof. Consider any set C of comonotone random variables. Then consider the family of half-bounded closed intervals $\{\omega \in \Omega : X(\omega) > \alpha\}$. $X \in C$, $\alpha \in \mathbb{R}$. By comonotonicity this family is *nested*, i.e., for each pair of such intervals, one is a subset of the other. On this family we define $P \equiv \mu$. By routine methods from measure theory, P can in a unique manner be extended to an additive probability measure on the algebra generated by the family. Since P coincides with μ on the family just defined, according to the definition of the Choquet integral the integrals w.r.t. μ and P coincide on C. □

6.4 Solutions to Allais and Ellsberg paradoxes

The Choquet integral of a nonnegative discrete random variable (monetary lottery)

$$X = \begin{pmatrix} x_1 & x_2 & ... & x_n \\ p_1 & p_2 & ... & p_n \end{pmatrix},$$

such that $0 \leq x_1 < x_2 < ... < x_n$ with respect to a distorted probability $\mu = g \circ Prob$ is defined to be

$$\mathbb{E}^\mu[X] = \sum_{i=1}^n \pi_i x_i,$$

where

$\pi_1 = g(p_1), \pi_i = g(p_1 + ... + p_{i-1} + p_i) - g(p_1 + ... + p_{i-1})$ for every $2 \leq i \leq n$.

Solution to Allais paradox

Consider the Allais paradox illustrated in Example 5.1.1. For the sake of convenience, we restate the gains:

$$X = \;\; 1.000.000 \quad \text{with probability } 1,$$

$$Y = \begin{cases} 5.000.000 & \text{with probability } 0,10 \\ 1.000.000 & \text{with probability } 0,89 \\ 0 & \text{with probability } 0,01 \end{cases},$$

$$V = \begin{cases} 1.000.000 & \text{with probability } 0,11 \\ 0 & \text{with probability } 0,89 \end{cases},$$

$$W = \begin{cases} 5.000.000 & \text{with probability } 0,10 \\ 0 & \text{with probability } 0,90 \end{cases}.$$

The paradox takes place when $X \succ Y$, $W \succ V$ and expected utility applies. Consider a distortion function g such that $g(0,90) = 0.99$, $g(0,01) = 0,005$, $g(0,1) = 0,05$, and $g(0,89) = 0,96$. Define $\mu = g \circ Prob$ and consider the utility functional $U(\cdot) = \mathbb{E}^{\mu}[\cdot]$. Then we have that

$$U(X) = 1.000.000,$$

$$U(Y) = 1.000.000(g(0,90) - g(0,01))) + 5.000.000(1 - g(0,90)) = 990.000,$$

$$U(W) = 5.000.000(1 - g(0,90)) = 50.000,$$

$$U(V) = 1.000.000(1 - g(0,89)) = 40.000.$$

Therefore, we have that $U(X) > U(Y)$ and $U(W) > U(V)$.

Solution to Ellsberg paradox

Consider the Ellsberg paradox in Example 5.1.2. For the sake of convenience, we restate the gains.

$X_r = A =$ you get 100 if you pick a red ball and nothing otherwise,

$X_b = B =$ you get 100 if you pick a blue ball and nothing otherwise,

$X_{r,y} = C =$ you get 100 if you pick a red or yellow ball and nothing otherwise,

$X_{b,y} = D =$ you get 100 if you pick a blue or yellow ball and nothing otherwise.

The paradox takes place when $X_r \succ X_b$ and $X_{b,y} \succsim X_{r,y}$. Define a capacity μ such that $\mu(\{r\}) = \frac{1}{3}$, $\mu(\{b\}) = \mu(\{y\}) = \frac{2}{9}$, $\mu(\{b,y\}) = \frac{2}{3}$, $\mu(\{r,b\}) = \mu(\{r,y\}) = \frac{5}{9}$. Then the Choquet integral is as follows:

$$\mathbb{E}^{\mu}[X_r] = \frac{1}{3} \cdot 100,\ \mathbb{E}^{\mu}[X_b] = \frac{2}{9} \cdot 100,\ \mathbb{E}^{\mu}[X_{r,y}] = \frac{5}{9} \cdot 100,\ \mathbb{E}^{\mu}[X_{b,y}] = \frac{2}{3} \cdot 100.$$

Therefore, we have that $\mathbb{E}^{\mu}[X_r] > [\mathbb{E}^{\mu}[X_b]$, and $\mathbb{E}^{\mu}[X_{b,y}] > \mathbb{E}^{\mu}[X_{r,y}]$. So, there is no contradiction.

6.5 Ordering of risks

We introduce the basic definitions concerning the most classical orderings of risks.

Definition 6.5.1. For any two nonnegative *risks* X, Y we say that

1. *Y dominates X in the sense of the first order stochastic dominance*, written $X \leq_{st} Y$ if and only if the following condition is verified for all $t \in \mathbb{R}_+$:

$$Prob(X \leq t) \geq Prob(Y \leq t) \Leftrightarrow F_X(t) \geq F_Y(t) \Leftrightarrow S_X(t) \leq S_Y(t);$$

2. *Y dominates X in the sense of the stop-loss order*, written $X \leq_{sl} Y$, if the following condition is verified for all $a \in \mathbb{R}_+$:

$$\mathbb{E}[(X-a)_+] \leq \mathbb{E}[(Y-a)_+] \Leftrightarrow \int_a^\infty S_X(t)dt \leq \int_a^\infty S_Y(t)dt.$$

Theorem 6.5.2 (characterization of first order stochastic dominance)**.** *Y dominates X in the sense of the first order stochastic dominance ($X \leq_{st} Y$) if and only if $\mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$ for every nondecreasing function g.*

Proof. We prove the theorem assuming that g is differentiable, and X, Y have compact support, so that there exists $a < b$ ($a, b \in \mathbb{R}$) such that $F_X(a) = F_Y(a) = 0$, and $F_X(b) = F_Y(b) = 1$. Integrating by parts, we get

$$\mathbb{E}[g(Y)] - \mathbb{E}[g(X)] = \int_a^b g'(t)[F_X(t) - F_Y(t)]dt.$$

Assume that

$$X \leq_{st} Y$$

. If g is nondecreasing, then $g'(t) \geq 0$ for every $t \in [a.b]$, and the difference $\mathbb{E}[g(Y)] - \mathbb{E}[g(X)$ is nonnegative since $F_X(t) \geq F_Y(t)$ for all $t \in [a, b]$. Conversely, assume by contraposition that $F_X(t) < F_Y(t)$ on some interval contained in $[a, b]$. Then we can pick a nondecreasing function g that is strictly increasing only on such interval, and constant elsewhere, so that the expression $\mathbb{E}[g(Y)] - \mathbb{E}[g(X)$ won't be nonnegative for every nondecreasing g. □

Theorem 6.5.3 (characterization of stop-loss order)**.** *Y dominates X in the sense of the stop-loss order ($X \leq_{sl} Y$) if and only if $\mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$ for every nondecreasing convex function g.*

Proof. If X and Y are two random variables such that $\mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$ for every nondecreasing convex function g, then in particular such property holds true when $g(x) = (x - d)_+$, so that $X \leq_{sl} Y$.

Conversely, consider two random variables X, Y such that $X \leq_{sl} Y$. As in the proof of the previous Theorem 6.5.2, we assume that that g is differentiable, and X, Y have compact support, so that there exists $a < b$ $(a, b \in \mathbb{R})$ such that $F_X(a) = F_Y(a) = 0$, and $F_X(b) = F_Y(b) = 1$. Then, if g is any real-valued function, we have that

$$\mathbb{E}[g(Y)] - \mathbb{E}[g(X)] = \int_a^b g'(t)[F_X(t) - F_Y(t)]dt.$$

Since $X \leq_{sl} Y$, we have, in particular, that $\int_a^b [F_X(t) - F_Y(t)]dt \geq 0$. Convexity of g implies that g' is nondecreasing on $[a, b]$, and therefore

$$\mathbb{E}[g(Y)] - \mathbb{E}[g(X)] \geq g'(a) \int_a^b [F_X(t) - F_Y(t)]dt \geq 0,$$

as a consequence of the previous considerations and the fact that g' is nonnegative. Therefore, $\mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$ for every nondecreasing convex function g, and the proof is complete. □

It is clear that we can now introduce the monotonicity of a premium functional $\mathbb{P}$ with respect to a stochastic partial order of the kind introduced above.

Definition 6.5.4 (monotonicity with respect to stochastic orders)**.** A premium principle $\mathbb{P}$ is said to satisfy

1. *Monotonicity* if $\mathbb{P}(X) \leq \mathbb{P}(Y)$ for all X, Y such that $X \leq Y$;
2. *Monotonicity with respect to first order stochastic dominance* if $\mathbb{P}(X) \leq \mathbb{P}(Y)$ for all X, Y such that $X \leq_{st} Y$;
3. *Monotonicity with respect to stop-loss order* if $\mathbb{P}(X) \leq \mathbb{P}(Y)$ for all X, Y such that $X \leq_{sl} Y$.

Remark 6.5.5. It is immediate to check that, for all risk X, Y,

$$X \leq Y \Rightarrow X \leq_{st} Y \Rightarrow X \leq_{sl} Y.$$

Therefore, if a premium functional $\mathbb{P}$ is monotone with respect to stop loss order, then it is monotone with respect to first order stochastic dominance. Indeed, in this case, for all risks X, Y such that $X \leq_{st} Y$, we have that $X \leq_{sl} Y$, implying that $\mathbb{P}(X) \leq \mathbb{P}(Y)$. It is clear that in this case the premium functional $\mathbb{P}$ also satisfies monotonicity.

We prove the following interesting property.

Proposition 6.5.6. *The zero-utility principle P_u defined in (5.2.2) satisfies monotonicity with respect to the stop-loss order whenever u is increasing and concave.*

Proof. Consider any two risks X, Y such that $X \leq_{sl} Y$. Then we have that $\mathbb{E}[u(c - X)] \geq \mathbb{E}[u(c - Y)]$ for every $c \in \mathbb{R}_+$, since $g(x) = -u(c - x)$ is an increasing and convex function of x. Therefore, $(0 =)\ \mathbb{E}[u(P_u(Y) - Y)] = \mathbb{E}[u(P_u(X) - X)] \geq \mathbb{E}[u(P_u(X) - Y)]$ implies that $P_u(X) \leq P_u(Y)$. □

Finally, it is immediate to check that the exponential premium principle $\mathbb{P}(X) = \frac{1}{\alpha} log \mathbb{E}[e^{\alpha X}]$ also satisfies monotonicity with respect to the stop-loss order, since $g(x) = e^{\alpha x}$ is an increasing and convex function of x. Further, the Choquet premium principle $\mathbb{P}(X) = \int_0^\infty g(S_X(t))dt$ satisfies monotonicity with respect to first order stochastic dominance, since $X \leq_{st} Y$ implies $S_X(t) \leq S_Y(t)$ for all $t \in \mathbb{R}_+$, which in turn implies that $\mathbb{P}(X) \leq \mathbb{P}(Y)$ since the probability distortion g is increasing.

We recall that a random variable X has a Bernoulli distribution with parameter $0 \leq p \leq 1$ ($X \sim Bernoulli(p)$) if X takes values 1 with probability p and 0 with probability $1 - p$. It is immediate to check that, for a random variable $X \sim Bernoulli(p)$), $\mathbb{E}[X] = p$ and $Var[X] = p(1 - p)$.

In the following example we show that the standard deviation premium principle may fail to respect monotonicity with respect to first order stochastic dominance.

Example 6.5.7. *Let* $\mathbb{P}(X) = \mathbb{E}[X] + \alpha\sqrt{var[X]}$ *be the standard deviation premium principle with* $\alpha > 1$. *If* $X \sim Bernoulli(\frac{1}{2})$ *and* $Y \equiv 1$, *then it is clear that* $Prob(X \leq Y) = 1$ *implies that* $X \leq_{st} Y$. *On the other hand* $\mathbb{P}(X) = \frac{1}{2} + \alpha\frac{1}{2}$ *and* $\mathbb{P}(Y) = 1$ *implies that* $\mathbb{P}(X) > \mathbb{P}(Y)$ *since* $\alpha > 1$.

Other properties of stop-loss order are in order now.

Proposition 6.5.8. *If for risks* X *and* Y *we have* $X \leq_{sl} Y$, *and risk* Z *is independent of* X *and* Y, *then* $X + Z \leq_{sl} Y + Z$. *If further* S_n *is the sum of* n *independent copies of* X *and* T_n *is the sum of* n *independent copies of* Y, *then* $S_n \leq_{sl} T_n$.

Proof. The first stochastic inequality can be proved by using the relation:

$$\mathbb{E}[X + Z - d)_+] = \int_0^\infty \mathbb{E}[(X + z - d)_+]dF_Z(z).$$

The second follows by iterating the first inequality. □

Proposition 6.5.9. *If* $M \sim Bernoulli(p)$, N *is a counting random variable with* $\mathbb{E}[N] \geq p$, X_1, X_2, . . . *are copies of a risk* X, *and all these risks are independent, then we have*

$$MX \leq_{sl} X_1 + X_2 + ... + X_N.$$

Proof. Let us first show that if $d \geq 0$ and $x_i \geq 0$ for all indexes i, the following relation always holds:

$$(x_1 + .. + x_n - d)_+ \geq (x_1 - d)_+ + ... + (x_n - d)_+.$$

Clearly, we can limit ourselves to the consideration of the case when the right hand side is greater than zero. Without loss of generality, assume that the first term is positive, i.e. $x_1 > d$. Then trhe above displayed inequality is equivalent to the following one:

$$x_2 + .. + x_n \geq (x_2 - d)_+ + ... + (x_n - d)_+,$$

and this is always true. If $p_n = Prob(N = n)$, replacing realizations by random variables and taking the expectation we get:

$$\mathbb{E}[X_1 + .. + X_N - d]_+ = \sum_{n=1}^{\infty} p_n \mathbb{E}[X_1 + .. + X_n - d]_+ \geq$$
$$\sum_{n=1}^{\infty} p_n \mathbb{E}[(X_1 - d)_+ + ... + (X_n - d)_+] = \sum_{n=1}^{\infty} n p_n \mathbb{E}[(X - d)_+] \geq$$
$$p\mathbb{E}[(X - d)_+] = \mathbb{E}[(MX - d)_+].$$

Notice that, in the last inequality, we used the fact that

$$\mathbb{E}[N] = \sum_{n=1}^{\infty} n p_n \geq p = \mathbb{E}[M].$$

□

CHAPTER 7

Reinsurance

7.1 Reinsurance types

Reinsurance occurs when the original insurer cedes a part of a risk X to another insurer, the *reinsurer*. Reinsurance, may concern a single risk or, more frequently, a *portfolio* of risks. We have that

$$X = \sum_{h=1}^{n} X^{(h)}$$

is the risk expressed as the sum of n individual risks $X^{(h)}$, each with a random number N_h of claims $X_i^{(h)}$ in the coverage period (i.e., $X^{(h)} = \sum_{i=1}^{N_h} X_i^{(h)}$). Then, a *reinsurance treaty* is specified whenever the part $\Gamma^{(h)}$ of the risk of the risk $X^{(h)}$ remaining with the original insurer is specified. Therefore, the functions

$$\phi_h : X^{(h)} \longrightarrow \Gamma^{(h)}$$

need to be specified.

Reinsurance can be either *proportional* or *non-proportional*. It is said to be proportional if either n numbers $0 \leq a_h \leq 1$ are determined in such a way that

$$\Gamma^{(h)} = a_h X^{(h)} \quad (h = 1, ..., n),$$

this case being indicated as the *individual case*, or else a single number $0 \leq a \leq 1$ is specified in such a way that

$$\Gamma^{(h)} = a X^{(h)} \quad (h = 1, ..., n),$$

this case being indicated as the *global case* (that is, $a_h \equiv a$ in the individual case). Sometimes the latter case is referred to as *quota share* and the former *surplus*.
It should be noted that, in the global case, the cumulative distribution function of the single loss that is retained by the initial insurer is determined as follows:

$$F_{aX}(x) = Prob(aX \leq x) = Prob(X \leq \frac{x}{a}) = F_X\left(\frac{x}{a}\right).$$

On the other hand, the non-proportional reinsurance can be of the *excess of loss* type, in which case we have that

$$\Gamma^{(h)} = \sum_{i=1}^{N_h} \min(X_i^{(h)}, L^{(h)}),$$

where a *retention level* (*priority*) $L^{(h)}$ is assigned for every single loss of the $h-th$ risk $X^{(h)}$, or else of the *stop-loss* type, in which case a unique retention level L is assigned and the reinsurer pays the top part $(X-L)_+$, that is

$$\Gamma = \min(X, L) = \min(\sum_{h=1}^{n} X^{(h)}, L),$$

where Γ is the retained payment over the whole portfolio.
Therefore, in the stop-loss case, the expected value of the single and total retained losses relative to the h-th risk, namely $X_i^{(h)ret}$ and $X^{(h)ret}$, are determined as follows (under the usual hypothesis when, for every $h \in \{1, ..., n\}$, the random variables $X_i^{(h)}$ are independent and identically distributed and N_h is independent from $X_1^{(h)}, X_2^{(h)}, ...$):

$$\mathbb{E}[X_i^{(h)ret}] = \mathbb{E}[\min(X_i^{(h)}, L)], \;\; \mathbb{E}[X^{(h)ret}] = \mathbb{E}[N_h]\mathbb{E}[\min(X_i^{(h)}, L)].$$

For the sake of exposition, let us concentrate our attention on a risk $X = X^{(h)}$, with single losses $X_i^{(h)} = X_i$. In this latter case, from the point of view of the reinsurer, the main two variables that come into consideration are the single loss

$$X_L = X - L \mid X > L$$

and the number of claims

$$N_L = \sum_{i=1}^{N} \chi_{\{X_i > L\}},$$

where as usual $\chi_{\{X_i > L\}}$ is the indicator function of the event $X_i > L$.

Since the cumulative distribution function of X_L is

$$\begin{aligned} F_{X_L}(x) &= Prob(X_L \leq x) = Prob(X - L \leq x \mid X > L) = \\ &= Prob(L < X \leq x + L \mid X > L) = \frac{F_X(x+L) - F_X(L)}{S_X(L)}, \end{aligned}$$

we have that the expectation of X_L is determined as follows:

$$\mathbb{E}[X_L] = \mathbb{E}[X - L \mid X > L] = \frac{\mathbb{E}[X] - \mathbb{E}[\min(X, L)]}{S_X(L)}.$$

The expected number of the claims $\mathbb{E}[N_L]$ can be determined as follows:

$$\mathbb{E}[N_L] = \mathbb{E}[\sum_{i=1}^{N} \chi_{\{X_i > L\}}] = \mathbb{E}[N] S_X(L).$$

Therefore, the expected payment of the reinsurer (equity premium) is

$$\begin{aligned} \mathbb{E}[\sum X_L] &= \mathbb{E}[X_L]\mathbb{E}[N_L] = \frac{\mathbb{E}[X] - \mathbb{E}[\min(X, L)]}{S_X(L)} \mathbb{E}[N] S_X(L) = \\ &= (\mathbb{E}[X] - \mathbb{E}[\min(X, L)])\mathbb{E}[N]. \end{aligned}$$

7.2 A motivation for reinsurance

We now present a possible motivation for reinsurance which is based on expected utility. From the considerations concerning the zero-utility principle, when the utility is quadratic and in particular we have that

$$u(x) = x - \frac{x^2}{2B} \quad (x \leq B),$$

the minimum safety loading $m = \mathbb{P}(X) - \mathbb{E}[X]$ appearing in the evaluation of a premium $\mathbb{P}(X)$ has been found as the smallest root of the equation

$$m - \frac{m^2 + \sigma^2}{2B} = 0,$$

where $\sigma^2 = Var[X]$.

If we now consider n risks $X^{(1)}, ..., X^{(n)}$ and m_h is the minimum safety loading appearing in the computation of the pure premium P_h for the risk $X^{(h)}$, then the portfolio $X = \sum_{h=1}^{n} X^{(h)}$ is acceptable for the insurer if the following condition is verified:

$$\sum_{h=1}^{n} m_h - \frac{(\sum_{h=1}^{n} m_h)^2 + Var[\sum_{h=1}^{n} X^{(h)}]}{2B} \geq 0.$$

In order to guarantee that the above inequality is satisfied, the insurer needs to reduce the correlation among the risks, since

$$Var[\sum_{h=1}^{n} X^{(h)}] = \sum_{h=1}^{n} Var[X^{(h)}] + 2\sum_{h>k} cov(X^{(h)}, X^{(k)}).$$

This may be performed by means of a reinsurance policy.

7.3 Stop-loss reinsurance

Let us observe that the expectation of the total loss for the reinsurer in the case of stop-loss reinsurance with retention level L is determined as follows

$$\pi_X(L) = \mathbb{E}[(X-L)_+] = \mathbb{E}[\max(X-L,0)] = \int_L^{\infty} S_X(t)dt, \tag{7.3.1}$$

with $S_X(\cdot)$ the decumulative distribution function of

$$X = \sum_{h=1}^{n} X^{(h)}.$$

The function $\pi_X(L)$ is said to be the *stop-loss transform* of X. The following theorem describes the properties of the stop-loss transform.

Theorem 7.3.1. *The stop-loss transform $\pi_X(L)$ is a continuous and convex function that is strictly decreasing in the retention level L as long as $F_X(L) < 1$. If $F_X(L) = 1$ then $\pi_X(L) = 0$. Further, $\pi_X(\infty) = 0$*

Proof. From the representation (7.3.1), we only need to prove that $\pi_X(L)$ is strictly decreasing in the retention level L as long as $F_X(L) < 1$ and convex. We have that $\pi'_X(L) = F_X(L) - 1$ is actually a right-hand derivative since $F_X(\cdot)$ is continuous to the right at every point. If $F_X(L) < 1$, then $\pi'_X(L) < 0$ implies that $\pi_X(L)$ is strictly decreasing. Finally, the fact that $F_X(\cdot)$ is non-decreasing (increasing) implies that $\pi'_X(L)$ is increasing, and therefore $\pi_X(L)$ is convex. □

The fact that stop-loss reinsurance may be viewed as the ideal type of reinsurance is illustrated in the following theorem. If some reinsurance treaty holds, denote by $I_r(X)$ the payment of the reinsurer on the global risk X ($0 \leq I_r(x) \leq x$ for every outcome x), so that $X_{ret} = X - I_r(X)$ is the retained payment on X of the original insurer.

Theorem 7.3.2. *If* $\mathbb{E}[I_r(X)] = \mathbb{E}[(X-L)_+]$, *then we have that*

$$Var[X - I_r(X)] \geq Var[X - (X-L)_+].$$

Proof. First observe that, since from Theorem 7.3.1 the stop-loss transform $\pi_X(L)$ is continuous and $\pi_X(\infty) = 0$, for every reinsurer's payment $I_r(X)$ it is always possible to find a retention level L such that

$$\mathbb{E}[I_r(X)] = \mathbb{E}[(X-L)_+] = \pi_X(L).$$

If this happens, then we have that $Var[X - I_r(X)] \geq Var[X - (X-L)_+]$ if and only if $\mathbb{E}[(X - I_r(X))^2] \geq \mathbb{E}[(X - (X-L)_+)^2]$ or equivalently $\mathbb{E}[(X - I_r(X) - L)^2] \geq \mathbb{E}[(X - (X-L)_+ - L)^2]$. This latter inequality would be implied by the following condition:

$$\mid X - I_r(X) - L \mid \geq \mid X - (X-L)_+ - L \mid \quad \text{with probability one.}$$

Such an inequality is satisfied. Indeed, if $X \geq L$ then $X - (X-L)_+ = L$. If it happens that $X < L$, then $X - (X-L)_+ = X$ and this implies that $X - I_r(X) - L \leq X - L = X - (X-L)_+ - L < 0$. This consideration completes the proof. □

7.4 Optimal reinsurance: the problem

In this section we describe the optimization problem of the original insurer who looks for an optimal reinsurance contract of some prescribed type (proportional or nonproportional). Denote by $G^{(r)}$ the *random gain* of the original insurer in the presence of a reinsurance contract. Then we have that

$$G^{(r)} = P - P_r + C - \Gamma, \tag{7.4.1}$$

where

1. P is the total premium,
2. P_r is the premium of the reinsurer,
3. C is the *retrocession* paid by the reinsurer to the original insurer,
4. $\Gamma = \sum_{i=1}^{n} \Gamma_i$ is the total retention of the original insurer in the presence of the reinsurance treaty.

If u is the utility function of the original insurer, then the optimal insurance contract is determined as the solution of the problem

$$\max\ \mathbb{E}[u(G^{(r)})] = \max\ \mathbb{E}[u(P - P_r + C - \Gamma)]. \tag{7.4.2}$$

For the sake of simplicity, it can be assumed that the reinsurer adopts the expectation principle, and further that

$$P_r - C = (1+\eta)\mathbb{E}[X_r],$$

where $X_r = X - \Gamma$ is the risk remaining with the reinsurer. In this case, the retrocession C is proportional to the expectation $\mathbb{E}[X_r]$ and it is absorbed by the loading of the reinsurer.
In the particular case of a quadratic utility $u(x) = x - \frac{x^2}{2B}$, the problem therefore reduces to

$$\max \left\{ \mathbb{E}[P - \Gamma - (1+\eta)\mathbb{E}[X_r]] - \frac{\mathbb{E}[(P - \Gamma - (1+\eta)\mathbb{E}[X_r])^2]}{2B} \right\},$$

where, as usual, we have that

$$\mathbb{E}[(P - \Gamma - (1+\eta)\mathbb{E}[X_r])^2] = var[\Gamma] + (\mathbb{E}[(P - \Gamma - (1+\eta)\mathbb{E}[X_r])])^2 .$$

In the stop-loss case with retention level L, we have that $\Gamma = \min(X, L)$ and therefore $\mathbb{E}[X_r] = \int_L^\infty S_X(t)dt$. It can be shown that

$$Var[\Gamma] = Var[\min(X, L)] = 2\int_0^L tS_X(t)dt - \left(\int_0^L S_X(t)dt\right)^2 .$$

Indeed, we have that

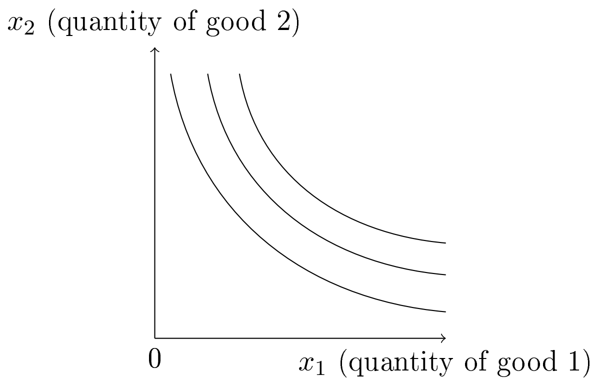
$$\begin{aligned}
\mathbb{E}[(\min(X, L))^2] &= \int_0^L x^2 dF_X(x) + L^2 \int_L^\infty dF_X(x) = \\
&= 2\left(\int_0^L dF_X(x) \int_0^x tdt + \int_L^\infty dF_X(x) \int_0^L tdt\right) = \\
&= 2\left(\int_0^L tdt \int_t^L dF_X(x) + \int_0^L tdt \int_L^\infty dF_X(x)\right) = \\
&= 2\int_0^L tS_X(t)dt.
\end{aligned}$$

Finally, in the cases of the global proportional scheme, when $\Gamma = aX = a\sum_{i=1}^n X_i$, we have that $Var[\Gamma] = a^2 Var[X]$ and $\mathbb{E}[X_r] = (1-a)\mathbb{E}[X]$.

CHAPTER 8

Expected utility and insurance in a two state model

8.1 Expected utility and insurance in a two state model



The simplest situation when dealing with expected utility *under uncertainty* is that of a two state set up with $p_1 = p$ and $p_2 = 1-p$. Therefore, in this case we consider a space $\Omega = \{\omega_1, \omega_2\}$ of elementary events, with $Prob(\omega_1) = p$ and $Prob(\omega_2) = 1 - p$. Individuals maximize

$$\mathbb{E}[u(X)] = pu(X(\omega_1)) + (1-p)u(X(\omega_2)) = pu(x_1) + (1-p)u(x_2),$$

$$X = \begin{pmatrix} x_1 & x_2 \\ p & 1-p \end{pmatrix}$$

An interesting way to represent preferences in this case is with a standard consumer model over two goods, consumption in state 1 and consumption in state 2. As usual, an *indifference curve* is given implicitly by setting utility to a fixed value and treating one variable (say x_2) as a function of the other (x_1) or formally

$$pu(x_1) + (1-p)u(x_2) = \bar{u}.$$

The above condition defines implicitly a function $x_2 = x_2^{\bar{u}}(x_1)$ of x_2 in terms of x_1. Such function clearly depends on the *utility level* $\bar{u}$.
Differentiating this condition implicitly once gives the condition

$$\frac{d\,x_2^{\bar{u}}(x_1)}{d\,x_1} = -\frac{p}{1-p}\,\frac{u'(x_1)}{u'(x_2)}.$$

Along the bisector of the first quadrant, where $x_1 = x_2$, then $u'(x_1) = u'(x_1)$, and so $\frac{d\,x_2^{\bar{u}}(x_1)}{d\,x_1} = -\frac{p}{1-p}$, the opposite of the *odds-ratio* $\frac{p}{1-p}$ (the proportion of state 1's to state 2's), no matter what the utility function looks like. This is one of the stronger implications of expected utility theory.
Notice that

$$MRS^{\bar{u}}_{x_1x_2} = -\frac{dx_2^{\bar{u}}(x_1)}{d\,x_1} = \frac{p}{1-p}\,\frac{u'(x_1)}{u'(x_2)}$$

is the *marginal rate of substitution* between good 1 and good 2.
Therefore, taking the derivative of $MRS^{\bar{u}}_{x_1x_2}$ with respect to x_1, since actually x_2 is a function of x_1 along the indifference curve corresponding to the the utility level $\bar{u}$ (namely, $x_2 = x_2^{\bar{u}}(x_1)$), we get

$$\begin{aligned}
\frac{d}{d\,x_1}MRS^{\bar{u}}_{x_1x_2} &= \frac{p}{1-p}\left[\frac{d}{d\,x_1}\,\frac{u'(x_1)}{u'(x_2^{\bar{u}}(x_1))}\right] \\
&= \frac{p}{1-p}\left[\frac{u''(x_1)u'(x_2) + u'(x_1)u''(x_2)\frac{p}{1-p}\frac{u'(x_1)}{u'(x_2)}}{[u'(x_2)]^2}\right].
\end{aligned}$$

Therefore, in the case of risk aversion, i.e. when $u''(x) < 0$ for every x, we have that $\frac{d}{d\,x_1}MRS^{\bar{u}}_{x_1x_2} < 0$, while in the case of risk neutrality, i.e. when $u''(x) = 0$ for every x, we have that $\frac{d}{d\,x_1}MRS^{\bar{u}}_{x_1x_2} = 0$, implying indifference curves are parallel lines, and the consumption of good 1 is perfect substitute of consumption of good 2.

Example 8.1.1 ($u(x) = \log x$). Assume that $u(x) = \log x$ is the utility function of the agent, so that the indifference curves are given by

$$p \log x_1 + (1-p) \log x_2 = \bar{u} \Rightarrow x_1^p x_2^{1-p} = e^{\bar{u}}.$$

Therefore, we have that

$$x_2 = e^{\frac{\bar{u}}{1-p}} x_1^{\frac{p}{p-1}},$$

$$MRS^{\bar{u}}_{x_1 x_2} = \frac{p}{1-p}\frac{x_2}{x_1} = \frac{p}{1-p}\frac{e^{\frac{\bar{u}}{1-p}} x_1^{\frac{p}{p-1}}}{x_1} = \frac{p}{1-p}\frac{e^{\frac{\bar{u}}{1-p}}}{x_1^{\frac{1}{1-p}}},$$

which is decreasing with respect to x_1.

8.2 Demand for insurance

Consider an agent with wealth w, who faces a probability p of incurring a loss L. She can insure against this loss by buying a policy that will pay out in the event the loss occurs. A policy that will pay a in the event of loss costs qa euros. How much insurance should she buy? This gives rise to an optimization problem under uncertainty, namely

$$\max_a U(a) = \max_a pu(w - qa - L + a) + (1-p)u(w - qa).$$

Denoting the objective function by $U(a)$, the first order condition is:

$$\frac{dU(a)}{da} = pu'(w - qa - L + a)(1-q) - (1-p)u'(w - qa)q = 0.$$

Assuming u is concave (so our consumer is risk-averse), this is necessary and sufficient for a solution. We say that insurance is *actuarily fair* if the expected payout of the insurance company just equals the cost of insurance, that is if $q = p$. We might expect a competitive insurance market to deliver actuarily fair insurance. In this event, the first-order condition simplifies to:

$$u'(w - qa - L + a) = u'(w - qa),$$

meaning that the consumer should fully insure and set $a = L$. In this case, she is completely covered against loss. This result - that a risk-averse consumer will always fully insure if insurance prices are actuarily fair - turns out to be quite important in many contexts.

What happens if the price of insurance is above the actuarily fair price, so $q > p$? Assuming an interior solution, the optimal coverage a^* satisfies the first order condition:

$$\frac{u'(w - qa - L + a)}{u'(w - qa)} = \frac{(1-p)q}{p(1-q)} > 1.$$

Therefore $a^* < L$ - the agent does not fully insure (notice that u' is decreasing). Or, put another way, the agent does not equate wealth in the two states of loss and no loss. Because transferring wealth to the loss state is costly, she makes due with less wealth there and more in the no loss state. We then have the following result.

Proposition 8.2.1. *If $p = q$, the agent will insure fully ($a^* = L$ for all wealth levels). If $p < q$, the agent does not insure fully (i.e., $a^* < L$ for all wealth levels).*

8.3 The Portfolio Problem

Consider an agent with wealth w who has to decide how to invest it. She has two choices: a safe asset that pays a (rate of) return r and a risky asset that pays a random (rate of) return Z, with cumulative distribution function F_Z . She has an increasing and concave utility function u. If she purchases a quantity a of the risky asset and invests the remaining $w-a$ in the safe asset, she will end up with $aZ + (w-a)r$. How should she allocate her wealth? Her optimization problem is:

$$\max_a U(a) = \max_a \mathbb{E}[u(aZ + (w-a)r)] = \max_a \int_0^\infty u(az + (w-a)r)dF_Z(z). \tag{8.3.1}$$

The first order condition for this problem is:

$$U'(a) = \int_0^\infty u'(az + (w-a)r)(z-r)dF_Z(z) = 0. \tag{8.3.2}$$

First, note that if the agent is risk-neutral, so that $u(x) = \alpha x$ for some positive constant α, the above condition becomes

$$\alpha(\mathbb{E}([Z] - r) = 0,$$

so that any choice of a is equally preferred if $\mathbb{E}[Z] = r$. On the other hand, a risk-neutral investor puts all her wealth into the asset with highest expected return, when $\mathbb{E}(Z) \neq r$.

Assuming now that the agent is risk averse, so that $u'' < 0$, it is easily seen that the agent's optimatization problem is strictly concave, and therefore the first order condition characterizes the solution. Indeed, we have that

$$U''(a) = \int_0^\infty u''(az + (w-a)r)(z-r)^2 dF_Z(z) < 0 \quad \text{for every } 0 \le a \le w.$$

This leads to an important observation:
if the risky asset has an expectation greater than r, a risk-averse investor will still invest some amount in it.
To see this, note that at $a = 0$, the derivative of the objective function U is valued

$$U'(0) = u'(wr) \int_0^\infty (z-r) dF_Z(z) = u'(wr)(\mathbb{E}(Z) - r),$$

and clearly $U'(0) > 0$ if $\mathbb{E}(Z) > r$.
So the optimal investment in the risky asset is some amount $a^* > 0$.
This result has a number of consequences. For example, in the insurance problem above the same argument implies that if insurance prices are above their actuarily fair levels, a risk-averse agent will not fully insure. Why? Well, buying full insurance is like putting everything in the safe asset - one gets the same utility regardless of the outcome of uncertainty. Starting from this position, not buying the last dollar of insurance has a positive rate of return, even though it adds risk. By the argument we've just seen, even the most risk-averse guy still wants to take at least a bit of risk at positive returns and hence will not fully insure.
Recall the following definition (already met in Section 3.4).

Definition 8.3.1 (index of absolute risk aversion)**.** Risk-averse utility functions u are characterized by their *(Arrow-Pratt) index of absolute risk aversion*

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

Denote by $\mathcal{C}_u$ the certainty equivalent under expected utility, corresponding to an expected utility representation $\mathbb{E}_u$ of a total preorder $\precsim$ on the set $\mathcal{F}$ of all cumulative distribution functions F on $[0, \infty[$, i.e.

$$\mathcal{C}_u(F) = u^{-1}\left(\mathbb{E}_u^F\right) = u^{-1}\left(\int_0^\infty u(x) dF(x)\right)$$

(see also Definition 2.3.16). Further, define the total preorder $\precsim_u$ on $\mathcal{F}$ by

$$F \precsim_u G \Leftrightarrow \mathcal{C}_u(F) \le \mathcal{C}_u(G) \quad (F,\ G \in \mathcal{F}).$$

As before, $\delta(x)$ stands for the cumulative distribution function corresponding to the degenerate distribution assigning probability 1 to some $x \in [0, \infty[$.

Proposition 8.3.2 (Pratt, 1964). *The following definitions of a utility function u being "more risk averse" than another utility function v are equivalent:*

(i) *For every $F \in \mathcal{F}$, and for every $x \in X$,*

$$\delta(x) \precsim_u F \Rightarrow \delta(x) \precsim_v F.$$

(ii) *For every $F \in \mathcal{F}$, and for every $x \in X$,*

$$\mathcal{C}_u(F) \leq \mathcal{C}_v(F).$$

(iii) *The function u is "more concave" than v, in the sense that there exists some increasing concave function g such that $u = g \circ v$.*

(iv) *For every $x \in X$,*

$$\alpha_u(x) \geq \alpha_v(x).$$

Proof. The equivalence (i) $\Leftrightarrow$ (ii) is clear. Now, because u and v are both increasing, there must be some nondecreasing function g such that $u = g \circ v$. In order to prove the equivalence (ii) $\Leftrightarrow$ (iii), using the definition of $\mathcal{C}_u$ and substituting for $u = g \circ v$, we get

$$\begin{aligned} u(\mathcal{C}_u(F)) \leq u(\mathcal{C}_v(F)) \Leftrightarrow \int_0^\infty g(v(x))dF(x) &\leq g \circ v(v^{-1}(\int_0^\infty v(x)dF(x))) \\ &= g(\int_0^\infty v(x)dF(x)). \end{aligned}$$

Therefore, the second inequality is Jensen's inequality (see Definition 2.3.3), and it holds if and only if g is concave.

In order to prove the equivalence (iii) $\Leftrightarrow$ (iv), consider that, by differentiating twice $u = g \circ v$, with v concave, we get

$$u'(x) = g'(v(x))v'(x), \quad u''(x) = g''(v(x))v'(x)^2 + g'(v(x))v''(x).$$

Dividing the second expression by the first, we obtain

$$\alpha_u(x) = \alpha_v(x) - \frac{g''(v(x))}{g'(v(x))}v'(x).$$

It follows that $\alpha_u(x) \geq \alpha_v(x)$ if and only if $g'' < 0$. □

Proposition 8.3.3. *If u is more risk-averse than v, then, in the portfolio problem (8.3.1), u will optimally invest less in the risky asset than v for any initial level of wealth.*

Proof. A sufficient condition for agent v to invest more in the risky asset than u is (see the first order condition (8.3.2)) is that

$$\int_0^\infty u'(az+(w-a)r)(z-r)dF_Z(z) = 0 \Rightarrow \int_0^\infty v'(az+(w-a)r)(z-r)dF_Z(z) \geq 0.$$

□

Now, if u is more risk-averse than v, then $v = h \circ u$ for some increasing convex function h. Given this, the second inequality above is satisfied, which can be re-written as

$$\int_0^\infty h'(u(az+(w-a)r))u'(az+(w-a)r)(z-r)dF_Z(z) \geq 0$$

Indeed, the first term $h'(\cdot)$ is positive and increasing in z, while the second set of terms is negative when $z < r$ and positive when $z > r$. So multiplying by $h'(\cdot)$ puts more weight on the positive terms and less on the negative term. Therefore the product of the two terms integrates up to more than the second term alone, or in other words, to more than zero. □

CHAPTER 9

Portfolio Theory

9.1 Notation

Consider n assets (investments), whose respective stochastic returns R_h ($h = 1, ..., n$) have mean values

$$\mu_h = \mathbb{E}[R_h],$$

and covariance matrix

$$V = (v_{hk})_{h,k} = (cov(R_h, R_k))_{h,k}.$$

Definition 9.1.1. A *portfolio* is a point $\underline{x}' = (x_1, ..., x_n) \in [0, 1]^n$, such that $x_h \geq 0$ for every $h \in \{1, ..., n\}$, and $\sum_{h=1}^{n} x_h = 1$.

Notice that the mean value

$$\mathbb{E}[R] = \mathbb{E}[\sum_{h=1}^{n} x_h R_h]$$

and the variance

$$Var[R] = Var[\sum_{h=1}^{n} x_h R_h]$$

of the *total investment*

$$R = \sum_{h=1}^{n} x_h R_h$$

associated to a portfolio $\underline{x}'$ are defined to be

$$\mu = \mathbb{E}[\sum_{h=1}^{n} x_h R_h] = \underline{\mu}'\underline{x},$$

$$\sigma^2 = Var[\sum_{h=1}^{n} x_h R_h] = \sum_{k=1}^{n}\sum_{h=1}^{n} x_h x_k cov(R_h, R_k) = \underline{x}'V\underline{x}.$$

Define the *unit vector* $\underline{e}' = (1, ...1) \in \mathbb{R}$.

9.2 Multi-objective optimization and portfolio selection

The *Multi-objective optimization problem* (MOP) is usually formulated by means of the standard notation[(1)]

$$\max_{x \in X}[u_1(x), ..., u_m(x)] = \max_{x \in X} \mathbf{u}(x), \quad m \geq 2, \tag{9.2.1}$$

where X is the *choice set* (or the *design space*), u_i is the *decision function* (in this case a *utility function*) associated with the i-th individual (or criterion), and $\mathbf{u} : X \mapsto \mathbb{R}^m$ is the vector-valued function defined by $\mathbf{u}(x) = (u_1(x), ..., u_m(x))$ for all $x \in X$.

(1)Needless to say, this formulation of the multi-objective optimization problem is equivalent, "mutatis mutandis", to $\min_{x \in X}[f_1(x), ..., f_m(x)] = \min_{x \in X} \mathbf{f}(x),\ m \geq 2$. We use the approach with the maximum for the sake of convenience.

Definition 9.2.1 (Pareto optimal solution). An element $x_0 \in X$ is a said to be

1. a *Pareto optimal* solution to problem (9.2.1) if one of the following two equivalent conditions is verified:

   (a) for no $x \in X$ it happens that $u_i(x_0) \leq u_i(x)$ for every $i \in \{1, ..., m\}$, and there exists an index $\bar{i} \in \{1, ..., m\}$ such that $u_{\bar{i}}(x_0) < u_{\bar{i}}(x)$;

   (b) For every $x \in X$, if $u_i(x_0) \leq u_i(x)$ for every $i \in \{1, ..., m\}$, then $u_i(x_0) = u_i(x)$ for every index i;

2. a *weak Pareto optimal* solution to problem (9.2.1) if one of the following two equivalent conditions is verified:

   (a) for no $x \in X$ it happens that $u_i(x_0) < u_i(x)$ for every $i \in \{1, ..., m\}$;

   (b) For every $x \in X$, if $u_i(x_0) \leq u_i(x)$ for every $i \in \{1, ..., m\}$, then $u_{\bar{i}}(x_0) = u_{\bar{i}}(x)$ for at least one index $\bar{i} \in \{1, ..., m\}$.

The point $x_0 \in X$ is said to be (*weakly*) *Pareto optimal* or a (*weakly*) *efficient point* for problem (9.2.1).

The *Markowitz portfolio selection problem* appears in the following form when expressed in terms of a multi-objective optimization problem:

$$\max_{x \in X} [\mu' x, -x' V x], \tag{9.2.2}$$

where $m = 2$ is the number of criteria, $X = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$ is the set of all portfolios, $n > 0$ is the number of risky assets considered, $\underline{\mu} \in \mathbb{R}^n$ is a vector of expected returns, and V is a covariance matrix of the returns. The existence of a solution to Markowitz portfolio selection problem (9.2.2) is guaranteed by the following general result.
We must present two definitions and a proposition.

Definition 9.2.2 (compact topology)**.** A topological space (X, τ) is said to be *compact* if every open cover of X admits a finite subcover (i.e., for every family of open sets $\{O_\alpha\}_{\alpha \in I}$ such that $\bigcup_{\alpha \in I} O_\alpha \supset X$ there exists a finite subfamily $\{O_1, ..., O_n\}$ such that $\bigcup_{i=1}^{n} O_i \supset X$).

Remark 9.2.3. A famous result guarantees that if (X, τ) is a metric space, then a subset A of X is compact if and only if A is closed and bounded.

Definition 9.2.4 (upper semicontinuous real-valued function)**.** A real-valued function u on an arbitrary topological space (X, τ) is said to be *upper semicontinuous* if

$$u^{-1}(]-\infty, \alpha[) = \{x \in X : u(x) < \alpha\}$$

is an open set for every $\alpha \in \mathbb{R}$.

Proposition 9.2.5. *Every upper semicontinuous real-valued function u on a compact topological space (X, τ) attains its maximum.*

Proof. Let u be any upper semicontinuous real-valued function on a compact topological space (X, τ). Assume that u has no maximum (i.e., for every $x \in X$ there exists $x' \in X$ such that $u(x) < u(x')$). Then we have that

$$\{u^{-1}(]-\infty, u(x)[)\}_{x \in X}$$

is an open cover of X. Since (X, τ) is compact, there exists a finite set $\{x_1, ..., x_n\} \subset X$ such that

$$\{u^{-1}(]-\infty, u(x_i)[)\}_{i \in \{1,...,n\}}$$

is still an open cover of X. Assume, without loss of generality, that $u(x_1) < u(x_2) < ... < u(x_n)$, so that $u^{-1}(]-\infty, u(x_1)[) \subsetneq u^{-1}(]-\infty, u(x_2)[) \subsetneq ... \subsetneq u^{-1}(]-\infty, u(x_n)[)$. Consider that $x_n \notin u^{-1}(]-\infty, u(x_n)[)$. However, there exists $i < n$ such that $x_n \in u^{-1}(]-\infty, u(x_i)[)$, implying that

$x_n \in u^{-1}(]-\infty, u(x_n)[)$, due to the fact that $u^{-1}(]-\infty, u(x_i)[) \subsetneq u^{-1}(]-\infty, u(x_n)[)$. This contradiction completes the proof. □

Let us now go back to the general multi-objective optimization problem (9.2.1) in order to state the following fundamental result.

Theorem 9.2.6. *Consider the general multi-objective optimization problem*

$$\max_{x \in X}[u_1(x), ..., u_m(x)] = \max_{x \in X} \mathbf{u}(x), \quad m \geq 2.$$

If τ is a topology on X, (X, τ) is a compact topological space and the functions $u_1, ...u_m$ are upper semicontinuous, then there exists a Pareto optimal solution $x_0 \in X$.

Proof. Since the functions $u_1, ...u_m$ are upper semicontinuous, it is easy to show that also the function $u = \sum_{i=1}^{m} u_i$ is upper semicontinuous. Then u attains its maximum on X by Proposition 9.2.5. Let $x_0 \in X$ be the point at which u attains its maximum. Then x_0 is also a Pareto optimal solution to the above problem. Otherwise, there exists $x \in X$ such that $u_i(x_0) \leq u_i(x)$ for every $i \in \{1, ..., m\}$, and there exists an index $\bar{i} \in \{1, ..., m\}$ such that $u_{\bar{i}}(x_0) < u_{\bar{i}}(x)$, so that $u(x_0) < u(x)$. This consideration completes the proof. □

9.3 Markowitz portfolio selection

The case of risky assets

The key assumptions of Markowitz portfolio selection are the following:

1. The investors are rational and risk averse.

2. The investors have access to the same information.

3. The investors base their decisions on expected return and variance.

4. The market is frictionless, i.e., there are no taxes or transaction costs.

5. All expected returns, variances and covariances of returns are known.

6. We are not considering risk free assets.

Theorem 9.3.1 (Markowitz). *Consider the mean-variance portfolio selection problem*

$$min\ \sigma^2 = \underline{x}'V\underline{x}$$

$$sub$$

$$\begin{cases} \underline{\mu}'\underline{x} = \mu_* \\ \underline{x}'\underline{e} = 1 \end{cases} \tag{9.3.1}$$

If the covariance matrix V is nonsingular, and the expected returns are nonidentical (i.e., $\mu_h \neq \mu_k$ for some $h, k \in \{1, ...n\}$), then the optimal portfolio $\underline{x}$ is defined to be, for every fixed global mean return μ_*,

$$\underline{x} = V^{-1} \begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix} A^{-1} \begin{bmatrix} \mu_* \\ 1 \end{bmatrix}, \tag{9.3.2}$$

where A is the symmetric matrix of order 2, defined as

$$A = \begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix}' V^{-1} \begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} = \begin{bmatrix} \underline{\mu}'V^{-1}\underline{\mu} & \underline{\mu}'V^{-1}\underline{e} \\ \underline{e}'V^{-1}\underline{\mu} & \underline{e}'V^{-1}\underline{e} \end{bmatrix}. \tag{9.3.3}$$

Further, the variance associated to the optimal portfolio (9.3.2) is

$$\sigma_*^2 = \underline{x}'V\underline{x} = \begin{bmatrix} \mu_* & 1 \end{bmatrix} A^{-1} \begin{bmatrix} \mu_* \\ 1 \end{bmatrix} = \frac{\gamma\mu_*^2 - 2\beta\mu_* + \alpha}{\alpha\gamma - \beta^2}. \tag{9.3.4}$$

Proof. Consider the Lagrangean function $L(\underline{x}, \lambda_1, \lambda_2)$ associated to problem (9.3.1), namely

$$L(\underline{x}, \lambda_1, \lambda_2) = \underline{x}' V \underline{x} - \lambda_1 (\underline{x}' \underline{\mu} - \mu_*) - \lambda_2 (\underline{x}' \underline{e} - 1).$$

The first order conditions $\frac{\partial L(\underline{x},\lambda_1,\lambda_2)}{\partial \underline{x}} = \underline{0}$, $\frac{\partial L(\underline{x},\lambda_1,\lambda_2)}{\partial \lambda_1} = 0$, $\frac{\partial L(\underline{x},\lambda_1,\lambda_2)}{\partial \lambda_2} = 0$ can be written as follows:

$$\begin{cases} \frac{\partial L(\underline{x},\lambda_1,\lambda_2)}{\partial \underline{x}} = 2V\underline{x} - \lambda_1 \underline{\mu} - \lambda_2 \underline{e} = \underline{0} \\ \frac{\partial L(\underline{x},\lambda_1,\lambda_2)}{\partial \lambda_1} = \underline{\mu}' \underline{x} - \mu_* = 0 \\ \frac{\partial L(\underline{x},\lambda_1,\lambda_2)}{\partial \lambda_2} = \underline{x}' \underline{e} - 1 = 0 \end{cases} . \tag{9.3.5}$$

From the first condition, we get

$$\underline{x} = \frac{1}{2} V^{-1} (\lambda_1 \underline{\mu} + \lambda_2 \underline{e}) = \frac{1}{2} V^{-1} \begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} .$$

The last two conditions can be written in compact form as follows:

$$\begin{bmatrix} \mu_* \\ 1 \end{bmatrix} = \begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix}' \underline{x},$$

so that

$$\begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix}' \underline{x} = \frac{1}{2} \begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix}' V^{-1} (\lambda_1 \underline{\mu} + \lambda_2 \underline{e}) = \frac{1}{2} \begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix}' V^{-1} \begin{bmatrix} \underline{\mu} & \underline{e} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \mu_* \\ 1 \end{bmatrix} . \tag{9.3.6}$$

Define the matrix A as in (9.3.3), and consider that, under our assumption, A is invertible. Then condition (9.3.6) can be written as

$$\frac{1}{2} A \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \mu_* \\ 1 \end{bmatrix},$$

and, in turn, as

$$\frac{1}{2}\begin{bmatrix}\lambda_1\\ \lambda_2\end{bmatrix} = A^{-1}\begin{bmatrix}\mu_*\\ 1\end{bmatrix}.$$

Finally, the portfolio which minimizes the variance for a given mean return μ_*, is precisely defined to be

$$\underline{x} = V^{-1}\begin{bmatrix}\underline{\mu} & \underline{e}\end{bmatrix} A^{-1}\begin{bmatrix}\mu_*\\ 1\end{bmatrix}.$$

The minimum variance is

$$\sigma_*^2 = \underline{x}'V\underline{x} = \begin{bmatrix}\mu_* & 1\end{bmatrix} A^{-1}\begin{bmatrix}\underline{\mu} & \underline{e}\end{bmatrix}' V^{-1}VV^{-1}\begin{bmatrix}\underline{\mu} & \underline{e}\end{bmatrix} A^{-1}\begin{bmatrix}\mu_*\\ 1\end{bmatrix}$$

$$= \begin{bmatrix}\mu_* & 1\end{bmatrix} A^{-1}\begin{bmatrix}\underline{\mu} & \underline{e}\end{bmatrix}' V^{-1}\begin{bmatrix}\underline{\mu} & \underline{e}\end{bmatrix} A^{-1}\begin{bmatrix}\mu_*\\ 1\end{bmatrix} = \begin{bmatrix}\mu_* & 1\end{bmatrix} A^{-1}AA^{-1}\begin{bmatrix}\mu_*\\ 1\end{bmatrix}$$

$$= \begin{bmatrix}\mu_* & 1\end{bmatrix} A^{-1}\begin{bmatrix}\mu_*\\ 1\end{bmatrix}.$$

□

Proposition 9.3.2 (Representation in the plane of the pairs (σ_*^2, μ_*) and (σ_*, μ_*)). *Notice that equation ((9.3.4) represents*

1. *a parabola with its axis parallel to the σ^2-axis in the plane (σ^2, μ). In this case the vertex is*

$$\left(\frac{1}{\gamma}, \frac{\beta}{\gamma}\right);$$

2. *an hyperbola in the plane (σ, μ). In this case the vertex is*

$$\left(\frac{1}{\sqrt{\gamma}}, \frac{\beta}{\gamma}\right),$$

and the asymptotes are

$$\mu = \frac{\beta}{\gamma} \pm \left[\frac{\alpha\gamma - \beta^2}{\gamma}\right]^{\frac{1}{2}}.$$

The corresponding curve is said to be the frontier of the admissible portfolio, *and its upper part is the* efficient frontier.

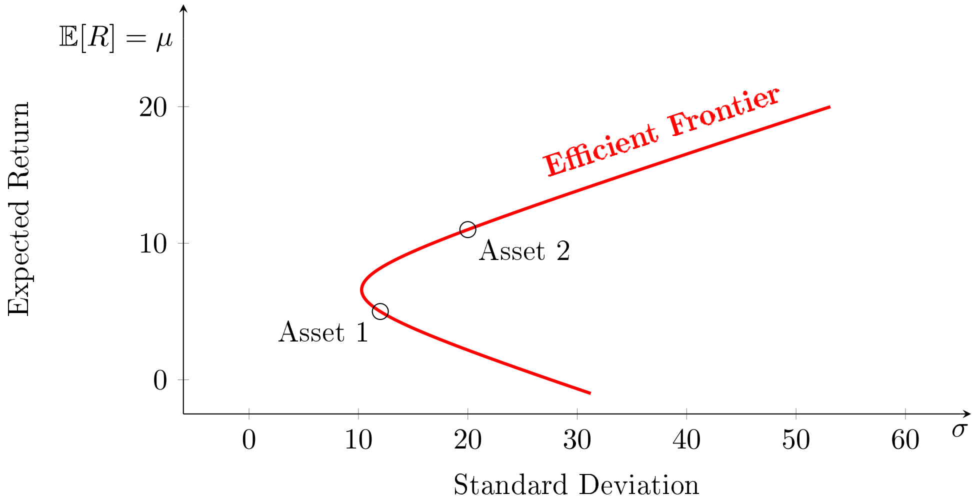
Definition 9.3.3. We say that *a portfolio $\underline{x}$ mean-variance dominates a portfolio $\underline{y}$* if either

$$\mu_{\underline{x}} > \mu_{\underline{y}} \text{ and } \sigma_{\underline{x}} \leq \sigma_{\underline{y}}$$

or

$$\mu_{\underline{x}} \geq \mu_{\underline{y}} \text{ and } \sigma_{\underline{x}} < \sigma_{\underline{y}}.$$

Definition 9.3.4. A *frontier portfolio* is one which displays minimum variance among all feasible portfolios with the same expected return. The *efficient frontier* for a given expected return is the set of all undominated portfolios.



Remark 9.3.5. The more common and better justification for the mean-variance approach is to assume that assets returns are multivariate normally distributed. Since the multivariate normal ditribution is completely characterized by its mean and variance, the third and higher moments can all be expressed in terms of mean and variance. Furthermore, any linear combination of normal random variables (a portfolio) is again a normal random variable. More is true, indeed one can show that only mean-variance efficient portfolios can maximize expected utility if returns are multivariate normal.

The case of two risky assets

Consider the particular case of 2 assets (investments), whose respective stochastic returns R_h $(h = 1, 2)$ have mean values

$$\mu_h = \mathbb{E}[R_h],$$

and covariance matrix The covariance matrix is

$$V = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{pmatrix}.$$

We have that $\sigma_{1,2} = cov(R_1, R_2) = \rho_{1.2}\sigma_1\sigma_2$, where $\rho_{1,2}$ is the linear correlation coefficient $(| \rho_{1,2} | \leq 1)$ between R_1 and R_2.

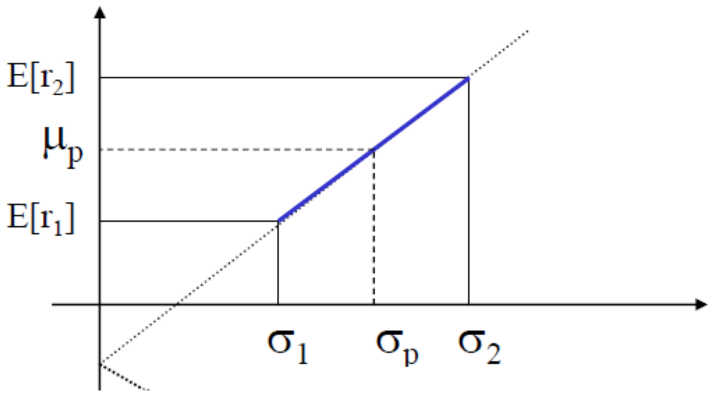


Figure 9.1 Efficient frontier of two perfectly correlated risky assets

If $\rho_{1,2} = 1$, we deal with two perfectly correlated risky assets, and

$$\sigma = \sqrt{Var(x_1 R_1 + (1 - x_1) R_2)} = | x_1 \sigma_1 + (1 - x_1) \sigma_2 |,$$

$$x_1 = \frac{\pm\sigma - \sigma_2}{\sigma_1 - \sigma_2},$$

$$\mu = x_1 \mu_1 + (1 - x_1)\mu_2 = \mu_1 + \frac{\mu_2 - \mu_1}{\sigma_2 - \sigma_1}(\pm\sigma - \sigma_1).$$

If $\rho_{1,2} = -1$, we deal with two perfectly negative correlated risky assets, and

$$\sigma = \sqrt{Var(x_1 R_1 + (1 - x_1) R_2)} = | x_1 \sigma_1 - (1 - x_1) \sigma_2 |,$$

$$x_1 = \frac{\pm\sigma + \sigma_2}{\sigma_1 + \sigma_2},$$

$$\mu = x_1 \mu_1 + (1 - x_1)\mu_2 = \frac{\sigma_2}{\sigma_1 + \sigma_2}\mu_1 + \frac{\sigma_2}{\sigma_1 + \sigma_2} \pm \frac{\mu_2 - \mu_1}{\sigma_1 + \sigma_2}\sigma.$$

The case of two imperfectly correlated risky assets, i.e. when $-1 < \rho_{1,2} < 1$, is represented in Figure 9.3.

The case of one risky and one riskless asset is represented in Figure 9.4.

The case of one riskless asset

Consider now the case when there is one riskless asset, with a rate of return equal to r, and n risky assets, whose respective stochastic returns R_h ($h =$

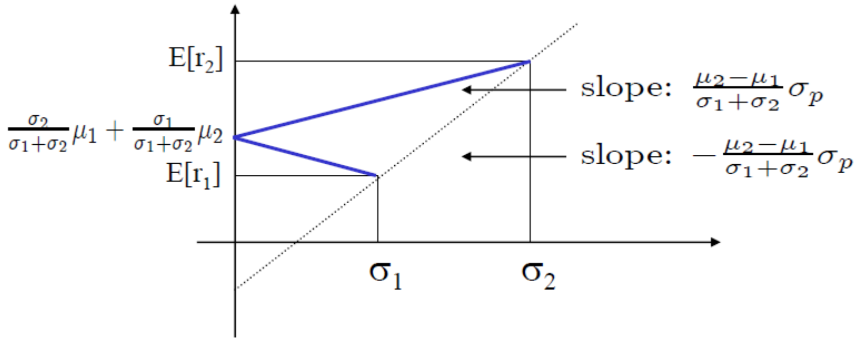


Figure 9.2 Efficient frontier of two negative correlated risky assets

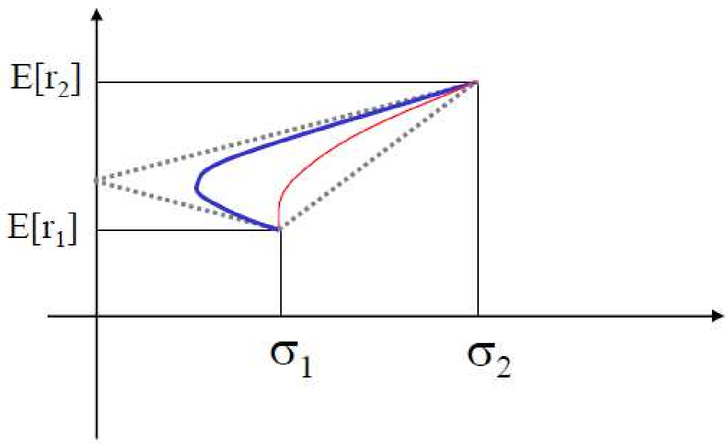


Figure 9.3 Efficient frontier of two imperfectly correlated risky assets

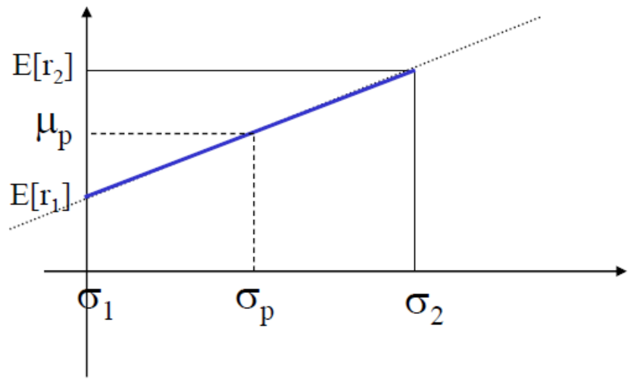


Figure 9.4 Efficient frontier of one risky and one riskless asset

$1, ..., n)$ have mean values

$$\mu_h = \mathbb{E}[R_h],$$

and covariance matrix

$$V = (v_{hk})_{h,k} = (cov(R_h, R_k))_{h,k}.$$

Define, for every $h \in \{1, ..., n\}$, the *expected excess of return of asset* h,

$$z_h = \mu_h - r,$$

and consider the n-components vector

$$\underline{z} = [z_h].$$

A portfolio is now a a point $\underline{w}' = (w_1, ..., w_n, w_{n+1}) \in [0,1]^{n+1}$, such that $w_h \geq 0$ for every $h \in \{1, ..., n+1\}$, and $\sum_{h=1}^{n+1} w_h = 1$, with w_{n+1} the weight associated to the riskless asset. Define $\underline{w}'_n = (w_1, ..., w_n) \in [0,1]^n$. Then the *global expected excess of return* is defined to be

$$\mathbf{z} = \underline{w}'_n \underline{z},$$

while the variance of the portfolio is still

$$\sigma^2 = \underline{w}_n{}' V \underline{w}_n,$$

since the return of the last asset is certain. Therefore the Markowitz mean-variance portfolio selection problem can be now written as

$$\begin{aligned} &\min\ \sigma^2 = \underline{w}_n{}' V \underline{w}_n \\ &\text{sub } \underline{w}_n{}' \underline{z} = \mathbf{z}_*, \end{aligned} \tag{9.3.7}$$

where the constraint $\sum_{h=1}^{n} w_h$ does not appear, since there is no obligation to invest the total wealth in the first n (risky) assets. Using the same procedure as before, the validity of the following theorem can be established.

Theorem 9.3.6. *The optimal solution to problem (9.3.7) is provided by the n-th vector*

$$\underline{w}_n = \left(\frac{\mathbf{z}_*}{\underline{z}' V^{-1} \underline{z}} \right) V^{-1} \underline{z}, \tag{9.3.8}$$

so that the minimum variance of a portfolio is

$$\sigma_*^2 = \underline{w}_n{}' V \underline{w}_n = \left(\frac{\mathbf{z}_*}{\underline{z}' V^{-1} \underline{z}} \right)^2 \underline{z}' V^{-1} V V^{-1} \underline{z} = \frac{\mathbf{z}_*^2}{\underline{z}' V^{-1} \underline{z}}, \tag{9.3.9}$$

Proof. Consider the Lagrangean function $L(\underline{w}_n, \lambda)$ associated to problem (9.3.7), namely

$$L(\underline{w}_n, \lambda) = \underline{w}_n{}' V \underline{w}_n - \lambda(\underline{w}_n{}' \underline{z} - \mathbf{z}_*).$$

The first order conditions $\frac{\partial L(\underline{w}_n, \lambda)}{\partial \underline{w}_n} = \underline{0}$, $\frac{\partial L(\underline{w}_n, \lambda)}{\partial \lambda} = 0$ can be written as follows:

$$\begin{cases} \frac{\partial L(\underline{w}_n, \lambda)}{\partial \underline{w}_n} = 2V \underline{w}_n - \lambda \underline{z} = \underline{0} \\ \frac{\partial L(\underline{w}_n, \lambda)}{\partial \lambda} = \underline{z}' \underline{w}_n - z_* = 0. \end{cases} . \tag{9.3.10}$$

From the first condition, we get

$$\underline{w}_n = \frac{1}{2}\lambda V^{-1}\underline{z}.$$

Therefore, using the second condition above,

$$\underline{z}'\underline{w}_n = z_* = \frac{1}{2}\lambda \underline{z}' V^{-1}\underline{z},$$

and, solving for λ, we arrive at

$$\lambda = \frac{2\underline{z}'\underline{w}_n}{\underline{z}'V^{-1}\underline{z}} = \frac{2z_*}{\underline{z}'V^{-1}\underline{z}},$$

and we finally obtain

$$\underline{w}_n = \left(\frac{\mathbf{z}_*}{\underline{z}'V^{-1}\underline{z}}\right) V^{-1}\underline{z}.$$

□

By equation (9.3.9), the corresponding expression of the minimum standard deviation portfolio is the following one:

$$\sigma_* = \frac{\mid \mu_* - r \mid}{(\underline{z}'V^{-1}\underline{z})^{\frac{1}{2}}}. \qquad (9.3.11)$$

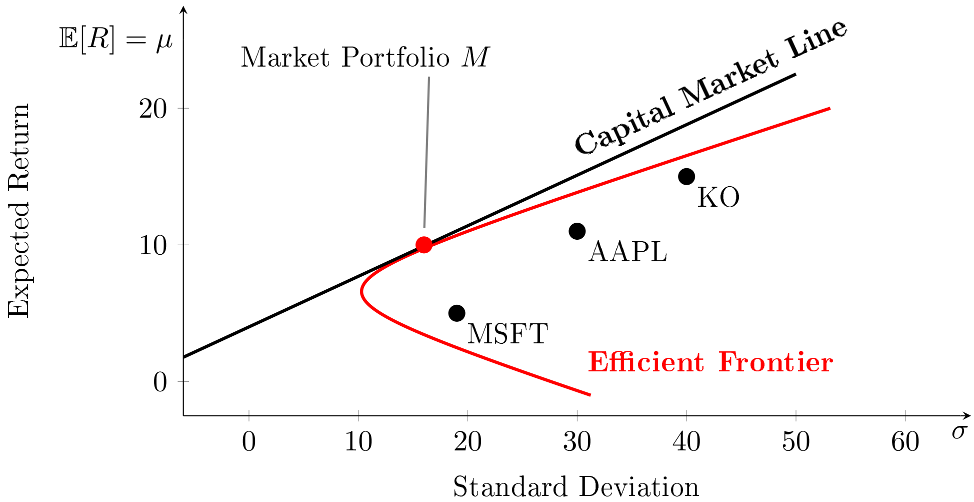
Expression (9.3.11) provides the equation of the efficient frontier, as an half-line with positive slope and with origin $(0, r)$. Such an half-line is said to be the *capital market line*.

By Theorem 9.3.1, there is a unique optimal portfolio which only consists of risky assets (i.e., when $w_{n+1} = 0$), and it is said to be the *market portfolio* M. Indeed, it could be shown that the capital market line is tangent to the efficient frontier relative to the n risky assets at the point M.

If we set $F = (0, r)$, we have that every optimal portfolio is of the form

$$\xi M + (1 - \xi)F,$$

with $\xi \geq 0$. in case that $0 \leq \xi \leq 1$, then clearly the portfolio is located on the segment with extremes F and M. If it happens that $\xi > 1$, the corresponding portfolio is obtained by getting into a debt at the riskless rate r, and then by investing the resulting sum in the market portfolio.



CHAPTER 10

Risk Sharing

10.1 Optimal Risk Allocations

Consider an uncertain payoff X and m agents endowed with their own initial exposures $(X_1, .., X_m)$, with $X = \sum_{i=1}^{m} X_i$. Agent i has preferences over her own risks which are expressed by a (not necessarily total) preorder $\precsim_i$ $(i = 1, ..., m)$.
Divide X into uncertain shares $Y_1, ..., Y_m$ in such a way that $X = \sum_{i=1}^{m} Y_i$, be the total exposure.

Definition 10.1.1 (feasible allocations)**.** *For every risk X, denote by $\mathcal{A}(X)$ the set of all the* feasible allocations *of X, i.e. the set*

$$\mathcal{A}(X) = \{(Y_1, ..., Y_m) \mid X = \sum_{i=1}^{m} Y_i\}. \tag{10.1.1}$$

We now present the central concept of *Pareto optimal allocation.*

Definition 10.1.2 (Pareto optimal allocation)**.** *An allocation $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ is said to be* Pareto optimal *if for no other allocation $(Y_1, ..., Y_m) \in \mathcal{A}(X)$ it occurs that $Y_1^* \precsim_1 Y_1, ..., Y_m^* \precsim_m Y_m$ with at least one index i such that $Y_i^* \prec_i Y_i$.*

We omit the immediate proof of the following proposition.

Proposition 10.1.3. *Assume that the individual preorder $\precsim_i$ is total for every $i \in \{1, ..., m\}$. Then the following conditions are equivalent concerning an allocation $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$:*

1. $(Y_1^*, ..., Y_m^*)$ *is Pareto optimal;*

2. *for every allocation $(Y_1, ..., Y_m) \in \mathcal{A}(X)$ such that $Y_i^* \precsim_i Y_i$ for $i \in \{1, ..., m\}$ it occurs that $Y_i^* \sim_i Y_i$ for $i \in \{1, ..., m\}$.*

The following definition is of basic importance.

Definition 10.1.4. *An allocation $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ is said to be* individually rational *if all agents are at least as well off under $(Y_1^*, ..., Y_m^*)$ as under the initial exposures X_i ($i \in \{1, ..., m\}$), so that $X_i \precsim_i Y_i^*$ for all $i \in \{1, ..., m\}$.*

Definition 10.1.5. *An allocation $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ is said to be* optimal *if it is both* Pareto optimal *and* individually rational.

The risk sharing problem basically consists in finding a Pareto optimal or in particular an optimal allocation $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$.
Let try now to qualitative understand individual preference decision making behaviour on feasible allocations. We know that each individual has preferences over her own risks expressed by a preorder; so if we consider two feasible allocations $Y = (Y_1, ..., Y_m)$ and $Z = (Z_1, ..., Z_m)$, that is $X = \sum_{i=1}^m Y_i = \sum_{i=1}^m Z_i$, agent i can only have preferences on her possible shares Y_i and Z_i. So, we can only say that $Z_i \precsim_i Y_i$ or $Y_i \precsim_i Z_i$ or else Y_i and Z_i are incomparable. This individual choice over single shares, implies an individual(selfish) preference over the entire allocations. Indeed, we can say that agent i prefers the allocation $Y = (Y_1, ..., Y_m)$ over $Z = (Z_1, ..., Z_m)$ if and only if she prefers her own share Y_i to Z_i i.e.

$$(Z_1, ..., Z_m) \precsim_i (Y_1, ..., Y_m) \;\Leftrightarrow\; Z_i \precsim_i Y_i \quad (i = 1, ..., m).$$

Therefore, in the sequel we shall assume that every individual preference $\precsim_i$ is actually defined on $\mathcal{A}(X)$.
Let us now introduce the definition of a *translation invariant* preorder.

Definition 10.1.6. *A preorder* $\precsim$ *on a vector space* $\mathcal{L}_+$ *of nonnegative random variables is said to be* translation invariant *if the following condition holds for every positive real number c (identified with the constant random variable equal to c), and all random variables* $X, Y \in \mathcal{L}_+$,

$$X \precsim Y \Leftrightarrow X + c \precsim Y + c. \tag{10.1.2}$$

Remark 10.1.7. It is easy to check that a preorder $\precsim$ is translation invariant if and only if actually the above condition (10.1.2) holds true for every constant random variable c.

Remark 10.1.8. It should be noted that, if $\precsim$ is a translation invariant total preorder on $\mathcal{L}$, then for all random variables X, Y, and every real number c,

$$X \prec Y \Leftrightarrow X + c \prec Y + c. \tag{10.1.3}$$

Indeed, in this case we have that $\neg(X + c \prec Y + c) \Leftrightarrow \neg((X + c \precsim Y + c)$ *and* $\neg(Y + c \precsim X + c)) \Leftrightarrow Y + c \precsim X + c \Leftrightarrow Y \precsim X \Leftrightarrow \neg(X \prec Y)$.

It is clear that a total preorder $\precsim$ on $\mathcal{L}$ is translation invariant provided that it admits a *translation invariant* utility function U (i.e., $U(X+c) = U(X)+c$ for all $X \in \mathcal{L}_+$ and $c \in \mathbb{R}$).
In the following proposition we present a simple but useful property exhibited by Pareto optimal allocations in case of translation invariant individual total preorders.

Proposition 10.1.9. *Assume that* $\precsim_i$ *is a translation invariant total preorder for all* $i \in \{1, ..., m\}$ *and consider any m-tuple of real numbers* $(\pi_i, ..., \pi_m)$ *such that* $\sum_{i=1}^{m} \pi_i = 0$. *Then the following conditions are equivalent.*

1. $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ *is Pareto optimal;*

2. $(Y_1^* + \pi_1, ..., Y_m^* + \pi_m) \in \mathcal{A}(X)$ *is Pareto optimal.*

Proof. (1) $\Rightarrow$ (2). By contraposition, consider $(Y_1^*+\pi_1, ..., Y_m^*+\pi_m) \in \mathcal{A}(X)$ which is not Pareto optimal. Then there exists $(Z_1^*+\pi_1, ..., Z_m^*+\pi_m) \in \mathcal{A}(X)$

such that $Y_i^* + \pi_i \precsim_i Z_i^* + \pi_i$ for all $i \in \{1, ..., m\}$ with at least one index i such that $Y_i^* + \pi_i \prec_i Z_i^* + \pi_i$. By translation invariance of the total preorder $\precsim_i$ we have that $Y_i^* \precsim_i Z_i^*$ for all $i \in \{1, ..., m\}$ with at least one index i such that $Y_i^* \prec_i Z_i^*$. Hence $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ is not Pareto optimal.
(2) $\Rightarrow$ (1). Analogous. □

Corollary 10.1.10. *Assume that $\precsim_i$ is a translation invariant total preorder for all $i \in \{1, ..., m\}$ and let $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ be a Pareto optimal allocation. Then $(Y_1^* + Z_1, ..., Y_m^* + Z_m) \in \mathcal{A}(X)$ is also a Pareto optimal allocation provided that the following condition holds for some uniquely determined m-tuple of real numbers $(\pi_i, ..., \pi_m)$ such that*

$$Z_i \sim_i \pi_i \text{ for all } i \in \{1, ..., m\} \text{ and } \sum_{i=1}^{m} \pi_i = 0. \tag{10.1.4}$$

Proof. We have that $Y_i^* + Z_i \sim_i Y_i^* + \pi_i$, implying that also $(Y_1^* + Z_1, ..., Y_m^* + Z_m) \in \mathcal{A}(X)$ is Pareto Optimal from the above Proposition 10.1.9. □

The existence of optimal allocations is guaranteed when there are Pareto optimal allocations and $\precsim_i$ has a translation invariant utility U_i for all $i \in \{1, ..., m\}$. This fact is illustrated in the following easy proposition.

Proposition 10.1.11. *Assume that $\precsim_i$ is a translation invariant total preorder for all $i \in \{1, ..., m\}$ with a translation invariant utility function U_i. Let $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ be Pareto optimal. Then the following conditions are equivalent for every m-tuple of real numbers $(\pi_i, ..., \pi_m)$ such that $\sum_{i=1}^m \pi_i = 0$:*

1. $(Y_1^* + \pi_1, ..., Y_m^* + \pi_m) \in \mathcal{A}(X)$ *is optimal;*
2. $U_i(X_i) - U_i(Y_i^*) \leq \pi_i$.

Proof. Just consider that, under our assumptions, $(Y_1^* + \pi_1, ..., Y_m^* + \pi_m) \in \mathcal{A}(X)$ is optimal if and only if, for all $i \in \{1, ..., m\}$, $U_i(X_i) \leq U_i(Y_i^*) + \pi_i = U_i(Y_i^* + \pi_i)$. □

Remark 10.1.12. In the case of individual total preorders with translation invariant utilities, the above Proposition 10.1.11 guarantees that determining Pareto optimal allocations is in fact equivalent to determining optimal allocations for every choice of the initial exposures.

We are now interested in considering a "coalition" preference decision making behaviour, expressed by the *coalition preorder* $\precsim$ on $\mathcal{A}(X)$.
In particular, we say that a coalition of m agents prefers the allocation $Y = (Y_1, ..., Y_m)$ over $Z = (Z_1, ..., Z_m)$ if and only if every agents prefers Y over Z, i.e.

$$(Z_1, ..., Z_m) \precsim (Y_1, ..., Y_m) \;\Leftrightarrow\; (Z_1, ..., Z_m) \precsim_i (Y_1, ..., Y_m) \;\forall i \in \{1, .., n\}. \tag{10.1.5}$$

Observe that this is equivalent to defining the *coalition preorder* $\precsim$ as the intersection of the individual total preorders. Indeed, let us consider the following definition.

Definition 10.1.13 (coalition preorder)**.** *The* coalition preorder $\precsim$ *on the set* $\mathcal{A}(X)$ *of all feasible allocations is defined to be*

$$\precsim = \bigcap_{i=1}^{m} \precsim_i . \tag{10.1.6}$$

We have previously seen that $(Z_1, ..., Z_m) \precsim_i (Y_1, ..., Y_m) \;\Leftrightarrow\; Z_i \precsim_i Y_i$, so we can write in more explicit terms:

$$(Z_1, ..., Z_m) \precsim (Y_1, ..., Y_m) \;\Leftrightarrow\; Z_i \precsim_i Y_i \;\forall i = (1, .., m).$$

Remark 10.1.14. Observe that the preorder $\precsim$ is not necessarily total, even if $\precsim_i$ is total for every i. Indeed, for two feasible allocations $Y = (Y_1, ..., Y_m)$ and $Z = (Z_1, ..., Z_m)$ there may exist two indexes i, j with $Y_i \prec_i Z_i$ and $Z_j \prec_j Y_j$. This consideration justifies in full the material and technique presented in the previous chapter and in particular the considerations on the existence of maximal elements for not necessarily total preorders.

Remark 10.1.15. It should be noted that in the particular case when every individual preorder $\precsim_i$ is total and admits a utility representation U_i, for all $Y = (Y_1, ..., Y_m)$ and $Z = (Z_1, ..., Z_m)$ in $\mathcal{A}(X)$ it occurs that $(Y_1, ..., Y_m) \precsim (Z_1, ..., Z_m)$ if and only if $U_i(Y_i) = U_i(Y) \leq U_i(Z) \leq U_i(Z_i)$. Therefore, we have that $\mathcal{U} = \{U_1, ..., U_m\}$ is a finite multi-utility representation of the coalition preorder $\precsim$.

The problem concerning the existence of Pareto optimal allocations can be related to the problem concerning the existence of maximal elements for the coalition preorder $\precsim$ defined in (10.1.6). The following proposition illustrates this possibility.

Proposition 10.1.16. *For every risk X and for every feasible allocation $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ the following conditions are equivalent:*

(i) $(Y_1^*, ..., Y_m^*)$ *is Pareto optimal;*

(ii) $(Y_1^*, ..., Y_m^*)$ *is maximal with respect to the coalition preorder* $\precsim = \bigcap_{i=1}^m \precsim_i$.

Proof. (ii) $\Rightarrow$ (i). By contraposition, consider a feasible allocation $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ which is not Pareto optimal. Then there exists $(Y_1, ..., Y_m) \in \mathcal{A}(X)$ such that $Y_i^* \precsim_i Y_i$ for all $i \in \{1, ..., n\}$ and there exists $\bar{i} \in \{1, ..., n\}$ such that $Y_{\bar{i}}^* \prec_i Y_{\bar{i}}$. Therefore, $(Y_1^*, ..., Y_m^*) \precsim (Y_1, ..., Y_m)$ and $\neg((Y_1, ..., Y_m) \precsim (Y_1^*, ..., Y_m^*))$ clearly implies that $(Y_1^*, ..., Y_m^*) \prec (Y_1, ..., Y_m)$. Hence, $(Y_1^*, ..., Y_m^*)$ is not maximal for $\precsim$.
(i) $\Rightarrow$ (ii). By contraposition, consider $(Y_1^*, ..., Y_m^*) \in \mathcal{A}(X)$ which is not maximal for $\precsim$. Then there exists $(Y_1', ..., Y_m') \in \mathcal{A}(X)$ such that $(Y_1^*, ..., Y_m^*) \prec (Y_1', ..., Y_m')$, and this is equivalent to require that $Y_i^* \precsim_i Y_i'$ for all $i \in \{1, ..., m\}$ with at least one index i such that $\neg(Y_i' \precsim_i Y_i^*)$. Hence, we have that $Y_i^* \prec_i Y_i'$. This means that $(Y_1^*, ..., Y_m^*)$ is not Pareto optimal. □

In the sequel we shall denote by

$$\mathcal{S} = \mathcal{B}((X_1, .., X_m)) = \{Y \in \mathcal{A}(X) \mid (X_1, .., X_m) \precsim Y\} \tag{10.1.7}$$

the set of all the feasible allocations for which each agent is at least as well as under the initial allocation $(X_1, .., X_m)$. From Definition 10.1.4 and

Proposition 10.1.16, a $\precsim$-maximal allocation of $\mathcal{S}$ is both individually rational and Pareto optimal.
Let us now introduce the so called *multi-objective maximization problem* associated to m real-valued functions $U_1, ..., U_m$.

Definition 10.1.17. A solution to the problem

$$\begin{aligned} &\max \ (U_1(Y_1), U_2(Y_2), ..., U_m(Y_m)) \\ &\text{sub} \\ &(Y_1, .., Y_m) \in \mathcal{S} \end{aligned} \tag{10.1.8}$$

is $(Y_1^*, ..., Y_m^*)$ provided that one of the following equivalent conditions hold:

1. for all $(Y_1, .., Y_m) \in \mathcal{S}$, $U_i(Y_i) \geq U_i(Y_i^*)$ for all $i \in \{1, ..., m\}$ imply $U_i(Y_i) = U_i(Y_i^*)$ for all $i \in \{1, ..., m\}$;
2. for no $(Y_1, .., Y_m) \in \mathcal{S}$ it holds that $U_i(Y_i) \geq U_i(Y_i^*)$ for all $i \in \{1, ..., m\}$ with at least one strict inequality;
3. for all $(Y_1, .., Y_m) \in \mathcal{S}$, if $U_i(Y_i) > U_i(Y_i^*)$ for some $i \in \{1, ..., m\}$, then there exists some $j \in \{1, ..., m\}$ such that $U_j(Y_j) < U_j(Y_j^*)$.

In the following proposition we are going to use the previous concept in order to produce sufficient conditions for the existence of an optimal solution in the risk sharing setting.

Proposition 10.1.18. *Let U_i be an order-preserving function for the individual preorder $\precsim_i$ $(i \in \{1, ..., m\})$ on $\mathcal{S}$. Then the following statements are valid:*

1. *If $(Y_1^*, ..., Y_m^*)$ is a solution to the problem (10.1.8), then it is an optimal solution;*
2. *If $\precsim_i$ is a total preorder for all $i \in \{1, ..., m\}$, then an optimal solution $(Y_1^*, ..., Y_m^*)$ is a solution to the problem (10.1.8).*

Proof. We prove statement 1. by contraposition. Assume that $(Y_1^*, ..., Y_m^*)$ is not an optimal solution. Then there exists $(Y_1', ..., Y_m') \in \mathcal{S}$

such that $Y_i^* \precsim_i Y_i'$ for all $i \in \{1, ..., m\}$ with one strict inequality. Therefore, since U_i is an order-preserving function for $\precsim_i$ for all $i \in \{1, ..., m\}$, it is clear that $U_i(Y_i^*) \leq U_i(Y_i')$ for all $i \in \{1, ..., m\}$ with one strict inequality, contradicting the fact that $(Y_1^*, ..., Y_m^*)$ is a solution to the problem (10.1.8). Statement 2. will be also proved by contraposition. Assume that $\precsim_i$ is a total preorder for all $i \in \{1, ..., m\}$ and that $(Y_1^*, ..., Y_m^*)$ is not a solution to the problem (10.1.8). Then $U_i(Y_i^*) \leq U_i(Y_i')$ for all $i \in \{1, ..., m\}$ with one strict inequality $U_{\bar{i}}(Y_i^*) < U_{\bar{i}}(Y_i')$. Since U_i is in this case a utility function for $\precsim_i$ for all $i \in \{1, ..., m\}$, we have that $Y_i^* \precsim_i Y_i'$ for all $i \in \{1, ..., m\}$ and $Y_{\bar{i}}^* \prec_{\bar{i}} Y_i'$, contradicting the fact that $(Y_1^*, ..., Y_m^*)$ is optimal. This consideration completes the proof. □

The following corollary concerning the case of total preorders and the corresponding utilities is immediate and we omit its proof.

Corollary 10.1.19. *Let $\precsim_i$ be a total preorder for all $i \in \{1, ..., m\}$ and let U_i be a utility function for $\precsim_i$ for all $i \in \{1, ..., m\}$. Then the following statements are equivalent:*

1. *$(Y_1^*, ..., Y_m^*)$ is a solution to the problem (10.1.8);*

2. *$(Y_1^*, ..., Y_m^*)$ is an optimal solution.*

CHAPTER 11

Domanda e offerta

11.1 Il bene assicurazione

Trattandosi di un puro servizio e non un accessorio di prodotto materiale, il bene assicurazione gode delle seguenti caratteristiche:

1. immaterialità;
2. contestualità tra produzione e vendita e quindi impossibilità di costituire scorte;
3. percezione, da parte dell'acquirente, del suo valore intrinseco solo al momento dell'utilizzo;
4. particolare situazione concorrenziale dal lato dell'offerta.

Immaterialità significa non tangibilità fisica. Ne discende l'importanza del fattore umano, da cui la frequenza del contatto fisico con l'utenza.
L'impossibilità di immagazzinamento si risolve, in questo caso, in un fattore di vantaggio, in quanto l'assicuratore resta libero da ogni problema relativo alla gestione ed all'imagazzinamento delle scorte.
La caratteristica di servizi di poter *essere apprezzati solo al momento del loro utilizzo* è aggravata, nel settore assicurativo, dal differimento temporale della prestazione al verificarsi del sinistro o al particolare evento della vita umana al quale essa è collegata. L'assicuratore deve mettere in atto un valido sistema di comunicazione affinché il cliente possa effettuare un confronto reale tra il sacrificio sopportato e l'utilità futura della garanzia promessa.

La particolare *situazione concorrenziale* è influenzata anche dal regime giuridico-vincolistico a cui è sottoposto in generale il settore servizi e il settore assicurativo in particolare.

Ai caratteri sopra accennati possiamo aggiungere i seguenti, che appaiono più pertinenti al bene assicurazione:

1. scarsa differenziabilità;

2. ridottissima ostentabilità.

Quanto alla *scarsa differenziabilità* del prodotto assicurativo, questa deriva solo in parte dalla tecnologia applicata per produrlo. In effetti, i metodi statistico-attuariali di gestione del rischio sono sotanzialmente gli stessi per tutte le compagnie. Si tenga presente, piuttosto, che l'utente, soprattutto per le assicurazioni di massa, non è in grado di differenziare nettamente la produzione di un'impresa da quella di un'altra.
La *ridottissima ostentabilità* del prodotto assicurativo è riconducibile alla constatazione che esso non può essere utilizzato per comprovare, o almeno per vantare, il raggiungimento di un certo "status" sociale (a differenza di un'automobile sportiva o di una casa per le vacanze).

Per lo studio del bene *assicurazione* è fondamentale la distinzione tra *prodotto danni* e *prodotto vita.*

Il prodotto danni è un prodotto a contenuto strettamente assicurativo, richiesto e fornito per neutralizzare gli effetti dannosi del manifestarsi di rischi statici. Come tale è sostituibile solo parzialmente mediante operazioni di prevenzione che eliminino o riducano le conseguenze dell'evento temuto. La sua sostituibilità parziale e non totale deriva non solo dall'alto costo della prevenzione, ma dall'impossibilità di eliminare, attraverso dispositivi di sicurezza, congegni e pratiche standardizzate, i rischi oggi assicurabili, a meno di non voler rinunciare allo svolgimento dell'attività da cui sorgono.

Il *prodotto vita* è invece un prodotto che ha connotazione solo in parte assicurativa. Infatti l'assicurazione vita, unitamente al rischi di mortalità, copre il *rischio di interesse*, garntendo una prestazione a lungo termine che deve avere un contenuto reale, per capitale e rendimento. La presenza della componente finanziaria impone che il prodotto vita si confronti sul mercato con gli altri tipici prodotti finanziari: titoli a reddito predeterminato, azioni, quote di fondi comuni, piani d'accumulo del risparmio, gestioni patrimoniali, ecc.. La *componente previdenziale* porta il prodotto vita a misurarsi con le prestazioni della previdenza sociale pubblica, di cui è strumento integrativo. Una caratteristica del prodotto vita è la sua non *redimibilità.* Quando l'assicurato stipula la polizza, contrae un impegno a lungo termine che va a intaccare una situazione futura a lui nota. Una delle ragioni che frenano l'acquisto di una polizza vita è la sua *non disinvestibilità*, se non sopportando una elevata penalizzazione, data dalla differenza tra la riserva matematica accumulata e il suo valore di riscatto previsto contrattualmente.

11.2 Domanda di assicurazione

I seguenti fattori influenzano la domanda di assicurazione:

1. reddito pro-capite e sua distribuzione territoriale;
2. esistenza di beni sostituibili;
3. congiuntura economica;
4. componente psicologica;
5. interazione assicurato-assicuratore;
6. tipo di attività esercitata o tipo di bene posseduto o prodotto;
7. incentivazione fiscale.

Il comparto assicurativo la cui domanda risente maggiormente del legame col reddito è quello vita. Si nota inoltre che nei Paesi in cui lo sviluppo economico è diversificato da regione a regione, la spesa media pro capite per assicurazioni è molto più bassa rispetto a quelli che godono di una distribuzione territoriale del reddito più equilibrata.
L'espansione delle *assicurazioni sociali obbligatorie* frena la spinta a contrarre la libera assicurazione, almeno fino a quando le prestazioni fornite dal comparto pubblico sono giudicate sufficientemente appetibili.
Il rapporto esistente fra domanda di assicurazione e congiuntura economica va esaminato separatamente per l'assicurazione danni e l'assicurazione vita. La sensibilità del monte premi vita al prodotto interno lordo è provata. Quello che invece recenti statistiche hanno messo in forse è il fatto che un alto tasso d'inflazione porti ad un decremento dei premi vita.

Nei rami danni invece, quando si riducono le prospettive di profitto per la lievitazione dei costi di produzione o, soprattutto, per una decelerazione della domanda, gli assicurati-imprese manifestano poco interesse ad estendere le coperture.

11.3 Offerta di assicurazione

A differenza di quanto avviene in altri settori, quali quello industriale, non c'è, in campo assicurativo, quel vincolo tecnico di capacità massima di produzione rappresentato dagli impianti e dall'attrezzatura fissa, per cui di fatto l'offerta è *espandibile* con grande facilità e *adattabile* alle richieste di mercato, almeno fino a quando si resta nell'ambito di rischi conosciuti e sperimentati dal mercato nazionale o internazionale.

I seguenti fattori condizionanao l'offerta di assicurazione.

1. mancanza di casualità dell'evento;
2. impossibiltà di valutare con sufficiente attendibilità la frequenza di accadimento e il costo medio del sinistro;
3. dimensione unitaria dei singoli rischi;
4. valore del massimo danno probabile;
5. capacità di assorbimento del mercato assicurativo.

Non si può, quindi, parlare di offerta illimitata di assicurazione, ma è altrettanto ovvio che più aumenta la numerosità del portafoglio, tanto più migliorano le condizioni di gestione tecnica del rischio, perchè la frequenza prevista tende alla probabilità dell'evento, non conosciuta dall'attuario.
La *divisibilità* del prodotto rende possibile trasferire, attraverso gli istituti della riassicurazione e dei pools, le quote di rischio che non possono essere trattenute in proprio dall'assicuratore originario.
Il servizio è prodotto solo quando le condizioni di mercato consentano di conseguire un premio adeguato nel lungo periodo. Da ciò consegue una tendenza dell'offerta ad assecondare la domanda nei mercati in cui il premio, per l'assicuratore, copre il costo industriale del rischio e le spese di gestione, e lascia un soddisfacente margine di profitto.

CHAPTER 12

Exercises

12.1 Risk aversion

1. Consider the following utility functions u (defined over wealth $x > 0$):

   (a) $u(x) = -\frac{1}{x}$ $\qquad [\alpha_R(x) = \frac{2}{x}]$;

   (b) $u(x) = log\, x$ $\qquad [\alpha_R(x) = \frac{1}{x}]$;

   (c) $u(x) = -x^{-\gamma}$ $\qquad [\alpha_R(x) = \frac{\gamma+1}{x}]$;

   (d) $u(x) = -e^{-\gamma x}$ $\qquad [\alpha_R(x) = \gamma]$;

   (e) $u(x) = \frac{x^\gamma}{\gamma}$ $\qquad [\alpha_R(x) = \frac{1-\gamma}{x}]$;

   (f) $u(x) = \alpha x - \beta x^2$ $\qquad [\alpha_R(x) = \frac{2\beta}{\alpha - 2\beta x}]$.

   Questions:

   (a) Check that they are well behaved ($u' > 0$, $u'' < 0$) or state restrictions on the parameters so that they are. For the last utility function, take positive α and β, and give the range of wealth over which the utility function is well behaved.

   (b) Compute the absolute risk-aversion coefficients.

   (c) What is the effect of the parameter γ (when relevant)?

   (d) Classify the functions as increasing/decreasing absolute risk-aversion utility functions .

Loss	Probability
1000	10%
2000	20%
3000	35%
5000	20%
6000	15%

2. An agent faces a risky situation in which he can lose some amount of money with probabilities given in the table:

   This agent has a utility function of wealth of the form

$$u(x) = \frac{x^{1-\gamma}}{1-\gamma} + 2$$

   His initial wealth level is 10000 and his is γ equal to 1.2.

   (a) Calculate the certainty equivalent of this prospect for this agent. Calculate the risk premium.

   (b) What would be the certainty equivalent of this agent if he would be risk neutral?

3. Suppose you have to pay €2 for a ticket to enter a competition. The prize is €19 and the probability that you win is $\frac{1}{3}$. You have an expected utility function $u(x) = log\, x$ and your current wealth is €10.

   (a) What is the certainty equivalent of this competition?

   (b) What is the risk premium?

   (c) Should you enter the competition?