FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
DIPARTIMENTO DI MATEMATICA “GUIDO CASTELNUOVO”
Dottorato di Ricerca in Matematica XXVI ciclo

The Target-Based Utility Model. The role of Copulas and of Non-Additive Measures

Candidate: *Fabio FANTOZZI*

Thesis advisor: *Prof. Fabio SPIZZICHINO*

Fabio Fantozzi

The Target-Based Utility Model
The role of Copulas and of Non-Additive Measures

Ph.D. Thesis. Sapienza – Università di Roma

email: `fantozzi@mat.uniroma1.it`

Contents

Introduction

This thesis covers topics that recently emerged in the field of decisions under risk and uncertainty. The main topic of this work is the *target-based approach* to utility theory. A rich literature has been devoted in the last decade to this approach to economic decisions (see [**19, 20, 28, 29, 138, 139**]). Originally, interest has been concentrated on the *single-attribute* case [**19, 28, 29**] and, more recently, extensions to multi-attribute case have been studied [**20, 138, 139**]. This literature is still growing, with a main focus on applied aspects (see, for example, [**13, 144, 145**]). We will, on the contrary, concentrate attention on some aspects of theoretical type, related with the multi-attribute case.

Various mathematical concepts, such as non-additive measures, aggregation functions, multivariate probability distributions, and notions of stochastic dependence emerge in the formulation and the analysis of target-based models, see [**38**]. It is to be said that notions in the field of non-additive measures and aggregation functions are quite common in the modern economic literature. They are used in game theory (see, for example, [**71, 142**]) and multi-criteria decision aid (see [**3, 62, 63, 69, 80, 85**]). In such fields, one aims to finding the best alternative for a Decision Maker (DM), or classifying the set of good alternatives in choices with many criteria, for situations where uncertainty is not present. These notions have generally been used to go beyond the classical principle of maximization of expected utility in decision theory [**59, 77, 79, 91, 121, 141**]. Along our work, on the contrary, we show how non-additive measures and aggregation functions are of interest even in the frame of the classical utility theory. More precisely we show that they emerge in a natural way in the target-based approach when considering the multi-attribute case. Furthermore we explain how they combine with the analysis of multivariate probability distributions and with concepts of stochastic dependence.

For what concerns non-additive measures, we pay particular attention to the concept of *capacity*, or *fuzzy measure*, that constitutes a specific class of such measures that enjoys the property of monotonicity. Capacities, on the family of subsets of a finite space, have been introduced by Choquet in [**34**] and independently defined by Sugeno in [**135**] in the context of fuzzy integrals. Given a finite set Ω, with corresponding power set 2^Ω, a capacity is a set function $m : 2^\Omega \to [0, 1]$ satisfying

- $m(\emptyset) = 0, m(\Omega) = 1;$
- $m(S) \leq m(T)$ for all sets $S, T \in 2^\Omega$ such that $S \subseteq T$.

Such capacities find many applications. For example, as mentioned above, in game theory, where they are used to assess the right importance to each component of a coalition, or in multi-criteria decision making, representing degrees of satisfaction of investors fulfilling a defined set of objectives (see, for example, [**64**]). Capacities can be better studied through the use of some algebraic transforms, like the *Möbius transform* [**115**], the *Shapley* [**125**] and the *interaction transforms* [**103**], and others.

In particular the Möbius transform M_m of a capacity m is a function satisfying the equality

$$M_m(S) = \sum_{T \subseteq S} (-1)^{|T|+1} m(T),$$

for any set $S \in 2^{\Omega}$. This object turns out to be very useful in multi-criteria decisional problems (see [**66**]) as, in particular, for problems described by the target-based model, as we will see later. A first application of the Möbius transform for capacities arises in the theory of aggregation functions and of non-additive integrals. Aggregation functions are built from capacities and inherit their basic feature of monotonicity. The idea of aggregation consists in summarizing the information contained in an $n-$dimentional vector to a single representative value. This value is a sort of average and it is expressed in terms of the underlying capacity. Also *non-additive integrals* are built by means of capacities, of which they represent a natural extension. They are also known as *fuzzy integrals* and take this name from the fuzzy measures from which they derive. An important feature of this kind of integrals is that, in their turn, they provide an extension of Lebesgue-kind integrals based on additive measures.

The most common fuzzy integral is the *Choquet integral*, introduced by Choquet in 1953 and rediscovered in 1986, when David Schmeidler [**121**] first put forward an axiomatic model of choice with non-additive beliefs. Let m be a capacity defined on a discrete set of indices $N := \{1, \ldots, n\}$ and let $x_1, \ldots, x_n \in \mathbb{R}_+$. The discrete Choquet integral of a function $x : N \to \mathbb{R}_+$ with respect to the capacity m is then defined as

$$Ch_m(x) := \sum_{i=1}^{n} [x_{(i)} - x_{(i-1)}] m(\{\sigma(i), \ldots, \sigma(n)\}),$$

where $\sigma(i)$ is the element of N corresponding to $x_{(i)}$, $x_{(1)} \leq \ldots \leq x_{(n)}$ and $x_{(0)} := 0$.

The Choquet integral, together with Sugeno integral [**135**] and other fuzzy integrals, has been largely used in the context of decision making and analysis of decisions under uncertainty, see [**62, 63, 69**]. In this paper we will show how the Choquet integral emerges as natural in the target-based approach to utilities, in the case in which the coordinates of the target vector manifest comonotonicity. In this view, we will show how our model may represent an extension of the Choquet integral for capacities.

The concept of decision analysis under uncertainty has seen a first formalization in 1944 [**142**], when Von Neumann and Morgenstern (NM) laid the foundations of what is known as the axiomatic theory *expected utility theory*. It should however be remember that a first hint to the use expected utility (instead of the simple average) was introduced by Bernoulli during the 18^{th} century, in the evaluation of the proceeds of a lottery.

Consider a set of random variables $\mathcal{X}$ with values in $\mathcal{Z}$ and consider a preference ordering $\succ$ that we want to use to describe our preferences on $\mathcal{X}$; consider $\mathcal{Z}$ as a complete and separable metric space with its sigma-algebra $\sigma(\mathcal{Z})$. The set $\mathcal{X}$, assumed finite for our purposes, takes the meaning of the set of possible choices, or the set of *lotteries*, while $\mathcal{Z}$ is the set of possible consequences of such choices, or possible outcomes of the lotteries, named *prospects*. The best possible choice will then be the one with best possible outcome. In this perspective, to give a qualitative analysis of the preferences, Decision Makers will try to measure, or at least to order, outcomes by means of some *utility function*.

First of all we have to notice that the goodness of the outcomes is not evaluated in the same way by all the Decision Makers, since the degree of satisfaction for a same result shall be different according to the feelings of each DM. The choice of the utility function is evidently subjective and linked to the behavior of the DM toward risk and uncertainty. So every DM is asked to choose his own function in order to express his preferences among elements of $\mathcal{Z}$. Hence, the utility function will be expressed as $u : \mathcal{Z} \to \mathbb{R}$, where u is assumed to be non-decreasing, following the idea that better outcomes shall be associated with bigger values. To be more precise, according to the preference relation $\succ$ and to NM principles, we have

$$X \succ Y \quad \Leftrightarrow \quad \mathbb{E}[u(X)] > \mathbb{E}[u(Y)] \quad \forall X, Y \in \mathcal{X},$$

where $\mathbb{E}$ has the meaning of the expected value of the function u.

Von Neumann and Morgenstern also devoted attention to the study of the attitudes of DMs towards risk. They classified them according to three categories of behavior, namely *risk neutral*, *risk-seeking* and *risk-averse* Decision Makers. The former are indifferent in choosing between two risky prospects, but with the same expected value; risk-averse Decision Makers, among prospects with the same expected value, prefer the less risky (for them you have $u(\mathbb{E}[X]) \geq \mathbb{E}[u(X)]$, then they make use of a concave utility function); risk-seeker DMs, finally, will manifest the opposite attitude towards risk (and hence will choose a convex utility function).

In this perspective it is interesting to compare investors through their attitudes toward risk. Between two DM playing the same game, but with two different utility functions, it is interesting, for example, to establish which of them is the more risk averse. De Finetti in [**37**] was the first to give a solution of this problem, by introducing the concept of *measure of risk aversion*. Such a concept is strictly linked to the one of *risk premium*, that is the quantity the DM is willing to pay in order to replace the utility of her risky prospect with its expected value. De Finetti's measure of risk aversion is a function that quantifies the risk premium of each DM on the basis of her utility function, describing as more risk averse an individual willing to pay a bigger risk premium.

Other substantial contributions in this direction, have been made over the years by Samuelson [**118**], Pratt [**109**], Rothschild and Stiglitz [**116**], Arrow [**7**], and Fishburn [**57**]. However, objections have also been made to the models based on the maximization of expected utility: first of all Allais, in 1953, offered a paradox in contrast to the theory proposed by Neumann and Morgenstern; Ellsberg [**52**] casted doubts on the axiomatic formulation of Savage [**119**], giving rise to a expected utility theory based on generalized Choquet integrals. However, it was only around 1980 that theories alternative to the expected utility began to be proposed (with the active participation of scholars from disciplines different from the traditional economic, statistical, and mathematical, as philosophers and psychologists). Among the main contributions in this period: Kahneman and Tversky [**77**], Machina [**91, 92**], Quiggin [**111**], Karni [**79**], Gilboa [**59**], Schmeidler [**121**], and others. The theory of choice under uncertainty has taken, since then, very different features than before.

Among the newest concepts of utilities, the one of *target-based utility* plays a central role in our work. Firstly introduced by Castagnoli and LiCalzi in 1996 [**28**], then extended by Bordley and LiCalzi in 2000 [**19**], it gives a quite innovative perspective in the frame of decision theory under risk. In such a model the classical utility function is seen as a distribution function of a (random) target, which the DM wants to overcome with the largest confidence possible. The principle of maximization of the expected utility, in these settings, will then be applied by the DM to the probability of achieving her target. It is interesting to notice that, in the one-dimensional case, the model built in this way is still a utility model in the sense of Von Neumann Morgenstern (NM), while in higher dimensions this parallel, in general, fails. One of the most important and amazing result in this paper is that, the multi-dimensional model that we are going to introduce perfectly fits with the utility models built according to the NM principles, although we make use of non-additive measures to describe preferences involved in it.

Consider, for instance, a utility function u, increasing and with values in $[0, 1]$. The degree of satisfaction of a DM adopting such a function u is then ranged, without loss of generality, in this interval, where 1 represents full accomplishment of DM's objective and 0 stand for a total failure. Now consider a random variable T, with values in $\mathbb{R}$, with the meaning of a target to fulfill, and consider its distribution function

$$F_T(x) = \mathbb{P}(T \leq x).$$

As a function ranging in $[0, 1]$, the utility u can be considered as the distribution function F_T of the target T. Then the degree of satisfaction of the

DM will be a non-decreasing function according to F_T. For any prospect $X \in \mathcal{Z}$ it will be of interest, then, to analyze and maximize the quantity

$$\mathbb{P}(T \leq X)$$

as the expected utility of the prospect X. In fact one has

$$\mathbb{E}(X) = \int u(X)\, dF_X(x) = \int \mathbb{P}(T \leq x)\, dF_X(x) = \mathbb{P}(T \leq X).$$

The one-dimensional model naturally follows the one of expected utility, but extensions of the target-based model to the multi-attribute case are not immediate and may not describe multi-attribute utility functions. Some proposal have been made in the recent years, for example by Bordley and Kirkwood [**20**], that considered multi-attribute target-based decisional model with independent targets, by Tsetlin and Winkler [**138, 139**], that look for correspondence between a target-oriented formulation corresponding to a multi-attribute utility function, with particular attention to the two-dimensional case. Given two targets T_1, T_2 with cdf F_1, F_2 and joint law F_{12}, Tsetlin and Winkler describe their target-oriented utility function by

$$u(x_1, x_2) = u_1 F_1(x_1) + u_2 F_2(x_2) + (1 - u_1 - u_2) F_{12}(x_1, x_2),$$

where u_1, u_2 are coefficient representing utilities of single targets achieved.

In our work we introduce and study a more extended version of this multi-attribute model. Our model considers n correlated targets $T_1, \ldots, T_n$ and describes the importance of achieving each one of them by means of a capacity m and its Möbius transform M_m. More in particular, let m be a capacity defined on a indices set $N = \{1, \ldots, n\}$; for any $B \subseteq N$, consider now $F_{B,\overline{B}}$ as the probability of achieving exactly the targets with indices in B and to fail with respect to the others. The utility function u can now be written as

$$u(\mathbf{x}) = \sum_{B \subseteq N} m(B) F_{B,\overline{B}}(\mathbf{x}).$$

By means of the Möbius transform of m, an analogous representation can be given by

$$u(\mathbf{x}) = \sum_{B \subseteq N} M_m(B) F_B(\mathbf{x}),$$

where F_B is the joint law of the targets whose indices are in B. The utility is then described by the capacity m and by the marginal contribution of F, both evaluated over all the subsets of N. The analysis of the capacity m can then be shifted to the study of its transform M_m, analogously the joint laws F_B of the targets can be rewritten in terms of their connecting copulas C_B, for any $B \subseteq N$.

The concept of copula constitutes a very important tool for this work. Properties of the copulas are for first studied to better describe the target-based multi-attribute model: to represent the interaction among goods in which a DM invests and to define properties of risk aversion and correlation aversion of the DM. Moreover they draw a link between target-based model and the Choquet integral, since, in the case when the connecting copula of the targets is the comonotonicity copula, the expression of the expected utility is actually given by a Choquet integral.

Copulas play an important role also in the last part of the work, in which we discuss the comparison between classical stochastic order and the concept of *stochastic precedence*. The stochastic precedence between two real-valued random variables has often emerged in different applied frameworks: it finds applications in various statistical contexts, including testing and sampling (see [**18**]), reliability modeling, tests for distributional equality versus various alternatives, and the relative performance of comparable tolerance bounds (see [**5, 118**]). Furthermore, this concept arises in the probabilistic context of Markov models for waiting times to occurrences of words in random sampling of letters from an alphabet (for references, see [**40, 41, 42, 43**]).

For two given random variables X_1 and X_2, with distributions F_1 and F_2, we have that $X_1 \prec_{st} X_2$ in the sense of the usual stochastic order if

$$F_1(x) \geq F_2(x), \text{ at any point } x,$$

while we say that X_1 *stochastically precedes* X_2 ($X_1 \preceq_{sp} X_2$) if

$$\mathbb{P}(X_1 \leq X_2) \geq \frac{1}{2}.$$

In this paper we consider a slightly more general, and completely natural, concept of stochastic precedence and analyze its relations with the notions of stochastic ordering. Motivations for our study arise from different fields, in particular from the frame of Target-Based Approach in decisions under risk. Although this approach has been mainly developed under the assumption of stochastic independence between Targets and Prospects, our analysis concerns the case of stochastic dependence, that we model by means of a special class of copulas, introduced for the purpose. Examples are provided to better explain the behavior of the target-based model under changes in the connecting copulas of the random variables, especially regarding their properties of symmetry and dependence.

Along our work we also trace connections to reliability theory, whose aim is studying the lifetime of a system through the analysis of the lifetime of its components. In these settings, the target-based model finds an application in representing the behavior of the whole considered system by means of the interaction of its components.

More in particular our work consists of five Chapters that are briefly summarized as follows:

- In the first Chapter we outline some basic notions of monotone (non-additive) measures and related concepts of integral. This topic has been of large importance in last decades and found many applications in decision theory under risk and game theory in particular. Here we introduce the basic concept of capacity and provide an insight to what is called "theory of aggregation".

- In Chapter 2 we fix attention on the concept of copula. Copulas are the most common aggregation function that are used for expressing joint laws of random variables in terms of their marginal distributions. We will review some of the main characteristics of such functions and provide examples useful for our work.
- Chapter 3 gives an overview of the theory of risk and decisions under risk and uncertainty. It introduces the von Neumann-Morgenstern theory of expected utility and gives a brief discussion about the main features of behavior of Decision Makers facing risky situations.
- In Chapter 4 we discuss the target-based approach to utility theory and we show the related role of capacities and multi-dimensional copulas. The multi-attribute model for target-based utility introduced in the work also provides connections with different fields, such as the ones of aggregation theory and system reliability. We provide extensions and application of such a model for both fields. Furthermore, we investigate properties of risk aversion and correlation aversion for Decision Makers who adopt this model for establishing their utility in investments involving more than one asset.
- The results presented in Chapter 5 are focused on the comparison between the classical *stochastic order* and the quite new concept of *stochastic precedence* among random variables. Such a relationship is explained in terms of their connecting copulas and relative properties and it is enclosed with an application to one-dimensional target-based model for utilities. We also provides several examples showing disagreement between stochastic order and stochastic precedence, principally due to properties concerning dependence and symmetry of connecting copulas.

At the end of this work, a final Section will present concluding remarks and perspectives for future work.

CHAPTER 1

Non-Additive Measures

Non-additive measure theory has made a significant progress in recent years and has been intensively used in many fields of applied mathematics, in economics, decision theory and artificial intelligence. In particular, non-additive measures are used when models based on classical measures are not appropriate.

In this work we will concentrate our attention in the possible applications for the study of expected utility models. Von Neumann and Morgenstern proposed in [**142**] a model that have been widely used for solving decision theoretical problems through decades, though it has its limitations. Savage in [**119**] improved it significantly by including subjective probabilities. However, probabilities used in his model remained additive. To make expected utility models more flexible, additive subjective probabilities were later replaced by non-additive probabilities, called *capacities* or *fuzzy measures*.

Capacities used in expected utility models prove to be a very flexible tool to model different kinds of behavior. Most Decision Makers, for example, overestimate small and underestimate large probabilities. Further, most Decision Makers prefer decisions where more information is available to decisions with less available information. Such a behavior is known as *uncertainty aversion* and turns out to be impossible to be expressed through an additive model. On the other side, it is possible to describe basic properties of *risk aversion* through additive model by transforming utility functions. For a deeper analysis of the aversion towards risk it is necessary to pass to non-additive measures.

Many other results and concepts related with additive measure or probability theory have natural generalizations to non-additive theory. Integration with respect to nonadditive measures, for example, ca be made by replacing the usual Lebesgue integral with the more general concept of *fuzzy integral*. Fuzzy integrals, in particular, are important tools used to solve problems of decision under risk in finance as well as in game theory.

We start this Chapter recalling the very basic aspects of probability measures. We will then introduce the more general concept of *capacity* or *fuzzy measure*, obtained by dropping some of the main properties of the probability measures, and many of its most important properties. Finally we will briefly discuss about integrals built with respect to fuzzy measures, with particular attention to the well known Choquet integral. For our purposes we restrict our study to the case of finite sets.

1. The Inclusion-Exclusion Principle

In combinatorics, the inclusionexclusion principle is a counting technique which generalizes the familiar method of obtaining the number of elements in the union of finite sets. Giancarlo Rota said in [**115**]: "One of the most useful principles of enumeration in discrete probability and combinatorial theory is the celebrated principle of inclusionexclusion. When skillfully applied, this principle has yielded the solution to many combinatorial problems". Actually, this basic combinatorial tool also finds many applications in number theory and in measure theory and, for our purposes, will be very useful for the statement of the Target-Based model for utility.

We start by introducing some useful notation. We consider the set of indices $N := \{1, \dots, n\}$, with I a subset of N with cardinality $|I|$. Furthermore, we consider a collection of finite sets $\{E_1, \dots, E_n\}$. We will denote with E_I the set $\cap_{i \in I} E_i$.

The inclusion-exclusion principle can be used to calculate the cardinality of the union of the sets $\{E_i\}_{i \in N}$, as follows.

THEOREM 1.1 (Inclusion-Exclusion Principle). *Let $E_1, \dots, E_n$ be finite sets. The cardinality of their union is given by*

$$\left| \bigcup_{i=1}^{n} E_i \right| = \sum_{I \subseteq N} (-1)^{|I|+1} |E_I|. \tag{1.1}$$

Notice that, when the cardinality of intersections is regular (namely $|E_I| = \alpha_I$, for all $I \subseteq N$), the formula can be rewritten as follows

$$\left| \bigcup_{k=1}^{n} E_k \right| = \sum_{I:|I|=1}^{n} (-1)^{|I|+1} \binom{n}{|I|} \alpha_I. \tag{1.2}$$

A similar formula can be found in probabilistic terms, when studying the probability of events in a given probability space. For this purpose fix a finite state space Ω and define by 2^Ω its power set. In these hypothesis we can introduce the following Definitions.

DEFINITION 1.2. *A* $\sigma-$algebra $\mathcal{F}$ *is a family of sets in* 2^Ω *such that*

(1) $\emptyset \in \mathcal{F}$*;*
(2) for any set E*, if* $E \in \mathcal{F}$ *then its complement* $E^c \in \mathcal{F}$*;*
(3) given a countable family of sets $\{E_n\}_{n \geq 1}$*,* $\cup_n E_n \in \mathcal{F}$*.*

DEFINITION 1.3. *A* probability measure *over* $(\Omega, \mathcal{F})$ *is a function* $\mathbb{P} : \mathcal{F} \to [0, 1]$ *such that*

(1) for all set $E \in \mathcal{F}$*,* $\mathbb{P}(E) \geq 0$*;*
(2) $\mathbb{P}(\Omega) = 1$ *(and* $\mathbb{P}(\emptyset) = 0$*);*
(3) for any countable collection of mutually disjoint sets $\{E_n\}_{n \geq 1}$*, one has that* $\mathbb{P}(\cup_n E_n) = \sum_n \mathbb{P}(E_n)$ *(*countably additivity*).*

A triplet $(\Omega, \mathcal{F}, \mathbb{P})$ will be called a (finite) *probability space*. Under these hypothesis we are ready to introduce the probabilistic version of the principle.

PROPOSITION 1.4. *Given a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a finite family of events* $\{E_1, \dots, E_n\} \in \mathcal{F}$*, the inclusion-exclusion principle reads*

$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{I \subseteq N} (-1)^{|I|+1} \mathbb{P}\left(\bigcap_{i \in I} E_i\right). \tag{1.3}$$

For practical purposes we give explicit formula for the case $n = 2$, for which the principle reduces to

$$\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2). \tag{1.4}$$

An analogous of formula (1.2) can be given in probabilistic settings, when the measure of each event depends only on its cardinality. We have

$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{I:|I|=1}^{n} (-1)^{|I|+1} \binom{n}{|I|} \mathbb{P}(E_I). \tag{1.5}$$

We remember that no hypothesis about dependence of events $\{E_n\}_n$ are given. In case that the events are mutually pairwise independent, for the countably additivity property of the probability measure, formula (1.3) reduces to

$$\mathbb{P}\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mathbb{P}(E_k).$$

2. Capacities

In this Section we will discuss about measures that do not benefit from the additivity, typical property of probability measures. By the way we will concentrate our attention on the weaker property of monotonicity, and to its extensions, by introducing the concept of $k-$*monotonicity*. This property is strictly linked to the inclusion-exclusion principle presented in the previous Section. We start considering the following

DEFINITION 1.5. *Given a set* Ω*, a relationship* $\subseteq$ *over* 2^Ω *with the properties of reflexivity, antisymmetry and transitivity is called a* partial order. *The set* 2^Ω *embedded with such relationship will be then called a partially ordered set, i.e. a* poset*.*

The notation $(2^\Omega, \subseteq)$ is sometimes used in literature to identify such poset. Over this structure the following property can be considered.

DEFINITION 1.6. *Let the poset* 2^Ω *be given. A function* $m : 2^\Omega \to \mathbb{R}$ *is called* monotone *if and only if, for all sets* $E, B \in 2^\Omega$ *such that* $E \subseteq B$*,*

$$m(E) \leq m(B). \tag{1.6}$$

It is straightforward to notice that any probability measure enjoys the property of monotonicity. Notice also that such definition can be given for set functions by replacing Ω with an set of indexes N. Let now k be an integer such that $k \geq 2$. We have the following

DEFINITION 1.7. *Let $K := \{1, \ldots, k\}$. A function m is said $k-$*monotone *if and only if*

$$m\Big(\bigcup_{j\in K} E_j\Big) \geq \sum_{I\subseteq K} (-1)^{|I|+1} m\Big(\bigcap_{i\in I} E_i\Big), \tag{1.7}$$

for all $E_j \in 2^\Omega$, $j \in K$. Furthermore we will say that such a function m is totally monotone *if it is $k-$monotone for all $k \geq 2$.*

For the special case $k = 2$ formula (1.7) can be rewritten as

$$m(E_1 \cup E_2) \geq m(E_1) + m(E_2) - m(E_1 \cap E_2). \tag{1.8}$$

Such a property of $2-$monotonicity is also called *supermodularity*. Notice that if m is $k-$monotone for some $k \geq 2$, than it is $k'-$monotone for any $k' \leq k$. Notice furthermore that, if m is a probability measure, we have an equality in (1.7) (and in (1.8)), and such an equation coincides with the one of inclusion-exclusion principle given in (1.3) (respectively in (1.4)). This fact is due to the additivity of the measure and can be interpreted in the sense that any probability measure is $\infty-$monotone.

As we were saying at the beginning of this Section, we want to study functions that are not assumed to be additive. Even for such functions, some useful properties can be given, such as the following result, due to Chateneuf and Jaffray (see [**31**]).

PROPOSITION 1.8. *Let $m : 2^\Omega \to \mathbb{R}$ be a $k-$monotone function. If $m(\{\omega\}) \geq 0$ for all $\omega \in \Omega$, then the function m is also monotone (and hence non-negative).*

By means of Definition 1.6 and Proposition 1.8 we are now ready to introduce the following

DEFINITION 1.9. *A* fuzzy measure *or* capacity *is a bounded function $m : 2^\Omega \to \mathbb{R}$ that satisfies*

(1) $m(\emptyset) = 0$;
(2) $m(A) \leq m(B)$ for any A, B sets in 2^Ω such that $A \subseteq B$.

Since a capacity is a bounded set function, it is usual to rescale it to the set of values $[0, 1]$, so that $m(\Omega) = 1$. We now give some basic example of capacities.

EXAMPLE 1.10. *Let $N = \{1, \ldots, n\}$ as usual. For any set $E \subseteq N$ define*

$$m^0(E) := \begin{cases} 1, & \textit{if } E = N; \\ 0, & \textit{otherwise.} \end{cases}$$

Such a capacity is the minimal possible over the set N. On the other side one can define the maximal capacity by a function

$$m^1(E) := \begin{cases} 0, & \text{if } E = \emptyset; \\ 1, & \text{otherwise.} \end{cases}$$

Both minimal and maximal capacities are examples of $0-1$ capacities, i.e. capacities assuming only values 0 and 1. This kind of capacities are very used in reliability theory and describe the functioning of series and parallel systems respectively. For further details on the topic see, for example, [**10**]. *Notice that the former enjoys the property of $\infty-$monotonicity, while the latter is $\infty-$alternating, property that we introduce in the following Definition.*

DEFINITION 1.11. *A function $m : 2^\Omega \to \mathbb{R}$ is said $k-$*alternating *if for all families of subsets of Ω of k elements*

$$m\Bigg(\bigcap_{j\in K} E_j\Bigg) \leq \sum_{I\subseteq K} (-1)^{|I|+1} m\Bigg(\bigcup_{i\in I} E_i\Bigg), \tag{1.9}$$

where once more K stands for the set $\{1, \ldots, k\}$. A totally alternating function is $k-$alternating for every $k \geq 2$.

A $2-$alternating set function is also called *submodular* and its expression reads

$$m(E_1 \cap E_2) \leq m(E_1) + m(E_2) - m(E_1 \cup E_2). \tag{1.10}$$

DEFINITION 1.12. *A fuzzy measure $m : 2^\Omega \to \mathbb{R}$ is said symmetric if its values depend only on the cardinality of the underlying sets, i.e if for any set $E \in 2^\Omega$, $m(E) = m(|E|)$.*

Generally speaking to know a capacity down pat one needs $2^{|\Omega|}$ pieces of information, for a symmetric one the amount of information needed is drastically reduced to $|\Omega|$. Under such condition one can rewrite equation (1.7) as

$$m\Bigg(\bigcup_{j\in K} E_j\Bigg) \geq \sum_{I:|I|=1}^{|K|} (-1)^{|I|+1} \binom{|K|}{|I|} m(|I|). \tag{1.11}$$

An analogous formula for condition (1.9) can be written in a similar way.

DEFINITION 1.13. *A function $m : 2^\Omega \to \mathbb{R}$ is said* superadditive *if, for any family of sets $\{E_n\}_{n\geq 1} \in 2^\Omega$,*

$$m\Bigg(\bigcup_{n\geq 1} E_n\Bigg) \geq \sum_{n\geq 1} m(E_n). \tag{1.12}$$

It will be called subadditive *if the inequality in* (1.12) *is reversed. The function m will be called additive if both superadditive and subadditive.*

A more general notion of additivity can be given for capacities as it follows.

DEFINITION 1.14. *A fuzzy measure* $m : 2^{\Omega} \to \mathbb{R}$ *is said* $k-$*additive if, for any family of sets of* Ω *with* k *elements* $A_1, \ldots, A_k$,

$$\sum_{I \subseteq K} (-1)^{|K \setminus I|} m\left(\bigcup_{i \in I} A_i\right) = 0. \tag{1.13}$$

We introduce now the concept of *dual* of a fuzzy measure.

DEFINITION 1.15. *Given a fuzzy measure* $m : 2^{\Omega} \to [0, 1]$, *its dual measure* m^* *is defined by*

$$m^*(A) = 1 - m(A^c), \tag{1.14}$$

for all sets $A \in 2^{\Omega}$. *The set* A^c, *as usual, stands for the complement of* A.

The dual m^* is a fuzzy measure itself and can enjoy all the properties of fuzzy measures. In particular, if a fuzzy measure m is superadditive its dual m^* will be subadditive and, viceversa, if m is subadditive its dual will enjoy the property of superadditivity; if m is supermodular then m^* is submodular and reciprocally; finally if one of them is $k-$monotone the other one will be $k-$alternating (see [**65**] for further details). An example in this direction is given by the minimal and maximal capacities introduced above. It is straightforward to notice that one is the dual of the other one.

Capacities may arise by manipulating probability measures, as follows.

EXAMPLE 1.16. *Let* $\mathcal{P}$ *a given class of probability measures defined on* $(\Omega, \mathcal{F})$. *For any given* $E \subseteq \Omega$, *the functions*

$$m_{\sup}(E) = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}(E)$$
$$m_{\inf}(E) = \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}(E)$$

are examples of capacities built in this way. Notice that on $\{1, 2\}$ *the first capacity is submodular while the second is supermodular. Furthermore they are reciprocally dual measures.*

Capacities can be also obtained through the composition of a probability measure $\mathbb{P}$ with a *distortion* γ, in the following way.

DEFINITION 1.17. *Let* $\mathbb{P}$ *a probability measure defined on a state space* Ω. *Let furthermore* $\gamma : [0, 1] \to [0, 1]$ *be an increasing function with* $\gamma(0) = 0$ *and* $\gamma(1) = 1$. *The function* $m = \gamma \circ \mathbb{P}$ *is called* distorted probability *while* γ *is the corresponding* distortion.

Such a capacity enjoy an important property.

PROPOSITION 1.18. *A function built by means of a probability measure and a distortion, as in Definition 1.17, is monotone and hence a capacity. Furthermore, if* γ *is convex then the capacity* m *is supermodular; with* γ *concave,* m *is submodular.*

For the proof of the proposition above and for further details see for example [**44**].

3. Möbius Transforms

Due to August Möbius (1790 - 1868), the so called Möbius Transform is a particular and very useful tool that belongs to number theory, but finds many applications also in other fields, especially in the one of non-additive measures. In this Section we introduce only the basic concepts needed for our dissertation. For further details see [**115**].

Let Ω be a poset, with 2^Ω the associated power set. We often refer to Ω like to an index set $N := \{1, \dots, n\}$. To any function $m : 2^\Omega \to [0,1]$ (or more in general with values in $\mathbb{R}$) it can be associated another function $M_m : 2^\Omega \to [0,1]$ by

$$M_m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} m\left(\bigcup_{i \in I} B_i \right) \quad \text{for all } A \in 2^\Omega. \tag{1.15}$$

If furthermore Ω is finite, equation (1.15) can be rewritten as

$$M_m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} m(B) \quad \text{for all } A \in 2^\Omega. \tag{1.16}$$

A particular feature of this correspondence is that it is one-to-one, since conversely

$$m(A) = \sum_{B \subseteq A} M_m(B) \quad \text{for all } A \in 2^\Omega. \tag{1.17}$$

The validity of formula (1.17) is proved by Shafer in [**123**].

The Möbius transform is very useful in the study of capacities since many of the properties of such measures can be expressed through their Möbius representation. First of all notice that any set of 2^n coefficients $\{m(A)\}_{A \subseteq \Omega}$ does not necessarily correspond to the Möbius transform of a capacity on Ω. The boundary and monotonicity conditions must be ensured (see [**31**]), i.e. we must have

$$M_m(\emptyset) = 0, \quad \sum_{B \subseteq \Omega} M_m(B) = 1, \quad \text{and} \quad \sum_{B \subseteq A} M_m(B) \geq 0 \ \forall A \in 2^\Omega. \tag{1.18}$$

A very important property concerns $k-$monotonicity and reads as follows.

PROPOSITION 1.19. *A fuzzy measure m is $k-$monotone if and only if its Möbius transform M_m is non-negative for any set of cardinality less or equal than k i.e., for all $E \in 2^\Omega$ with $|E| \leq k$, $M_m(E) \geq 0$.*

As a corollary of the above proposition, we can say that the Möbius transform of a totally monotone fuzzy measure is always non-negative.

PROPOSITION 1.20. *A fuzzy measure m is $k-$additive if and only if its Möbius transform M_m of order greater than k are null i.e., for all $A \in 2^\Omega$ with $|A| > k$, $M_m(A) = 0$, and $M_m(B) > 0$ for at least one element B with $|B| = k$.*

The result above follows directly from Definition 1.14 of $k-$additivity. Finally, an alternative useful representation, given by Shafer in [**123**], is the following.

DEFINITION 1.21. *The co-Möbius representation* $\check{M}_m$ *of* m *is defined by*

$$\check{M}_m(T) := \sum_{S \supset T} m(S). \tag{1.19}$$

This definition let to an useful property linking the transform of a capacity m to its dual m^*.

PROPOSITION 1.22. *Let* m *and* m^* *a pair of dual measures and* M_m *and* M_{m^*} *be their Möbius representation respectively. Then, for any* $T \subseteq N$,

$$M_{m^*}(T) = (-1)^{|T|+1} \sum_{S \supset T} m(S) = (-1)^{|T|+1} \check{M}_m(T). \tag{1.20}$$

4. Interaction Indices

In the framework of cooperative game theory, the concept of interaction index, which can be regarded as an extension of that of value, has been recently proposed to measure the *interaction phenomena* among players. The expression "interaction phenomena" refers to either complementarity or redundancy effects among players of coalitions resulting from the non-additivity of the underlying game. Thus far, the notion of interaction index has been primarily applied to multi-criteria decision making in the framework of aggregation by the Choquet integral. We will provide an insight of these concepts in the following Sections.

For a better comprehension of the interaction phenomena modeled by a capacity, several numerical indices can be computed (see [**94, 95**]). In the sequel, we present two of them in details, the *Shapley value* and the *interaction transform*. The Shapley value was introduced in 1953 by Lloyd Shapley and it is a very important tool in cooperative games. Its main function is that of defining the importance of a single player within the coalition to which he belongs. As an extension of the Shapley value, the interaction transform assigns importance to subsets of any cardinality concerning such a given coalition. Further information on the topic can be found in [**103**] and [**125**].

Shapley noticed in [**125**] that the overall importance of a criterion $i \in N$ into a decision problem is not solely determined by the number $m(\{i\})$, but also by all $m(T)$ such that $i \in T$. Indeed, we may have $m(\{i\}) = 0$, suggesting that element i is unimportant, but it may happen that for many subsets $T \in N \setminus \{i\}$, $m(T \cup \{i\})$ is much greater than $m(\{i\})$, suggesting that i is actually an important element in the decision. To overcome the difficulties in attributing the right weight to each component i of a coalition, Shapley proposed a definition of a coefficient of importance like follows.

DEFINITION 1.23. *The* importance index *of criterion i with respect to m is defined by:*

$$\Phi_m(i) := \sum_{T \subseteq N \setminus \{i\}} \frac{(n-t-1)!t!}{n!} [m(T \cup \{i\}) - m(T)], \tag{1.21}$$

where is intended that $t = |T|$. The Shapley value *is the vector of importance indices $\{\Phi_m(1), \ldots, \Phi_m(N)\}$.*

Having in mind that, for each subset of criteria $T \in N$, $m(T)$ can be interpreted as the importance of T in the decision problem, the Shapley value of i can be thought of as an average value of the marginal contribution $m(T \cup \{i\}) - m(T)$ of criterion i to a subset T not containing it. To make this clearer, it is informative to rewrite the index as follows:

$$\Phi_m(i) := \frac{1}{n} \sum_{t=0}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{T \subseteq N \setminus \{i\} : |T| = t} [m(T \cup \{i\}) - m(T)]. \tag{1.22}$$

A fundamental property is that the numbers $\Phi_m(1), \ldots, \Phi_m(n)$ form a probability distribution over N, in fact

$$\Phi_m(i) \geq 0 \, \forall i \in N \quad \text{and} \quad \sum_{i=1}^{n} \Phi_m(i) = 1.$$

The best known axiomatic supporting the Shapley value is given in the following

THEOREM 1.24. *The numbers $\Phi_m(i)$, with $m : 2^N \to [0,1]$, $i = 1, \ldots, n$, satisfy the following conditions:*

(1) are linear w.r.t. the fuzzy measure, that is, there exist real constants p_T^i ($T \subseteq N$) such that

$$\Phi_m(i) = \sum_{T \subseteq N} p_T^i m(T);$$

(2) are symmetric, that is, for any permutation σ on N, we have

$$\Phi_m(i) = \Phi_{\sigma m}(\sigma(i));$$

(3) fulfill the null criterion *axiom, that is,*

$$m(T \cup \{i\}) = m(T) \, \forall T \subseteq N \setminus \{i\} \Rightarrow \Phi_m(i) = 0;$$

(4) fulfill the efficiency axiom*, that is*

$$\sum_{i=1}^{n} \Phi_m(i) = 1.$$

Let us comment on the axioms presented in this characterization. First of all we ask the importance indices to be linear w.r.t. the corresponding fuzzy measure. Next, the symmetry axiom demands that the indices are independent of the name (label) given to each criterion. The third axiom, which is quite natural, says that when a criterion does not contribute in the decision problem then it has a zero global importance. The last axiom naturally acts as a normalization property.

A very useful property consists in the possibility of rewriting Shapley value in terms of the Möbius representation of m, as

$$\Phi_{M_m}(i) := \sum_{T\subseteq N\setminus\{i\}} \frac{1}{t+1} M_m(T\cup\{i\}). \tag{1.23}$$

Another interesting concept is that of interaction among criteria. Of course, it would be interesting to appraise the degree of interaction among any subset of criteria. Let's start considering a pair of criteria $\{i,j\}\in N$. It may happen that $m(i)$ and $m(j)$ are small and at the same time $m(\{i,j\})$ is large. Clearly, the number $\Phi_m(i)$ merely measures the average contribution that criterion i brings to all possible combinations, but it does not explain why criterion i may have a large importance. In other words, it gives no information on the interaction phenomena existing among criteria. Suppose that i and j are positively correlated or *substitutable* (resp. negatively correlated or *complementary*). Then the marginal contribution of j to every combination of criteria that contains i should be strictly less than (resp. greater than) the marginal contribution of j to the same combination when i is excluded. Thus, depending on whether the correlation between i and j is positive or negative, the quantity

$$(\Delta_{i,j}m)(T) := (T\cup\{i,j\}) - m(T\cup\{i\}) - m(T\cup\{j\}) + m(T)$$

is ≤ 0 or ≥ 0 for all $T\subseteq N\setminus\{i,j\}$, respectively. We call this expression the *marginal interaction* between i and j. Now, an interaction index for $\{i,j\}$ is given by an average value of this marginal interaction. Murofushi and Soneda in [**103**] proposed to calculate this average value as for the Shapley value.

DEFINITION 1.25. *The* interaction index *of criteria i and j related to m is defined by*

$$I_m(i,j) := \sum_{T\subseteq N\setminus\{i,j\}} \frac{(n-t-2)!t!}{(n-1)!}(\Delta_{i,j}m)(T). \tag{1.24}$$

We immediately see that this index is negative as soon as i and j are positively correlated or substitutable. Similarly, it is positive when i and j are negatively correlated or complementary. Moreover, it has been shown in [**65**] that $I_m(i,j)\in[-1,1]$ for all $i,j\in 2^N$. The interaction index among a combination S of criteria was introduced by Grabisch in [**65**] as a natural extension of the case $|S|=2$ and lately axiomatized by Grabisch and Roubens [**71**].

DEFINITION 1.26. *The interaction index of S ($|s|\geq 2$) related to m, is defined by*

$$I_m(S) := \sum_{T\subseteq N\setminus S} \frac{(n-t-s)!t!}{(n-s+1)!}(\Delta_S m)(T), \tag{1.25}$$

where $s = |S|$ *and*

$$(\Delta_S m)(T) := \sum_{L \subseteq S} (-1)^{s-l} m(L \cup T).$$

Finally, it can be also written, in terms of the Möbius representation, as

$$I_{M_m}(S) := \sum_{T \subseteq N \setminus S} \frac{1}{t+1} M_m(T \cup S). \tag{1.26}$$

There is a rich literature regarding this kind of index including, for example, the *Banzhaf index*, *andness* and *orness* indices, *veto* and *favor* and others more. For such a literature reference is made to [**69, 93**].

5. Aggregation Functions

Aggregation functions became in the last decade a very important field of mathematics and information sciences. The idea of aggregation functions is rather simple: they aim to summarize the information contained in a vector of n values by means of a single representative one. Starting from the most simple example, the arithmetic mean, many other kinds aggregation functions were applied in various sectors of research.

The basic feature of all aggregation functions is their nondecreasing monotonicity, as fuzzy measures have. Another axiomatic constraint of aggregation functions concerns the boundary conditions, expressing the idea that "minimal (or maximal) inputs are aggregated into minimal (maximal) output of the scale we work on".

By these first definitions, the class of aggregation functions results really huge and the problem of choosing the right function for a given application really difficult. The study of the main classes of aggregation functions is then very complex, so we just report some of the main examples and features relative to such operators. More information about aggregation functions and operator can be found, for example, in [**12**].

Before recalling the basic definitions, it is opportune to introduce some notations. We will use $\overline{\mathbb{R}}$ for the extended real line $[-\infty, \infty]$, while $\mathbb{I}$ will stand for a generic closed subset of $\overline{\mathbb{R}}$. The symbol N, when not differently specified, will refer to a set of indices with n elements, namely $N := \{1, \ldots, n\}$.

DEFINITION 1.27. *An aggregation function in* $\mathbb{I}^n$ *is a function* $A^{(n)} : \mathbb{I}^n \to \mathbb{I}$ *that*

(1) is non-decreasing in each variable;
(2) satisfies $\inf_{\mathbf{x} \in \mathbb{I}} A^{(n)}(\mathbf{x}) = \inf \mathbb{I}$ *and* $\sup_{\mathbf{x} \in \mathbb{I}} A^{(n)}(\mathbf{x}) = \sup \mathbb{I}$*;*
(3) $A^{(1)}(x) = x$ *for all* $x \in \mathbb{I}$*.*

The integer (n) represents the number of variables considered for A. From now on, when no possibility of mistakes may occur, we will omit to write it. Now we introduce some basic aggregation functions.

- the arithmetic mean AM, defined by

$$AM(\mathbf{x}) := \frac{1}{n}\sum_{i=1}^{n} x_i,$$

that represents an aggregation function for any domain $\mathbb{I}^n$;
- the product $\Pi(\mathbf{x}) = \prod_{i=1}^{n} x_i$ on $[0,1]$ or on $[1,\infty]$;
- the minimum and the maximum, defined on any $\mathbb{I}$, respectively by

$$Min(\mathbf{x}) = \min\{x_1,\dots,x_n\} \text{ and } Max(\mathbf{x}) = \max\{x_1,\dots,x_n\};$$

- the $k-$order statistics $OS_k : \mathbb{I}^n \to \mathbb{I}$, defined for any choice of $\mathbb{I}$ as $OS_k(\mathbf{x}) = x_{(k)}$, where $x_{(k)}$ is the $k-$th elements of the ordered vector $(x_{(1)},\dots,x_{(n)})$;
- the $k-$th projection $P_k : \mathbb{I}^n \to \mathbb{I}$ with $P_k(\mathbf{x}) = x_k$;
- for any $i \in N$, the *Dirac measure* centered on i, defined for any $A \subseteq \Omega$ as

$$\delta_i(A) = \begin{cases} 1 \text{ if } i \in A, \\ 0 \text{ otherwise;} \end{cases}$$

- the *threshold measure* τ_k defined, for any integer $k \in N$, by

$$\tau_k(A) = \begin{cases} 1 \text{ if } |A| \geq k, \\ 0 \text{ otherwise;} \end{cases}$$

As for fuzzy measures, it can be introduced the *dual* of the aggregation function, in the special case in which $\mathbb{I}$ is limited. If not specified, from now on we will assume $\mathbb{I} = [0,1]$.

DEFINITION 1.28. *Let $A : \mathbb{I}^n \to \mathbb{I}$ be an aggregation function. The dual of A is a function $A^d : \mathbb{I}^n \to \mathbb{I}$ such that*

$$A^d(\mathbf{x}) = 1 - A(1-x_1,\dots,1-x_n). \tag{1.27}$$

Notice that the dual of an aggregation function is an aggregation function itself. Moreover it can be easily extended to any limited interval $[a,b] \subset \mathbb{R}$, as $A^d(\mathbf{x}) = a+b-A(a+b-x_1,\dots,a+b-x_n)$.

The aggregation functions may have many properties that we briefly list below.

DEFINITION 1.29 (Monotonicity). *The aggregation function $A : \mathbb{I}^n \to \mathbb{I}$ is* strictly increasing *in each argument if for any two different vectors $\mathbf{x}$ and $\mathbf{x}'$ with $\mathbf{x} < \mathbf{x}'$ ($x_i < x'_i$ for at least for one index i) one has $A(\mathbf{x}) < A(\mathbf{x}')$. It is called* jointly strictly increasing *if for any $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$, such that $x_i < x'_i$ for all entries $i = 1,\dots,n$, $A(\mathbf{x}) < A(\mathbf{x}')$.*

It is immediate to notice that any strictly increasing aggregation function is also jointly strictly increasing, while the viceversa is not true. The product Π on $[0,1]$ is an example of aggregation function that has the latter property but not the former.

DEFINITION 1.30 (Lipschitz condition). *Let* $\|\cdot\|: \mathbb{R}^n \to \mathbb{R}_+$ *be a norm. If* $A : \mathbb{I}^n \to \mathbb{I}$ *satisfies*

$$|A(\mathbf{x}) - A(\mathbf{y})| \leq c\|\mathbf{x} - \mathbf{y}\| \tag{1.28}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and for some positive constant* c*, then* A *is called Lipschitzian. The infimum value* c *for which equation* (1.28) *holds is called the* Lipschitz constant.

Important examples of norms are given by the L_p norm, i.e. the Minkowski norm of order p

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \tag{1.29}$$

and its limit case $\|\mathbf{x}\|_\infty := \max_i |x_i|$ which is the Chebyshev norm. Notice that the aggregation functions Min, Max, AM are Chebyshev norms of constant 1, while Π on $[0,1]$ is $1-$Lipschitz w.r.t. to norm L_p but no more than $n-$Chebyshev.

DEFINITION 1.31 (Symmetry). *The aggregation function* $A : \mathbb{I}^n \to \mathbb{I}$ *is* symmetric *if* $A(\mathbf{x}) = A(\sigma(\mathbf{x}))$ *for any vector* $\mathbf{x} \in \mathbb{I}^n$ *and any permutation* σ *of the elements of the vector* $\mathbf{x}$*, namely* $\sigma(\mathbf{x}) = (x_{(1)}, \ldots, x_{(n)})$.

The symmetry property is essential when considering criteria that do not depend on the order in which they are chosen, maybe because they have the same importance or the original importance attributed by an anonymous Decision Maker is unknown. Notice that all the aggregation functions introduced so far, as Min, Max, AM, Π and so on, are symmetric. An example of non-symmetric aggregation function is given by the *Weighted Arithmetic Mean*

$$WAM_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_i, \tag{1.30}$$

where the weights w_i are such that $\sum_{i=1}^{n} w_i = 1$. This aggregation function represents the simplest way to assess importance to different criteria in a decision problem. Notice that it is a simple extension of the arithmetic mean in fact, when all weights are equal to $1/n$, it trivially reduces to AM. Any non-symmetric function can, anyway summarized by replacing the variables x_i with the corresponding order statistics $x_{(i)}$, $i = 1, \ldots, n$. One of the simplest examples in this direction is given by the *Ordered Weighted Average* function defined as

$$OWA_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_{(i)}. \tag{1.31}$$

This function trivially reduces to WAM when considering an ordered vector and, in turn, to AM if symmetrized.

DEFINITION 1.32 (Idempotence). *An* idempotent *aggregation function $A : \mathbb{I}^n \to \mathbb{I}$ is one that satisfies $A(n \cdot x) = x$, where with $n \cdot x$ we stands for a vector with all identical components x, i.e. $(x_1, \dots, x_n) = (x, \dots, x)$.*

Many of the aggregation functions mentioned above, like AM, WAM, OS_k, P_k, Min, and Max enjoy this property while, for example, Π doesn't.

DEFINITION 1.33 (Associativity). *Let $A : \mathbb{I}^2 \to \mathbb{I}$ an aggregation function. Then it is called* associative *if for all $x_1, x_2, x_3 \in \mathbb{I}$ we have*

$$A(A(x_1, x_2), x_3) = A(x_1, A(x_2, x_3)). \tag{1.32}$$

This property can be suitably extended to generic $n-$ary aggregation functions, as shown in [**70**]. Such functions are easy to build, starting from any $2-$ary associative one, once all inputs to be aggregated are known.

Other interesting properties to highlight regard the elements to be aggregated.

DEFINITION 1.34 (Neutral element). *An element $e \in \mathbb{I}$ is called* neutral element *of an aggregation function $A : \mathbb{I}^n \to \mathbb{I}$ if $A(x_{\{i\}}e) = x$, where the vector $x_{\{i\}}e$ is the one with all components equal to e except the $i-$th one which is x.*

DEFINITION 1.35 (Annihilator). *An element $a \in \mathbb{I}$ is called* annihilator element *of an aggregation function $A : \mathbb{I}^n \to \mathbb{I}$ if for any vector $\mathbf{x} \in \mathbb{I}^n$ such that $a \in \{x_1, \dots, x_n\}$ (at least one element of the vector $\mathbf{x}$ is equal to a) we have $A(\mathbf{x}) = a$.*

Finally, like fuzzy measures do, aggregation functions may enjoy the following properties.

DEFINITION 1.36. *An aggregation function is called*

- *additive, if for any $\mathbf{x}, \mathbf{y} \in \mathbb{I}^n$ such that $\mathbf{x} + \mathbf{y} \in \mathbb{I}^n$ we have*

$$A(\mathbf{x} + \mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y});$$

  it is then superadditive (subadditive) if the equality is replaced with the symbol $\geq$ ($\leq$);
- *modular, if for any $\mathbf{x}, \mathbf{y} \in \mathbb{I}^n$ we have*

$$A(\mathbf{x} \vee \mathbf{y}) + A(\mathbf{x} \wedge \mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y});$$

  it is supermodular (submodular) if the equality is replaced with the symbol $\geq$ ($\leq$).

The arithmetic mean AM satisfies all the four properties mentioned above, while Π on $[0, 1]$ is supermodular and superadditive, but neither modular nor additive.

6. Fuzzy Integrals based on Aggregation Functions

Fuzzy measures can be seen as a tool useful to resume all the values of a function to a single point. To this aim Sugeno in [**135**] extended such concept to the one of fuzzy integrals. These integrals are built on the real with respect to a fuzzy measure, like Lebesgue integral is built with an ordinary (additive) one. As an ordinary integral can be seen in a certain sense as the average of a function, a fuzzy integral can be seen as an averaging aggregation operator. At the same time the classical notion of measure extends the notion of weight to infinite universes, and the Lebesgue integral on a finite universe coincides with the weighted arithmetic mean. Therefore, the existence of more general notions of measure than the classical additive one, together with the appropriate integrals, offer a new realm of aggregation functions when these integrals are limited to a finite universe. Since additivity is replaced by monotonicity, we deal with monotone measures although the most common name, which we will use, is capacity, introduced by Choquet in [**34**] and resumed in Section 2. The term fuzzy measure introduced by Sugeno is often used in the fuzzy set community.

There are many types of integrals defined with respect to a capacity. The most common ones are the Choquet integral and the Sugeno integral, leading to two interesting classes of aggregation functions, developed in this section. To introduce these arguments we will need some notation first. Let $N := \{1, \dots, n\}$ an set of indexes of n elements.

DEFINITION 1.37. *For any subset $A \subseteq N$, e_S represents the* characteristic vector *of A, i.e. the vector of $\{0,1\}^n$ whose $i-th$ component is 1 if and only if $i \in A$.Geometrically, the characteristic vectors are the 2^n vertices of the hypercube $[0,1]^n$.*

In game theory N represents a group of n players, whose subgroups As indicate *coalitions* among such players. The function v allows to assign to each coalition the proper worth (for example the amount of money earned if the game is played). One can also define the *unanimity game* for $A \subseteq N$ as the game v_B such that $v_B(A) = 1$ if and only if $B \subseteq A$, and 0 otherwise.

DEFINITION 1.38. *A* pseudo-Boolean *function is a function defined as $f : \{0,1\}^n \to \mathbb{R}$.*

Any real valued set function $m : 2^N \to \mathbb{R}$ can be assimilated unambiguously with a pseudo-Boolean function. The correspondence is straightforward: we have

$$f(x) = \sum_{A \subseteq N} m(A) \prod_{i \in A} x_i \prod_{i \notin A} (1 - x_i), \tag{1.33}$$

for $x \in \{0,1\}^n$, and $m(A) = f(e_A)$ for all $A \in N$. In particular, a pseudo-Boolean function that corresponds to a fuzzy measure, is increasing in each variable and fulfils the boundary conditions $f(\mathbf{0}) = 0$ and $f(\mathbf{1}) = 1$, where $\mathbf{0}$ indicates the vector with all null components, while $\mathbf{1}$'s components are all equal to 1. Hammer et al. in [**72**] showed that any pseudo-Boolean function has a unique expression as a multilinear polynomial in n variables:

$$f(x) = \sum_{A \subseteq N} M_m(A) \prod_{i \in A} x_i, \tag{1.34}$$

for $x \in \{0,1\}^n$. The coefficients $M_m(A)$ are the ones of the Möbius transform, defined in (1.16). In game theory, these coefficients are called the *dividends* of the coalitions in *game* m (for further details see, for example, [**108**]). In view of Definition 1.37, equation (1.34) can be seen w.r.t. unanimity games as

$$v(A) = f(e_A) = \sum_{B \subseteq N} M_m(B) \prod_{i \in B} (e_A)_i = \sum_{B \subseteq N} M_m(B) v_B(A). \tag{1.35}$$

Thus, any game v has a canonical representation in terms of unanimity games that determine a linear basis for v (extensions of this topic to general (infinite) spaces of players can be found in [**60**]).

Let now m be a fuzzy measure defined on a discrete set N and let $x_1, \ldots, x_n \in \mathbb{R}$. We are now ready to introduce the following

DEFINITION 1.39. *The (discrete) Choquet integral of a function* $x : N \to \mathbb{R}$*, with respect to a fuzzy measure* m *on* N*, is defined by*

$$Ch_m(x) := \sum_{i=1}^{n} [x_{(i)} - x_{(i-1)}] m(\sigma(i), \ldots, \sigma(n)) \tag{1.36}$$

where, as usual, $x_{(1)} \leq \ldots \leq x_{(n)}$ *and* $x_{(0)} := 0$*. An equivalent formulation of the integral can also be given*

$$Ch_m(x) := \sum_{i=1}^{n} x_{(i)} [m(\sigma(i), \ldots, \sigma(n)) - m(\sigma(i+1), \ldots, \sigma(n))] \tag{1.37}$$

Notice that the link with the Lebesgue integral is strong, since both coincide when the measure m is additive:

$$Ch_m(x) := \sum_{i=1}^{n} m_i x_i.$$

In this sense the Choquet integral can be seen as a generalization of the Lebesgue integral.

DEFINITION 1.40. *The (discrete) Sugeno integral of a function* $x : N \to [0,1]$*, with respect to a fuzzy measure* m *on* N*, is defined by*

$$Su_m(x) := \bigvee_{i=1}^{n} [x_{(i)} \wedge m(\sigma(i), \ldots, \sigma(n))]. \tag{1.38}$$

Given a fuzzy measure m on N, the Choquet and Sugeno integrals can be regarded as aggregation operators defined on $\mathbb{R}^n$ and $[0,1]^n$, respectively. But they are essentially different in nature, since the latter is based on nonlinear operators ($\min$ and $\max$), and the former on usual linear operators.

It can be said that the Choquet integral is suitable for cardinal aggregation (where numbers have a real meaning), while the Sugeno integral seems to be more suitable for ordinal aggregation (where only order makes sense). One of the common properties of the two integrals, instead, is that both compute a kind of distorted average of $x_1, \ldots, x_n$.

Let now introduce another kind of integral useful for its simplicity in aggregation models. Let us consider first the additive fuzzy measure p, derived from the Shapley value Φ_m defined in 1.23:

$$p(S) := \sum_{i \in S} \Phi_m(\{i\}), \tag{1.39}$$

for any set $S \subseteq N$. Then we can define the *Shapley integral* as follows.

DEFINITION 1.41. *The Shapley integral of a function* $x : N \to [0,1]$ *with respect to a fuzzy measure* m *is defined by*

$$Sh_m(x) = \sum_{i \in N} \Phi_m(\{i\}) x_i. \tag{1.40}$$

Thus defined, the Shapley integral is actually a weighted arithmetic mean operator WAM_ω whose weights are the Shapley power indices $\omega_i = \Phi_m(\{i\})$, for all $i = 1, \ldots, n$. Starting from any fuzzy measure, we can define the Shapley additive measure and aggregate by the corresponding weighted arithmetic mean. Note that, contrary to the Choquet and Sugeno integrals, the Shapley integral w.r.t. the fuzzy measure m is not an extension of m. Indeed, for any $S \subseteq N$, we generally have

$$Sh_m(e_S) = \sum_{i \in S} \Phi_m(\{i\}) \neq m(S).$$

More general definitions and properties can be found, for example, in [**62**] and [**63**].

7. The Choquet Integral and its Extensions

In what follows, we give particular attention to the Choquet integral, its extensions and properties. We start recalling that Lovász in [**90**] observed that any $x \in \mathbb{R}^n_+ \setminus \{0\}$ can be written uniquely in the form

$$x = \sum_{i=1}^{k} \lambda_i e_{A_i}, \tag{1.41}$$

with $\lambda_i \geq 0$ for all $i = 1, \ldots, k$ and $\emptyset \neq A_1 \subsetneq \ldots \subsetneq A_k \subseteq N$. Hence any function $f : \{0,1\}^n \to \mathbb{R}$ with $f(0) = 0$ can be extended to $\hat{f} : \mathbb{R}^n_+ \to \mathbb{R}$, with $\hat{f}(0) = 0$ and

$$\hat{f}(x) = \sum_{i=1}^{k} \lambda_i f(e_{A_i}). \tag{1.42}$$

Notice that $\hat{f}$ it is unique and represents an extension of f since $\hat{f} = f$ on $\{0,1\}^n$. Such an extension $\hat{f}$ is called the Lovász extension of the function

f and it benefits of many interesting properties (that on may find, for example, in [**96**]). The most important, for this dissertation, are the following.

THEOREM 1.42. *Let $f : \{0,1\}^n \to \mathbb{R}$, with Lovász extension $\hat{f} : \mathbb{R}^n_+ \to \mathbb{R}$. For any $\sigma \in \Sigma_N$, set of all the permutation of $N = \{1, \ldots, n\}$, and for any $x \in \mathbb{R}^n$ we set*

$$\Psi_\sigma(x) := \sum_{i=1}^{n} x_{\sigma(i)}[f(e_{\{\sigma(i),\ldots,\sigma(n)\}}) - f(e_{\{\sigma(i+1),\ldots,\sigma(n)\}})]. \tag{1.43}$$

Then the following are equivalent:

(1) f is submodular;
(2) $\hat{f}$ is convex;
(3) we have

$$\hat{f}(x) = f(0) + \max_{\sigma \in \Sigma_N} \Psi_\sigma(x), \text{ for } x \in \mathbb{R}^n_+;$$

(4) we have

$$f(x) = f(0) + \max_{\sigma \in \Sigma_N} \Psi_\sigma(x), \text{ for } x \in \{0,1\}^n.$$

In this view we have that the convexity (concavity, linearity) of $\hat{f}$ corresponds to the submodularity (supermodularity, modularity)of f. The proof of the Theorem can be found in [**128**]. From (1.43) we get a useful formulation for the extension of pseudo-Boolean functions, that we give in the following

PROPOSITION 1.43. *Let f a pseudo-Boolean function. Then its Lovász extension $\hat{f}$ is given by*

$$\hat{f}(x) = \sum_{A \subseteq N} M_m(A) \bigwedge_{i \in A} x_i, \tag{1.44}$$

for $x \in \mathbb{R}^n_+$. The coefficients M_m are the Möbius representation of f.

What immediately follows from (1.44) is that, when m is a fuzzy measure on N, the Choquet integral Ch_m on $\mathbb{R}^n_+$ defined in (1.36) is nothing else than the Lovász extension of the pseudo-Boolean function f_m which represents m:

$$Ch_m = \hat{f}_m \tag{1.45}$$

on $\mathbb{R}^n_+$. Thus, the Choquet integral is a piecewise affine function on $[0,1]^n$; moreover it can be seen as the unique liner interpolation the vertices of the hypercube $[0,1]^n$. In fact, the vertices of $[0,1]^n$ correspond to the vectors e_A, so that

$$Ch_m(e_A) = m(A) \quad \text{for all } A \subseteq N.$$

Moreover, we clearly see that Ch_m is an increasing function if and only if m is as well. Proposition 1.43 can be rewritten as follows (see also [**31**]).

PROPOSITION 1.44. *Assume $E \supseteq [0,1]$. Any Choquet integral $Ch_m : E^n \to \mathbb{R}$ can be rewritten as*

$$Ch_m(x) = \sum_{A \subseteq N} M_m(A) \bigwedge_{i \in A} x_i, \tag{1.46}$$

where $x \in E^n$ and M_m is the Möbius transform of the capacity m.

Notice that the integral is unique since the representation M_m of m is. Many other properties of Choquet integral can be found in literature, as the ones listed below.

PROPOSITION 1.45. *The Choquet integral satisfy the following properties (see* **[94]** *for further details).*

(1) The Choquet integral is linear with respect to the capacity: for any two capacities m_1, m_2 on N and any two real numbers $\alpha, \beta \geq 0$ we have

$$Ch_{\alpha m_1 + \beta m_2} = \alpha \cdot Ch_{m_1} + \beta \cdot Ch_{m_2}.$$

(2) The Choquet integral is monotone w.r.t. capacities: for any two capacities m and m' we have that $m \leq m'$ if and only if $Ch_m \leq Ch_{m'}$.

(3) If m is a $0-1$ capacity then

$$Ch_m(\mathbf{x}) = \bigvee_{A \subseteq N : m(A)=1} \bigwedge_{i \in A} x_i, \quad \forall \mathbf{x} \in [0,1]^n.$$

(4) The Choquet integral Ch_m is symmetric if and only if the capacity m is symmetric.

(5) The Choquet integral is invariant under positive affine transformations, that is, for any $c > 0$ and any $a \in \mathbb{R}$,

$$Ch_m(c\mathbf{x} + a\mathbb{1}_{\mathbf{x}}) = c \cdot Ch_m(\mathbf{x}) + a.$$

(6) For any capacity m we have $Ch_m^* = Ch_{m^*}$, i.e. the dual of the Choquet integral with respect to the capacity m is the Choquet integral with respect to the dual of the capacity m.

A property that, in general, a Choquet integral lacks of is that one of additivity, since the corresponding capacity m is not additive itself. However there is a particular situation in which the property of additivity is granted, that is when the integrand vectors satisfy *comonotonic additivity*. There are many definitions of comonotonic additivity (briefly said comonotonicity) as, for example, the following regarding real vectors.

DEFINITION 1.46. *Two vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ are said comonotonic if there exists a permutation σ on N that gives the same order to both vectors, i.e. $x_{\sigma(1)} \leq \ldots \leq x_{\sigma(n)}$ and $x'_{\sigma(1)} \leq \ldots \leq x'_{\sigma(n)}$. Equivalently we can say that there are no couples of indices i, j for which $x_i < x_j$ and $x'_i > x'_j$ at the same time.*

Under such hypothesis we have the following result.

PROPOSITION 1.47. *If* $\mathbf{x}, \mathbf{x}'$ *are comonotonic vectors of* $\mathbb{R}^n_+$ *then, for any capacity* m*,*

$$Ch_m(\mathbf{x} + \mathbf{x}') = Ch_m(\mathbf{x}) + Ch_m(\mathbf{x}'). \tag{1.47}$$

The following gives a characterization of the Choquet integral.

THEOREM 1.48. *Let* $F : \mathbb{R}^n \to \mathbb{R}$ *be a given function. Then there exists a unique capacity* m *such that* $F = Ch_m$ *if and only if the function* F *satisfies the following properties:*

(1) comonotonic additivity;
(2) nondecreasing monotonicity;
(3) boundary conditions, i.e. $F(\mathbf{0}) = 0, F(\mathbf{1}) = 1$.

Moreover, m *is defined through* F *as* $m(A) := F(\mathbb{1}_A)$ *for any* $A \subseteq N$.

This result was showed by De Campos in [**35**], assuming in addition positive homogeneity, condition that can be deduced from hypothesis (1) and (2). The proof in the continuous case is due to Schmeidler and can be found in [**121**].

We now present the connections between the Choquet integral and the most common aggregation functions introduced in the previous Section.

PROPOSITION 1.49. *Let* m *be a capacity and consider* $\mathbb{I} = \mathbb{R}$*. The following holds*

(1) $Ch_m = Min$ *if and only if* $m = m_{Min}$ *is the minimal capacity; in the same way we can state that* $Ch_m = Max$ *if and only if* $m = m_{Max}$*;*
(2) $Ch_m = OS_k$*, the* $k-$*th order statistic, if and only if the capacity* m *is the threshold measure* τ_{n-k+1}*;*
(3) $Ch_m = P_k$*, the* $k-$*th projection, if and only if* m *is the Dirac measure* δ_k*;*
(4) $Ch_m = WAM_{\mathbf{w}}$ *if and only if* m *is additive, with* $w_i = m(\{i\})$ *for all* $i \in N$*;*
(5) $Ch_m = OWA_{\mathbf{w}}$ *if and only if* m *is symmetric, with weights* $w_i = m(A_{n-i+1}) - m(A_{n-i})$ *for all* $i = 2, \ldots, n$ *and* $w_1 = 1 - \sum_{i=2}^{n} w_i$*; any subset* A_i *of* Ω *is such that* $|A_i| = i$ *and its measure* $m(A) = \sum_{j=0}^{i-1} w_{n-j}$.

We already mentioned the problem of the complexity of the capacity m: one requires, in fact, $2^n - 2$ information to know it completely. To reduce this problem significatively one can make use of capacities owning the property of $k-$additivity, like proposed by Grabisch in [**65**]. The Choquet integral considerably simplifies in this case, in particular, when the underlying capacity is $2-$additive (or equivalently, m is a $2-$order fuzzy measure), we have the following result, du to Marichal [**93**].

THEOREM 1.50. *Let m a $2-$order fuzzy measure on N. Then the best weighted arithmetic mean that minimizes*

$$\int_{[0,1]^n} [Ch_m(x) - WAM_\omega(x)]^2 \, dx \tag{1.48}$$

is given by the Shapley integral Sh_m. Moreover, if $E \supseteq [0,1]$, we have

$$Ch_m(x) = Sh_m(x) - \frac{1}{2} \sum_{\{i,j\} \subseteq N} I_m(\{i,j\})[x_i \vee x_j - x_i \wedge x_j], \tag{1.49}$$

for $x \in E^n$.

Equation (1.49) shows that the Choquet integral can be decomposed in a linear part, represented by Sh_m, and a non-linear part divided, in turn, into 2 components: the one considering positive indices $I(\{i,j\})$ and the second one consisting in the negative indices. The positive part, for which $I(\{i,j\}) \in [0,1]$ implies a complementary behavior, that means that both criteria need to be satisfied to have a better global score; the negative components $I(\{i,j\}) \in [-1,0]$ describe negative interaction between the couples $\{i,j\}$, that means that the components are substitutable, i.e. the satisfaction of either i or j is sufficient to have a significant effect on the global score. A more specific result in this direction is showed by Grabisch in [**64**].

THEOREM 1.51. *Let m be a $2-$order fuzzy measure on N and assume $E \supseteq [0,1]$. Then we have*

$$\begin{aligned} Ch_m(x) = & \sum_{i \in N} \Big(\Phi_m(\{i\}) - \frac{1}{2} \sum_{j \in N \setminus \{i\}} |I_m(\{i,j\})| \Big) x_i \\ & + \sum_{I_m(\{i,j\}) \geq 0} I_m(\{i,j\})(x_i \vee x_j) - \sum_{I_m(\{i,j\}) \leq 0} I_m(\{i,j\})(x_i \wedge x_j) \end{aligned} \tag{1.50}$$

for all $x \in E^n$. Moreover, we have $\Phi_m(\{i\}) - \frac{1}{2} \sum_{j \in N \setminus \{i\}} |I_m(\{i,j\})| \geq 0$ for all $i \in N$.

This decomposition emphasizes the role of the positive and negative components: a Choquet integral with strong positive (negative) component will be strongly conjunctive (disjunctive); if the values $I_m(\{i,j\})$ are low the integral will be, with good approximation, linear. In this view it is possible to write the integral as the sum of two components, namely

$$Ch_m(x) = Ch_{m^+}(x) + Ch_{m^-}(x), \tag{1.51}$$

where m^+ and m^- are defined through their interaction representation

$$I_{m^+}(T) = \max\{I_m(T), 0\} \quad I_{m^-}(T) = \min\{I_m(T), 0\}.$$

Due to linearity of I on set functions we have $m = m^+ + m^-$ and equation (1.51) holds true. More in particular it can be rewritten as

$$Ch_m(x) = \sum_{T \subseteq N} M_{m^+}(T) \bigwedge_{i \in T} x_i + \sum_{T \subseteq N} \check{M}_{m^-}(T) \bigvee_{i \in T} x_i, \tag{1.52}$$

where M_{m^+} represents the Möbius transform of the positive component m^+ while $\check{M}_{m^-}$ is the co-Möbius representation of m^-.

A more general extension of formula (1.50) can be given for any $k-$additive measure m. Grabisch in [**64**] reports an example of order 3. For such measures it seems that an interpretation similar to the previous one can no longer be given.

Many other extensions of the Choquet integral, that we do not consider in this work, have been presented in literature in the last years, see for example the *concave integral* proposed by Lehrer [**88**], or the *universal integral* by Klement et al. in [**83**].

CHAPTER 2

Copulas: an overview

Copulas are specific aggregation operators, that are applied to aggregate marginal distribution functions into an output joint distribution function. Nelsen in [**106**] referred to copulas as "functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions" and as "distribution functions whose one-dimensional margins are uniform". But neither of these statements is a definition, hence we will devote this Chapter to giving a precise definition of copulas and to examining some of their elementary properties.

1. Basic Concepts and Definitions

We first define subcopulas as a certain class of grounded $2-$increasing functions with margins; then we define copulas as subcopulas with domain $I^2 = [0,1]^2$.

DEFINITION 2.1. *A two-dimensional subcopula (or $2-$subcopula, or briefly, a subcopula) is a function C' with the following properties*

(1) $Dom(C') = S_1 \times S_2$ where S_1 and S_2 are subset of I containing 0;
(2) C' is grounded, namely $C'(0,v) = C'(u,0) = 0$ for all $u, v \in I$, and $2-$increasing, i.e. increasing in each variable;
(3) $C'(u,1) = u$ and $C'(1,v) = v$, for all $u, v \in I$.

DEFINITION 2.2. *A two-dimensional copula (or briefly a copula) is a $2-$subcopula C whose domain is I^2. Equivalently, a copula is a function C from I^2 to I with the following properties:*

(1) For every $u, v \in I$, $C(u,0) = C(0,v) = 0$ and $C(u,1) = u$ and $C(1,v) = v$;
(2) 2. For every $u_1, u_2, v_1, v_2 \in I$ such that $u_1 < u_2$ and $v_1 < v_2$,

$$V_C([\mathbf{u}, \mathbf{v}]) = C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

The value $V_C([\mathbf{u}, \mathbf{v}])$ is the volume of the $2-$copula over the set $[u_1, u_2] \times [v_1, v_2]$.

We give now some general properties of copulas, for reference see [**106**].

PROPOSITION 2.3. *The following hold for any copula C.*

- *C is increasing in each argument;*
- *C is Lipschitz (and hence uniformly) continuous;*
- *for $i = 1, 2$ $\partial_i C$ exists a.e. and $0 \leq \partial_i C \leq 1$;*

- *The functions $t \to \partial_1 C(u,t)$ and $t \to \partial_2 C(t,v)$ are defined and increasing a.e. on I.*

There are three distinguished copulas, namely

$$W(u,v) = \max(u+v-1,0); \quad M(u,v) = \min(u,v); \quad \Pi(u,v) = uv, \tag{2.1}$$

see Figure 2.1. Copulas M and W are called the *Fréchet-Hoeffding* upper and lower bounds, respectively, since for any copula C and any $u,v \in I$ we have

$$W(u,v) \leq C(u,v) \leq M(u,v). \tag{2.2}$$

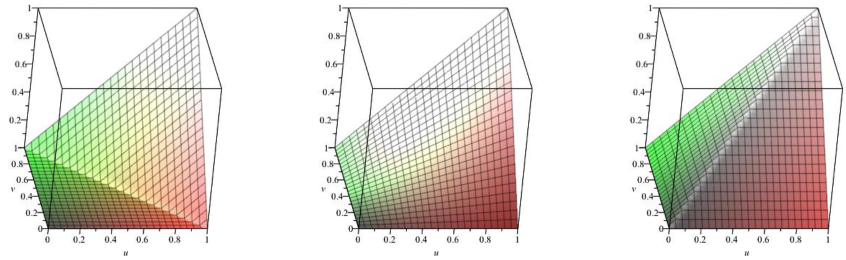

FIGURE 2.1. Copulas W, Π, M respectively

One of the most important results about copulas, that one that links them with the concepts of joint distribution function, is the following due to Sklar.

THEOREM 2.4. *Let H be a joint distribution function with margins F and G. Then there exists a copula C such that for all $x,y \in \mathbb{R}$,*

$$H(x,y) = C(F(x),G(y)). \tag{2.3}$$

If F and G are continuous, then C is unique; otherwise, C is uniquely determined on $RanF \times RanG$. Conversely, if C is a copula and F and G are distribution functions, then the function H defined by (2.3) *is a joint distribution function with margins F and G.*

LEMMA 2.5. *Let H be a joint distribution function with margins F and G. Then there exists a unique subcopula C' such that*

(1) $DomC' = RanF \times RanG$;
(2) for all $x,y \in \mathbb{R}$, $H(x,y) = C'(F(x),G(y))$.

DEFINITION 2.6. *Let F be a distribution function. Then a quasi-inverse of F is any function $F^{(-1)}$ with domain I such that*

- *if t is in $RanF$, then $F^{(-1)}(t)$ is any number x in $\mathbb{R}$ such that $F(x) = t$, i.e., for all $t \in RanF$,*

$$F(F^{(-1)}(t)) = t;$$

- *if $t \notin RanF$, then*

$$F^{(-1)}(t) = \inf\{x | F(x) \geq t\} = \sup\{x | F(x) \leq t\}.$$

If F is strictly increasing, then it has but a single quasi-inverse, which is of course the ordinary inverse, for which we use the customary notation F^{-1}.

COROLLARY 2.7. *Let H, F, G, and C' be as in Lemma 2.5, and let $F^{(-1)}$ and $G^{(-1)}$ be quasi-inverses of F and G, respectively. Then for any (u, v) in $DomC$,*

$$C'(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)).$$

When F and G are continuous, the above result holds for copulas as well and provides a method of constructing copulas from joint distribution functions.

LEMMA 2.8. *Let C' be a subcopula. Then there exists a copula C such that $C(u, v) = C'(u, v)$ for all $(u, v) \in DomC'$; i.e., any subcopula can be extended to a copula. The extension is generally non-unique.*

THEOREM 2.9. *Let X and Y be continuous random variables. Then X and Y are independent if and only if $C_{XY} = \Pi$.*

THEOREM 2.10. *Let X and Y be continuous random variables with copula C_{XY}. If a and b are strictly increasing on $RanX$ and $RanY$, respectively, then $C_{a(X)b(Y)} = C_{XY}$. Thus C_{XY} is invariant under strictly increasing transformations of X and Y.*

THEOREM 2.11. *Let X and Y be continuous random variables with copula C_{XY}. Let a and b be strictly monotone on $RanX$ and $RanY$, respectively. Then*

(1) if a is strictly increasing and b is strictly decreasing, then

$$C_{a(X)b(Y)}(u, v) = u - C_{XY}(u, 1 - v);$$

(2) if a is strictly decreasing and b is strictly increasing, then

$$C_{a(X)b(Y)}(u, v) = v - C_{XY}(1 - u, v);$$

(3) if a and b are both strictly decreasing, then

$$C_{a(X)b(Y)}(u, v) = u + v - 1 + C_{XY}(1 - u, 1 - v).$$

A copula is a continuous function $C : [0, 1]^n \to [0, 1]$, but is not necessarily absolutely continuous. Any copula C can, in fact, be written as $C(u, v) = A_C(u, v) + S_C(u, v)$, where

$$A_C(u, v) = \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} C(s, t)\, dt ds, \quad S_C(u, v) = C(u, v) - A_C(u, v).$$

A copula C coinciding with A_C ($S_C = 0$) is then absolutely continuous, while if $C = S_C$ ($A_C = 0$) the copula is said singular. Otherwise it has a singular component S_C and an absolutely continuous one A_C. The Fréchet-Hoeffding bounds W and M are singular copulas: the mass of M is concentrated on the line $u = v$ while W is distributed on the line $u + v = 1$; on the other hand, the independence copula Π is absolutely continuous.

In many applications, the random variables of interest represent the lifetimes of individuals or objects in some population. The probability of an individual living or surviving beyond time x is given by the survival function (or reliability function) $F(x) = \mathbb{P}(X > x) = 1 - F(x)$, where as before, F denotes the distribution function of X. For a pair (X, Y) of random variables with joint distribution function H, the joint survival function is given by $H(x, y) = \mathbb{P}(X > x, Y > y)$. The margins of H are the functions are the univariate survival functions $\overline{F}$ and $\overline{G}$, respectively. We are going to show the relationship between univariate and joint survival functions. Suppose that C is the copula between the variables X and Y. Then

$$\begin{aligned}\overline{H}(x,y) &= 1 - F(x) - G(x) + H(x,y)\\ &= \overline{F}(x) + \overline{G}(x) - 1 + C(F(x), G(x))\\ &= \overline{F}(x) + \overline{G}(x) - 1 + C(1 - \overline{F}(x), 1 - \overline{G}(x))\end{aligned}$$

so we can define $\widehat{C} : [0,1]^2 \to [0,1]$ by

$$\widehat{C}(u,v) = u + v - 1 + C(1-u, 1-v), \tag{2.4}$$

then we have $\overline{H}(x,y) = \widehat{C}(\overline{F}(x), \overline{G}(x))$.

Two other functions closely related to copulas (and survival copulas) are the *dual of a copula* and the *co-copula* (Schweizer and Sklar 1983). The dual of a copula C is the function $\widetilde{C}$ defined by

$$\widetilde{C}(u,v) = u + v - C(u,v) \tag{2.5}$$

and the co-copula is the function C^* defined by

$$C^*(u,v) = 1 - C(1-u, 1-v). \tag{2.6}$$

Neither of these is a copula, but when C is the copula of a pair of random variables X and Y, the dual of the copula and the co-copula each express a probability of an event involving X and Y. More in details we know that

$$\begin{aligned}C(F(x), G(y)) &= \mathbb{P}(X \leq x, Y \leq y),\\ \widetilde{C}(\overline{F}(x), \overline{G}(y)) &= \mathbb{P}(X > x, Y > y),\end{aligned}$$

and we also have

$$\begin{aligned}\widetilde{C}(F(x), G(y)) &= \mathbb{P}(X \leq x \vee Y \leq y),\\ C^*(\overline{F}(x), \overline{G}(y)) &= \mathbb{P}(X > x \vee Y > y).\end{aligned}$$

Extension to generic dimension n can be given as we are going to report below.

DEFINITION 2.12. *An $n-$dimensional subcopula (or $n-$subcopula) is a function C' with the following properties:*

(1) $DomC' = S_1 \times \ldots \times S_n$, where each S_i is a subset of I^n containing 0 and 1;
(2) C' is grounded and $n-$increasing;

(3) C' *has (one-dimensional) margins* C'_i, $i = 1, \dots, n$, *which satisfy* $C'_i(u) = u$ *for all* $u \in S_i$.

Note that for every $\mathbf{u} \in DomC'$, $0 \le C'(\mathbf{u}) \le 1$, *so that* $RanC'$ *is also a subset of* I.

DEFINITION 2.13. *An* $n-$*dimensional copula (or simply an* $n-$*copula) is an* $n-$*subcopula* C *whose domain is* I^n. *Equivalently, an* $n-$*copula is a function* $C : I^n \to I$ *with the following properties:*

(1) *for every* $\mathbf{u} \in I^n$, $C(\mathbf{u}) = 0$ *if at least one coordinate of* u *is* 0, *and if all coordinates of* u *are* 1 *except* u_k, *then* $C(\mathbf{u}) = u_k$;
(2) *for every* $\mathbf{a}, \mathbf{b} \in I^n$ *such that* $\mathbf{a} \le \mathbf{b}$, *the* $n-$*volume* $V_C([\mathbf{a}, \mathbf{b}]) \ge 0$.

It is easy to show that for any $n-$*copula* C *with* $n \ge 3$, *each* $k-$*margin of* C *is a* $k-$*copula,* $2 \le k \le n$.

The main properties of the copulas as well as Sklar's Theorem are still valid in dimension n. Any $n-$dimensional copula C satisfy the Fréchet-Hoeffding upper and lower bounds, so takes values between

$$W(u_1, \dots, u_n) = \max(u_1 + \dots + u_n - n + 1, 0) \tag{2.7}$$

and

$$M(u_1, \dots, u_n) = \min(u_1, \dots, u_n).$$

It is important to notice that, for $n > 2$, $W(u_1, \dots, u_n)$ is no longer a copula.

For the $2-$dimensional case, the Fréchet-Hoeffding bounds inequality introduced in (2.2), suggests a partial order on the set of copulas.

DEFINITION 2.14. *If* C_1 *and* C_2 *are copulas, we say that* C_1 *is smaller than* C_2, *and write* $C_1 \prec C_2$ *(or* $C_2 \succ C_1$*), if* $C_1(u, v) \le C_2(u, v)$ *for all* $u, v \in \mathbb{I}$.

Recalling that $W(u, v) \le C(u, v) \le M(u, v)$ for every copula C and all $u, v \in \mathbb{I}$, the lower bound copula W is smaller than every copula, and the upper bound copula M is larger than every copula. This point-wise partial ordering of the set of copulas is called the *concordance ordering*. It is a partial order rather than a total order because not every pair of copulas is comparable. It assumes importance in the study of the dependence of random variables, through the use of their connecting copulas. A similar definition can be given for the multi-dimensional case. We delay it discussion to next Section, in which we introduce the main ideas of dependence among random variables.

1.1. Archimedean Copulas. An important class of copulas is that one of Archimedean copulas. This class has a wide range of application, due to the great variety of copulas belonging to it, the ease with which they can be built and the many properties they enjoy. We encountered the particular case of independence between variables, whose copula is expressed by the product copula Π. From a practical point of view, we are interested in similar simple expressions useful for the construction of copulas, like $\phi(C(u,v)) = \phi(u) + \phi(v)$, for some function ϕ. So we need to find an appropriate inverse $\phi^{[-1]}$, with opportune properties, that solves

$$C(u,v) = \phi^{[-1]}(\phi(u) + \phi(v)). \tag{2.8}$$

DEFINITION 2.15. *Let $\phi : I \to [0,\infty]$ be continuous, strictly decreasing and with $\phi(1) = 0$. The* pseudo-inverse *of ϕ is the function $\phi^{[-1]} : [0,\infty] \to I$ given by*

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & \text{if } 0 \leq t \leq \phi(0); \\ 0, & \text{if } \phi(0) \leq t \leq \infty. \end{cases} \tag{2.9}$$

Notice that $\phi(\phi^{[-1]}(t)) = \min(t, \phi(0))$ and if $\phi(0) = \infty$ then $\phi^{[-1]}(t) = \phi^{-1}(t)$.

A function C defined as in (2.8) is, indeed, a copula, since the following holds.

LEMMA 2.16. *Let ϕ and $\phi^{[-1]}$ as in Definition 2.15 and let $C : I^2 \to I$ be a function satisfying Eq.* (2.8)*. Then C is $2-$increasing and satisfies the Fréchet-Hoeffding boundary conditions, so it is a copula.*

Moreover, it can be given a characterization as follows.

THEOREM 2.17. *Let ϕ and $\phi^{[-1]}$ and C as in the previous Lemma. Then C is a copula if and only if ϕ is convex.*

Some important properties of Archimedean copulas are the following.

THEOREM 2.18. *Let C be an Archimedean copula with generator ϕ. Then:*

(1) C is symmetric, i.e. $C(u,v) = C(v,u)$ for all $u, v \in I$;
(2) C is associative, namely $C(C(u,v),w) = C(u, C(v,w))$ for all $u, v, w \in I$;
(3) $c\phi$ is a generator of C, for any constant $c > 0$.

A first simple example of Archimedean copulas is given by the independence copula Π. Consider $\phi = -\ln t$, so $\phi^{[-1]} = \exp(-t)$ and with straightforward calculation we get, from (2.8),

$$C(u,v) = \phi^{[-1]}(\phi(u) + \phi(v)) = \exp(\ln u + \ln v) = uv = \Pi(u,v).$$

In a similar way one can prove that also the minimal copula W is Archimedean, while M is not. Other important families of Archimedean copulas the ones attributed to Clayton, Ali-Mikhail-Haq, Frank and Gumbel. These classes of copulas are called one-parameters families, since all the copulas belonging to any of this families can be obtained by changing the value of the generating parameter. Consider, for example, the copulas from the Ali-Mikhail-Haq family (see Figure 2.2), namely those one that can be written as

$$C_\theta(u,v) = \frac{uv}{1-\theta(1-u)(1-v)}, \tag{2.10}$$

with parameter θ ranging in $[-1,1]$ and generator given by

$$\ln(1-\theta(1-t)/t).$$

The independence copula belongs to this family, since can be obtained substituting to θ the value 0. Many other examples and interesting properties won't be discussed in this paper, but can be found, for example, in [**106**].

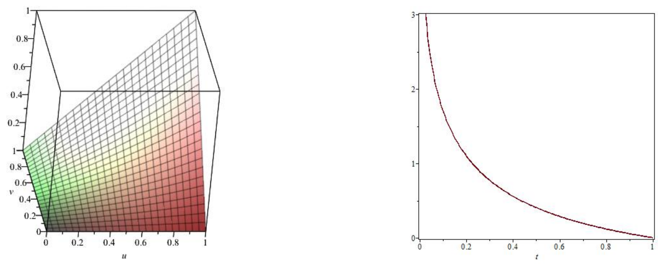

FIGURE 2.2. Copula from the Ali-Mikhail-Haq family, with parameter $\theta = -0.5$ and its support

1.2. Copulas for capacities. It is interesting to highlight connections between copulas and non-additive measures, especially with capacities. As for additive probabilities, copulas for non-additive measures can be defined, with just some minor requirements.

Consider, for instance, the extended line $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ and a capacity μ on $(\overline{\mathbb{R}}^d, \mathcal{B}(\overline{\mathbb{R}}^d))$. Let μ_i the $i-$th projection of μ: for $S \in \mathcal{B}(\overline{\mathbb{R}})$,

$$\mu_i(S) = \mu(\overline{\mathbb{R}} \times \overline{\mathbb{R}} \times \ldots \times S \times \ldots \overline{\mathbb{R}}),$$

where S is in the $i-$th coordinate of the vector. It can be defined, as natural, the distribution function $F_\mu : \overline{\mathbb{R}}^d \to \mathbb{R}$ associated to μ as follows:

$$F_\mu(x_1, \ldots, x_d) = \mu([-\infty, x_1] \times \ldots \times [-\infty, x_d]).$$

Marginal components are defined as well, namely $F_{\mu_i} : \overline{\mathbb{R}} \to \mathbb{R}$ such that

$$F_{\mu_i}(x) = \mu_i([-\infty, x]).$$

It is easy to see that F_μ is increasing, since the capacity μ, for definition, is monotone. In general F_μ is not right continuous and, of course, does not characterize μ on the whole Borel $\sigma-$field, since even the distribution function of a finitely additive probability measure in general does not have these properties. Some properties can, in any case, be attributed to μ, as the following.

DEFINITION 2.19. *Let*

$$\Delta_{s_i=x_i}^{y_i} f(s_1,\dots,s_i,\dots,s_d) = f(s_1,\dots,y_i,\dots,s_d) - f(s_1,\dots,x_i,\dots,s_d).$$

A function $F : \overline{\mathbb{R}}^d \to \mathbb{R}$ *is called* $n-$increasing $(n \le d)$ *if*

$$\Delta_{s_{i_1}=x_{i_1}}^{y_{i_1}} \cdots \Delta_{s_{i_n}=x_{i_n}}^{y_{i_n}} f(\dots,s_{i_1},\dots,s_{i_n},\dots) \ge 0$$

for any possible values of the indices $i_1,\dots,i_n$.

Any $d-$variate distribution function associate with a finitely additive probability measure is n-increasing for all $n \le d$. This is true also for $d-$monotone capacities. In the case of probability measures the result is a consequence of the fact that the probability of any $d-$dimensional rectangle in $\overline{\mathbb{R}}^d$ is nonnegative and this probability can be expressed as the multiple finite difference of the distribution function. The same procedure cannot be applied for capacities, due to lack of additivity, but the definition of $d-$monotonicity gives the result, resumed as follows.

LEMMA 2.20. *If* μ *is* $d-$*monotone, then* F_μ *is* $n-$*increasing for every* $n \le d$.

This result is due to Scarsini, see [**120**]. In particular for the case $d = 2$, the distribution function of any convex capacity is increasing and $2-$increasing. Other important properties are the following.

COROLLARY 2.21. *If* μ *is a* $d-$*monotone capacity on* $(\overline{\mathbb{R}}^d, \mathcal{B}(\overline{\mathbb{R}}^d))$*, then there exists a finitely additive probability measure* ν *on* $(\overline{\mathbb{R}}^d, \mathcal{B}(\overline{\mathbb{R}}^d))$ *such that* $F_\mu = F_\nu$.

THEOREM 2.22. *If* μ *is convex then* F_μ *satisfies the Fréchet-Hoeffding upper and lower bounds (see* (2.2)*), namely*

$$\max(F_{\mu_1}(x_1) + F_{\mu_2}(x_2) - 1, 0) \le F_\mu(x_1,x_2) \le \min(F_{\mu_1}(x_1), F_{\mu_2}(x_2)).$$

The same result holds for $d-$*dimensional bounds, like in* (2.7).

These two results allow us to state the following Theorem.

THEOREM 2.23. *Let* μ *be a convex capacity on* $(\overline{\mathbb{R}}^d, \mathcal{B}(\overline{\mathbb{R}}^d))$*. Then there exists a function* $C_\mu : I^d \to I$*, called a* generalized copula *of* μ*, such that*

(1) $F_\mu(x_1,\dots,x_d) = C_\mu(F_{\mu_1}(x_1),\dots,F_{\mu_d}(x_d))$*;*
(2) $C_\mu(x_1,\dots,x_d) = 0$ *if* $x_i = 0$ *for at least one index* $i \in \{1,\dots,n\}$*;*
(3) $C_\mu(1,\dots,1,x_i,1,\dots,1) = x_i$*;*
(4) C_μ *is increasing.*

The proof of this statement is left to [**120**]. Notice that, when the underlying measure μ is a probability measure, then the generalized copula coincide with the usual copula, defined in 2.2. If the measure μ is $d-$monotone, then the generalized copula C_μ is $n-$increasing for all $n \le d$. Generally, it is enough to assume convexity of the capacity to establish the existence of a function that relates the joint distribution function to its margins, but

$d-$monotonicity is required for this function to have all the analytic properties of a copula.

2. Dependence

The concepts of dependence are needed in the analysis of multivariate models. The literature is rich of such models, so we list some of the most important concept of dependence that we have found useful for our study. These are:

- the *positive quadrant dependence* (PQD) and the *concordance ordering*, basic for copulas to determine wherever a multivariate parameter is a dependence parameter;
- the *stochastic increasing* positive dependence (SI);
- the TP_2 dependence, necessary for constructing families of closed-form copulas with wide range of dependence;
- the *tail dependence* for extreme values copulas;
- *Kendall's tau*, *Spearman's rho* and *Gini's gamma*, as functions to study concordance among variables from the analysis of their connecting copulas.

We will consider principally dependence concepts for bivariate distributions. For reference about this literature see, for example, [**10**], [**76**], and [**106**]. We start with the following

DEFINITION 2.24 (Lehmann [**87**]). *Let* $\mathbf{X} = (X_1, X_2)$ *a bivariate random vector with cdf* F*. We say that* $\mathbf{X}$ *(or* F*) is positive quadrant dependent (PQD) if*

$$\mathbb{P}(X_1 > x_1, X_2 > x_2) \geq \mathbb{P}(X_1 > x_1)\mathbb{P}(X_2 > x_2) \quad \forall x_1, x_2 \in \mathbb{R}. \quad (2.11)$$

Condition (2.11) *is equivalent to*

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) \geq \mathbb{P}(X_1 \leq x_1)\mathbb{P}(X_2 \leq x_2) \quad \forall x_1, x_2 \in \mathbb{R}. \quad (2.12)$$

The reason why this consists in a positive dependence concept is that X_1 and X_2 are more likely to be large (or small) together than two other variables with same marginal laws but independent. If the inequalities in (2.11) and (2.12) are reversed we will talk about *negative quadrant dependence* (NQD). A similar definition can be given for multidimensional random vectors.

DEFINITION 2.25. *Let* $\mathbf{X} = (X_1, \ldots, X_n)$ *a multivariate random vector with cdf* F*. We say that* $\mathbf{X}$ *(or* F*) is* positive upper orthant dependent *(PUOD) if*

$$\mathbb{P}(X_1 > x_1, \ldots, X_n > x_n) \geq \prod_{i=1}^{n} \mathbb{P}(X_i > x_i) \quad \forall \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}, \quad (2.13)$$

and that is positive lower orthant dependent *(PLOD) if*

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) \geq \prod_{i=1}^{n} \mathbb{P}(X_i \leq x_i) \quad \forall \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}. \tag{2.14}$$

If both conditions (2.13) *and* (2.14) *hold, then* $\mathbf{X}$ *(F) is said* positive orthant dependent *(POD). Notice that, in the multivariate case, these two expressions are not necessarily equivalent.*

If the inequality are reversed, we can state, in a similar way, the concepts of *negative upper orthant dependence* (NUOD), *negative lower orthant dependence* (NLOD), and *negative orthant dependence*.

The Definitions given above can be restated in terms of copulas. For example PQD condition becomes as follows.

DEFINITION 2.26. *Consider two random variables X_1, X_2 with continuous marginal distributions G_1, G_2, cdf F, and connecting copula C. We say that C is PQD (as X_1, X_2 are) if*

$$C(G_1(u), G_2(v)) \geq \Pi(G_1(u), G_2(v)), \tag{2.15}$$

for all $(u, v) \in \mathbb{I}^2$. If the inequality in (2.15) *is reversed, then the copula C is said NQD.*

Similar arguments can be used to define PLOD, PUOD, and POD conditions in terms of multivariate copulas (and NLOD, NUOD, and NOD too). According to the definition of concordance ordering for copulas, given in 2.14, we can make comparisons between couples of random variables to establish their degree of concordance on the basis of the degree of concordance expressed by their connecting copulas respectively. For example a $2-$copula C_1 is more PQD than another $2-$copula C_2 if $C_1(u, v) \geq C_2(u, v)$ for all $(u, v) \in [0, 1]^2$. In dimension n, C_1 will be more PLOD than C_2 if $C_1(\mathbf{u}) \geq C_2(\mathbf{u})$, and more PUOD if $\overline{C}_1(\mathbf{u}) \geq \overline{C}_2(\mathbf{u})$, for every $\mathbf{u} \in [0, 1]^n$ (then POD if both hold).

An other concepts of dependence is the following, regarding tail monotonicity of copulas.

DEFINITION 2.27. *Let $\mathbf{X} = (X_1, X_2)$ a bivariate random vector with cdf F. We say that X_2 is* stochastically increasing *in X_1, or the conditional distribution $F_{2|1}$ is stochastically increasing SI($X_2|X_1$), if*

$$\mathbb{P}(X_2 > x_2 \,|\, X_1 = x_1) = 1 - F_{2|1}(x_2|x_1) \tag{2.16}$$

is an increasing function of x_1, for all $x_2 \in \mathbb{R}$. By reversing the direction of the monotonicity in (2.16)*, the* stochastically decreasing *(SD) condition results.*

In terms of copulas, this result can be restated as follows.

DEFINITION 2.28. *Let X_1 and X_2 be continuous random variables with copula C. Then SI($X_2|X_1$) holds if and only if for any $v \in [0, 1]$ and for* *almost all u, $\frac{\partial}{\partial u}C(u,v)$ is non-increasing in u or, equivalently, if C is a concave function of u.*

Other important definitions are that ones of *right tail increasing* (RTI), *right tail decreasing* (RTD) or *totally positivity of order 2* (TP2), with some properties defining connections among them, can be found in [**106**].

Dependance is also modeled by some concept better known as *measure of concordance* between random variables. Given a pair of random variables (X, Y), we say that two observation (x_1, y_1) and (x_2, y_2) from the pair are concordant if $(x_1 - x_2)(y_1 - y_2) > 0$, discordant if the inequality is reversed. This means that the values of one of the random variables tends to be big or small in the same way as the values of the other variable do. We give for first the following

DEFINITION 2.29. *A numeric measure k of association between two continuous random variables X and Y whose copula is C is a measure of concordance if it satisfies the following properties (we write $k_{X,Y}$ or k_C):*

- *k is defined for every pair of continuous random variables;*
- *$1 \leq k_{X,Y} \leq 1$, $k_{X,X} = 1$, and $k_{X,-X} = 1$;*
- *$k_{X,Y} = k_{Y,X}$;*
- *if X, Y are independent then $k_{X,Y} = k_\Pi = 0$;*
- *$k_{-X,Y} = k_{X,-Y} = -k_{X,Y}$;*
- *if $C_1 \prec C_2$ then $k_{C_1} \leq k_{C_2}$;*
- *if $\{(X_n, Y_n)\}$ is a sequence of continuous random variables with copulas C_n, and if $\{C_n\}$ converges pointwise to C, then we have that $\lim_{n\to\infty} k_{C_n} = k_C$.*

Now we introduce some of the most known examples of measures of concordance.

DEFINITION 2.30. *Let (X_1, Y_1) and (X_2, Y_2) be independent and identically distributed random vectors with joint distribution F.* Kendall's tau *measure of concordance is defined as*

$$\tau_{X,Y} = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0]. \quad (2.17)$$

For independent vectors of random variables not sharing the same joint distributions, one can define a "concordance function" Q as follows. If C_1, C_2 are the connecting copulas for the couples $(X_1, Y_1), (X_2, Y_2)$ then we can write

$$Q = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

and obtain the following integral representation for Q:

$$Q = Q(C_1, C_2) = 4 \int_{\mathbb{I}^2} C_2(u, v) dC_1(u, v) - 1. \quad (2.18)$$

Such a representation can be used to express Kendall's tau by means of the copula C:

$$\tau_{X,Y} = \tau_C = Q(C,C) = 4\int_{\mathbb{I}^2} C(u,v)dC(u,v) - 1. \tag{2.19}$$

It is interesting to notice that this value, although expressed in terms of integrals, can be used to compute dependance also for copulas that contains a singular component. To do this we just need to rewrite the integral in (2.19) as

$$\tau_C = 1 - 4\int_{\mathbb{I}^2} \frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v)\, dudv. \tag{2.20}$$

The proof of this result rises from an application of integration by parts. An interesting example is given in [**106**] of Kendall's tau for the copulas of Marshall-Olkin family (see section 5). In this particular case the measure of concordance coincides with the measure of the singular component of the copula.

Another important measures are given by *Spearman's rho* and *Gini's Gamma*. Both of them can be defined by means of function Q. Namely $\rho = 3Q(C,\Pi)$ while $\gamma = Q(C,M) + Q(C,W)$. In a sense, Spearman's rho measures a concordance relationship or "distance" between the distribution of X and Y as represented by their copula C and independence as represented by the copula Π. On the other hand, Gini's gamma measures a concordance relationship or "distance" between C and monotone dependence, as represented by the copulas M and W.

Notice that, for Archimedean copulas, this expressions can be written by means of their generators, and assume a simpler form. For example, given an Archimedean copula C with generator ϕ, Kendall's Tau can be written as

$$T_C = 1 + 4\int_0^1 \frac{\phi(t)}{\phi'(t)}dt. \tag{2.21}$$

For further details and properties of dependance measures see [**106**].

3. Methods of constructing copulas

Nelsen in [**106**] presents several general methods of constructing bivariate copulas. By means of Sklar's theorem one can produce copulas directly from joint distribution functions. Using geometric methods, one may construct singular copulas whose support lies in a specified set and copulas with sections given by simple functions such as polynomials. He also discusses some geometrical procedures that produce copulas known as ordinal sums, shuffles of M, and convex sums. In the algebraic method, he constructs copulas from relationships involving the bivariate and marginal distributions functions. In this Section we briefly report some of these methods and provide examples for them.

We start introducing the "inversion method", based on Sklar's inversion theorem presented in 2.7. Given a bivariate distribution function H with continuous margins G_1 and G_2, we can obtain a copula by "inverting" via the expression 2.7:

$$C(u,v) = H(G_1^{(-1)}(u), G_2^{(-1)}(v)).$$

With this copula, we can construct new bivariate distributions with arbitrary margins, say

$$H'(x,y) = C(G_1'(x), G_2'(y)).$$

Of course, this can be done equally as well using survival functions: by recalling 2.4

$$\widehat{C}(u,v) = \overline{H}(\overline{G}_1^{(-1)}(u), \overline{G}_2^{(-1)}(v)).$$

where $\overline{G}^{(-1)}$ denotes a quasi-inverse of $\overline{G}$, defined analogously to $G^{(-1)}$ in (2.6), like

$$\overline{G}^{(-1)}(t) = \overline{G}^{(-1)}(1-t).$$

An example of a family of copulas built in this way is given by the *Marshall-Olkin* system of bivariate exponential distributions (see e.g [**76, 102, 106**]). This family is modeled by two parameter, say α_1, α_2, with values in $[0,1]$, and its expression reads

$$\widehat{C}^{(\alpha_1,\alpha_2)}(u,v) := u\,v\,\min\{u^{-\alpha_1},\, v^{-\alpha_2}\}. \tag{2.22}$$

This model is suitable to describe the lifetime a system with two components, which are subject to shocks that are fatal to one or both of them. For this reason it is a model that fits well with reliability problems and finds many applications in such field. We discuss its construction in Chapter 5, also giving some of its properties, useful for our purposes.

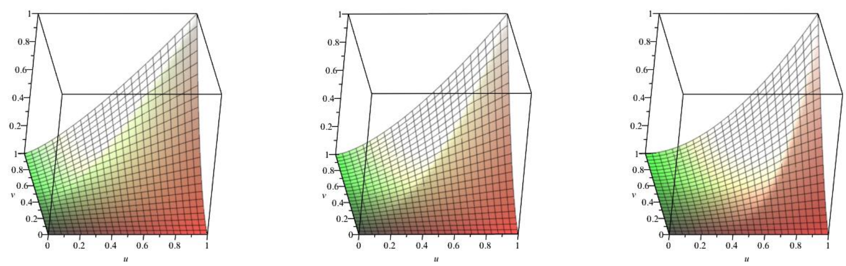

FIGURE 2.3. Copulas from the Marshall-Olkin family with parameters (α_1, α_2) respectively $(0.3, 0.6)$, $(0.5, 0.5)$ and $(0.9, 0.3)$

Other examples of copulas constructed by using the inversion method can be found in [**106**], like the one from the circular uniform distribution and others with Gaussian or Cauchy margins.

Another kind of approach in building copulas arise for geometric-type methods. One can, indeed, construct grounded $2-$increasing functions on $\mathbb{I}^2$ with uniform margins, by using some information of a geometric nature, such as the shape of the graphs of horizontal, vertical, or diagonal sections. Nelsen in [**106**] examines *ordinal sum* construction, wherein the members of a set of copulas are scaled and translated in order to construct a new copula; the *shuffles of* M, which are constructed from the Fréchet-Hoeffding upper bound; and the *convex sum* construction, a continuous analog of convex linear combinations. For our purposes we will discuss about the second kind of geometric method mentioned, the shuffles of M, the maximal copula.

It is known that M and W are singular copulas, whose support consists in a single segment. For M it consists of the line connecting $(0, 0)$ with $(1, 1)$, with slope 1, while W is supported by the line connecting $(0, 1)$ to $(1, 0)$, with slope -1. All copulas with support consisting in segments of slope -1 and 1 are called shuffles of M.

Informally speaking, we can say that such copulas are obtained by cutting the support of M in small parts and rearranging into the unit square by translating and flipping them.

More formally, a shuffle of M is determined by a positive integer n, a finite partition $\{J_i\}_{i=1,...,n}$ of $\mathbb{I}$ into n closed subintervals, a permutation σ on $N = 1, 2, \ldots, n$, and a function $\omega : N \to \{-1, 1\}$ where $\omega(i)$ is -1 or 1 according to whether or not the strip $J_i \times mI$ is flipped. The shuffle of M resulting from a permutation σ will be denoted by $M(n, \{J_i\}, \sigma, \omega)$. A shuffle of M with $\omega = 1$, i.e., for which none of the strips are flipped, is a straight shuffle, and a shuffle of M with $\omega = -1$ is called a flipped shuffle. We will also write I_n for $\{J_i\}$ when it is a regular partition of $\mathbb{I}$, i.e., when the width of each subinterval J_i is $1/n$.

As an example of shuffle consider the following copula C_γ, described by a parameter $\gamma \in (0, 1)$.

$$C_\gamma(u, v) = \begin{cases} \min(u, v - \gamma), & \text{if } (u, v) \in [0, 1 - \gamma] \times [\gamma, 1] \\ \min(u + \gamma - 1, v), & \text{if } (u, v) \in [1 - \gamma, 1] \times [0, \gamma] \\ W(u, v), & \text{otherwise.} \end{cases} \tag{2.23}$$

This copulas are built by one single cut at γ and represents a straight shuffle of M, since no strips are flipped. The graph of the support consists in two lines with slope 1 connecting $(0, \gamma)$ to $(1 - \gamma, 1)$ and $(1 - \gamma, 0)$ to $(1, \gamma)$. This example arise when considering two uniform random variables U and V, with $V = U \oplus \gamma$, with the meaning that the value of V is given by the fractional part of the sum $U + \gamma$. Then one can see that the joint behavior of U and V is expressed in terms of the copula C_γ. We make a deeper investigation of this copula in Section 5, see Proposition 5.10.

Many other methods of constructing copulas exist in literature, as the algebraic methods, involving both bivariate and marginal distributions of the random variables considered. Some examples of copulas built with these methods are given by the Plackett family and the Ali-Mikhail-Haq family of distributions. These and other examples are studied accurately in [**106**].

4. Symmetry and exchangeability

An important concept in the study of copulas is due to its property of symmetry. Symmetry of copulas is strictly liked with the concept of *exchangeability* of the random variables described by them. The first to introduce the concept of exchangeability was De Finetti in [**36**].

DEFINITION 2.31. *We say that two random variables X_1 and X_2, with marginal laws G_1 and G_2 and joint law $F_{1,2}$, are exchangeable if and only if $G_1 = G_2$ and $F_{1,2} = F_{2,1}$.*

A similar definition can be given for groups of random variables. If we consider n identically distributed random variables $X_1, \dots, X_n$, they are exchangeable if $F_{1,\dots,n} = F_{\sigma(1),\dots,\sigma(n)}$, for any permutation σ of the indices $1, \dots, n$.

It is immediate to think to, when two random variables are not exchangeable, how to measure they degree of non-exchangeability.

DEFINITION 2.32. *Let $H(F)$ be the class of all random pairs (X_1, X_2) such that X_1 and X_2 are identically distributed with continuous d.f. F. A function $\widehat{\mu} : H(F) \to \mathbb{R}_+$ is a* measure of non-exchangeability *for $H(F)$ if it satisfies the following properties:*

- **A1:** *$\widehat{\mu}$ is bounded, viz. there exists $K \in \mathbb{R}_+$ such that, for all $(X_1, X_2) \in H(F)$, $\widehat{\mu}(X_1, X_2) \leq K$;*
- **A2:** *$\widehat{\mu}(X_1, X_2) = 0$ if, and only if, (X_1, X_2) is exchangeable;*
- **A3:** *$\widehat{\mu}(X_1, X_2)$ is symmetric, i.e., for all $(X_1, X_2) \in H(F)$, one has $\widehat{\mu}(X_1, X_2) = \widehat{\mu}(X_2, X_1)$;*
- **A4:** *$\widehat{\mu}(X_1, X_2) = \widehat{\mu}(f(X_1), f(X_2))$ for every strictly monotone function f and for all $(X_1, X_2) \in H(F)$;*
- **A5:** *if (X_1^n, X_2^n) and (X_1, X_2) are pairs of random variables with joint distribution functions H_n and H, respectively, and if H_n converges weakly to H as n tends to ∞, then $\widehat{\mu}(X_1^n, X_2^n)$ converges to $\widehat{\mu}(X_1, X_2)$ as n tends to ∞.*

Axioms **A1** and **A2** ensures that the measure is bounded and non always equal to 0. The other axioms state that the measure must be invariant under permutation of components, strictly monotone transformations and distributional limit. This Definition is due to Durante et al. (see [**49**]), where they showed in addition that, by means of Sklar's Theorem, an equivalent formulation of measure of non-exchangeability can be given w.r.t. the connecting copula of the random variables X_1 and X_2.

PROPOSITION 2.33. *Let X_1, X_2 be continuous r.v.s and let C_{X_1,X_2} be their connecting copula. The random variables X_1 and X_2 are exchangeable if, and only if, they are identically distributed, i.e. $F_{X_1} = F_{X_2}$, and C_{X_1,X_2} is symmetric, viz. $C_{X_1,X_2}(u, v) = C_{X_1,X_2}(v, u)$ for every $u, v \in [0, 1]$.*

In this view one can rewrite Definition 2.32 by

DEFINITION 2.34. *Let $\mathcal{C}$ the class of all copulas. A function $\mu : \mathcal{C} \to \mathbb{R}_+$ is a* measure of non-exchangeability *for $\mathcal{C}$ if it satisfies the following properties:*

- **B1:** *μ is bounded, viz. there exists $K \in \mathbb{R}_+$ such that, for all $C \in \mathcal{C}$, $\mu(C) \le K$;*
- **B2:** *$\mu(C) = 0$ if, and only if, C is symmetric;*
- **B3:** *$\mu(C) = \mu(C^t)$ for every $C \in \mathcal{C}$;*
- **B4:** *$\mu(C) = \mu(\widehat{C})$ for every $C \in \mathcal{C}$;*
- **B5:** *if (C_n) and C are in $\mathcal{C}$ and if C_n converges uniformly to C, then $\mu(C_n)$ converges to $\mu(C)$ as n tends to ∞.*

Several measures of non-exchangeability, that satisfy $B1 - B5$, have been presented in [**49**]. Consider, for example, d_p, the classical L_p distance in $\mathcal{C}$ (with $p \in [1, \infty]$). For all $A, B \in \mathcal{C}$ one has

$$d_p(A,B) := \left(\int_0^1 \int_0^1 |A(u,v) - B(u,v)|^p du dv \right)^{1/p} \tag{2.24}$$

for p finite and, for $p = \infty$,

$$d_\infty(A,B) := \max_{(u,v)\in I^2} |A(u,v) - B(u,v)|. \tag{2.25}$$

It has been showed in [**49**] that $\mu_p : \mathcal{C} \to \mathbb{R}_+$ is a measure of non-exchangeability for every $p \in [1, \infty]$. Klement and Mesiar in [**82**] and Nelsen in [**107**] showed that, for every copula C, $\mu_\infty(C) \le 1/3$ and that the upper bound is attained. More in particular two copulas are considered for this purpose, namely

$$C_1(u,v) = \min\left(u, v, \max\left(u - \frac{2}{3}, 0\right) + \max\left(v - \frac{1}{3}, 0\right)\right),$$
$$C_2(u,v) = \max\left(u + v - 1, \frac{1}{3} - \max\left(\frac{1}{3} - u, 0\right) - \max\left(\frac{2}{3} - v, 0\right), 0\right),$$

whose support are described by Figure 2.4.

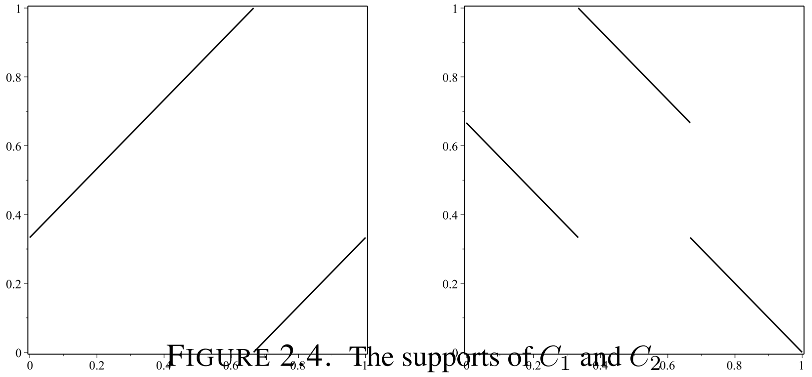

FIGURE 2.4. The supports of C_1 and C_2.

Notice that copula C_1 is part of the one-parameter family C_γ introduced in (2.23) (in particular, it corresponds to the value $\gamma = 1/3$). In Chapter 5 we will investigate, among other things, some properties of symmetry and dependence of copulas, also by means of the family C_γ cited above (see Proposition 5.10).

CHAPTER 3

Decisions under Risk and Utility Theory

The field of decision theory under risk and uncertainty represents a very important area of interest in economics and finance. Conditions of uncertainty are typically considered in situations involving a Decision Maker (from now on, a DM) that is facing a choice among different opportunities, or *acts*, whose consequences are not deterministic. These choices may involve investments in financial assets or insurances as well as bets in gambles or lotteries, and so on. All these situations involve objective facts (known or unknown), regarding the possible choices and its consequences, and subjective matters, as the will of the DM when facing risky or uncertain situations, as well.

The utility theory had arisen and developed in time, to describe diversities among DMs and their attitudes toward risk and uncertainty. The main practical argument studied in this Chapter is the one of *utility functions*, that are used to describe the behavior of the DM by attempting to order the set of consequences corresponding to her choices.

1. Choice under Uncertainty

The modern analysis of decisions under uncertainty has seen its first formalization in 1944 [**142**], when Von Neumann and Morgenstern (NM) laid the foundations of what is known as the *axiomatic theory of expected utility*. Their starting point is the study of people's preferences with regard to choices that have uncertain outcomes, namely *gambles*. Their hypothesis state that if certain axioms are satisfied, the subjective value associated with a gamble by an individual is the expectation of that individual's valuations of the outcomes of that gamble. According to the principle of maximization of the utility, Decision Makers make use of their (subjective) utility functions to evaluate gambles and then they try to maximize their expected outcome.

In their work, Von Neumann and Morgenstern also devoted attention to the attitude of investors towards risk. Decision Makers can be, in particular, classified according to three categories of behavior: *risk neutral*, *risk averse* and *risk seeker*. The first one represents people that, when facing two risky prospects with the same expected value, will feel indifferent in the choice between them. The second one is the attitude of DMs that, when asked to choose between two such prospects, will prefer the less risky one, in contraposition to the risk seeker people that will behave in the complete opposite way.

Consider, for instance, the set of random variables taking values in the finite set $\mathcal{Z}$. Since variables are characterized by their probability laws, we are led to consider a preference relation $\succ$ on $\mathcal{P}(\mathcal{Z})$, the set of all probability measures on $\mathcal{Z}$ or, in fact, on its $\sigma-$algebra $\sigma(\mathcal{Z})$; by abuse of language, each element $p \in \mathcal{P}(\mathcal{Z})$ will be called a lottery. The representation of the expected utility according to NM principles for the preference relations consists now in establishing a utility function $U : \mathcal{P}(\mathcal{Z}) \to \mathbb{R}$ built from a function $u : \mathcal{Z} \to \mathbb{R}$ such that

$$U(p) = \mathbb{E}(u(X)) = \sum_{z \in \mathcal{Z}} p(z)u(z) \tag{3.1}$$

for any random variable $X \in \mathcal{Z}$ with law p, being $\mathbb{E}$ the expected value evaluated over $u(X)$.

Let now recall some basic properties that are required for preference relations.

Asymmetry: $x \succ y$ implies $y \not\succ x$;
Negative transitivity: $x \not\succ y$ and $y \not\succ z$ imply $x \not\succ z$.

From a *strict* preference relation we can immediately define some related ones: the equivalence $\sim$, the *weak* relation $\succeq$, and the reverse relation $\prec$, with same characteristics of $\succ$. Concerning the first one, we write $x \sim y$ when $x \not\succ y$ and $y \not\succ x$ simultaneously. This is an *indifference relation* (or *equivalence relation*), as it is reflexive, symmetric and transitive. On the other hand, we write $x \succeq y$ when $y \not\succ x$ or, equivalently, when both $x \succ y$ or $x \sim y$ may occur: the weak relation is a complete and transitive relation. We now focus on the relation $\succ$ and give some of its basic properties in what follows.

DEFINITION 3.1. *The preference relation $\succ$ is said to be* rational *if it satisfies the following two axioms:*

Continuity or Archimedean: *For all $p, q, r \in \mathcal{P}(\mathcal{Z})$, if $p \succ q \succ r$, then there exist $\alpha, \beta \in (0, 1)$ such that*

$$\alpha p + (1 - \alpha)r \succ q \succ \beta p + (1 - \beta)r \tag{3.2}$$

Independence: *For all $p, q, r \in \mathcal{P}(\mathcal{Z})$ and $\alpha \in (0, 1]$, if $p \succ q$ then*

$$\alpha p + (1 - \alpha)r \succ \alpha q + (1 - \alpha)r. \tag{3.3}$$

Furthermore we introduce the following two lemmas.

LEMMA 3.2. *If the preference relation $\succ$ is rational, then the following hold:*

(1) for two given real numbers $a, b \in [0, 1]$, with $a < b$, one has

$$p \succ q \Rightarrow bp + (1 - b)q \succ ap + (1 - a)q;$$

(2) the conditions $p \succeq q \succeq r$ and $p \succ r$ imply that there exist an unique $a^* \in [0, 1]$ such that

$$q \sim a^* p + (1 - a^*)r;$$

(3) for $a \in [0,1]$, $p \sim q$ *imply, for any* $r \in \mathcal{P}(\mathcal{Z})$

$$ap + (1-a)r \sim aq + (1-a)r.$$

In the following we indicate, with the symbol $\delta_z(A)$, Dirac's delta for the element z into set A, as a quantity equal to 1 if $z \in A$ and 0 otherwise.

LEMMA 3.3. *If* $\succ$ *satisfy the axioms of Definition 3.1, then there exist* z_0 *and* z^0 *in* $\mathcal{Z}$ *such that* $\delta_{z^0} \succeq p \succeq \delta_{z_0}$ *for any* $p \in \mathcal{P}(\mathcal{Z})$.

Finally, we can state the following

THEOREM 3.4. *Any rational preference relation* $\succ$ *can be uniquely represented by a function* u *up to a positive linear transformation, i.e., if* $u : \mathcal{Z} \to \mathbb{R}$ *is such that*

$$p \succ q \Leftrightarrow \sum_{z \in \mathcal{Z}} p(z)u(z) > \sum_{z \in \mathcal{Z}} q(z)u(z)$$

then the same holds with u *being replaced by* $v(\cdot) = au(\cdot) + b$*, for any two real numbers* a, b *with* $a > 0$.

Proofs for the two lemmas together with Theorem 3.4 can be found in [**130**].

We turn now to consider more realistic random quantities with the meaning of economic tools, as possible losses in investment portfolios. Consider $\mathcal{Z}$ as a possibly infinite set of random variables, with range on $\mathbb{R}$. Let $\mathcal{Z}$ be a complete and separable metric space with its sigma-algebra $\sigma(\mathcal{Z})$. As a generalization of the finite case, the numerical representation of a preference relation $\succ$ on a class of random variables $\mathcal{X}$, according to the NM-principles, is the following. For any DM with preference relation $\succ$ there exists an unique (or unique up to positive linear transformations) utility function $u : \mathcal{Z} \to \mathbb{R}$ such that

$$X \succ Y \Leftrightarrow \mathbb{E}(u(X)) > \mathbb{E}(u(Y)),$$

for any $X, Y \in \mathcal{X}$, provided that the expectations exist. This is known as the *von Neumann-Morgenstern expected utility representation of preference relations* (for references, see [**142**]).

Let now $\mathcal{Z} = \mathbb{R}$ for simplicity, with its usual topology generated by open intervals; the Borel $\sigma-$field of $\mathbb{R}$ is denoted as $\mathcal{B}(\mathbb{R})$. So $\mathcal{P}(\mathcal{Z})$ ($\mathcal{P}(\mathbb{R})$) will be the set of all probability measures defined on $\mathcal{B}(\mathbb{R})$. A topology on $\mathcal{P}(\mathbb{R})$ is defined by specifying a concept of convergence in it. Here we say that p_n converges weakly to p in $\mathcal{P}(\mathbb{R})$ if, for any $f \in C_b(\mathbb{R})$ (bounded continuous functions with support $\mathbb{R}$),

$$\int_{\mathbb{R}} f(x)\, dp_n(x) \to \int_{\mathbb{R}} f(x)\, dp(x). \tag{3.4}$$

The associated topology, generated by the neighborhoods of each p, will be called the weak topology on $\mathcal{P}(\mathbb{R})$.

We now need to extend the axiom of continuity defined in (3.2).

DEFINITION 3.5. *Given a separable and complete metric space $\mathcal{P}(\mathcal{Z})$, for any $p_n, p \in \mathcal{P}(\mathcal{Z})$ with $p_n \to p$ in the weak topology, the following hold:*

(1) if $p \succ q$ for some $q \in \mathcal{P}(\mathcal{Z})$, then $p_n \succ q$ for sufficiently large n;
(2) if $q \succ p$, then $q \succ p_n$ for all sufficiently large n.

By means of this new Definition of weak continuity, we can extend result given in Theorem 3.4 to $\mathcal{P}(\mathbb{R})$.

THEOREM 3.6. *Let X, Y random variables in $\mathcal{X}$ with probability laws $p, q \in \mathcal{P}(\mathbb{R})$ respectively. A preference relation $\succ$ on $\mathcal{P}(\mathbb{R})$ satisfies the independence and the weak continuity axioms if and only if there exists a bounded and continuous utility function $u : \mathbb{R} \to \mathbb{R}$ such that*

$$p \succ q \Leftrightarrow \mathbb{E}_p(u(X)) = \int_{\mathcal{Z}} u(z)\, dp(z) > \int_{\mathcal{Z}} u(z)\, dq(z) = \mathbb{E}_q(u(Y)). \tag{3.5}$$

Moreover this representation is unique up to positive affine transformations.

2. The Expected Utility Principle and Risk Attitudes

In this Section we sketch the basics of the attitudes of Decision Makers facing risky or uncertain situations. From a qualitative point of view, we can say that there are essentially three types of risk attitudes (or three kind of people), namely *risk neutral*, *risk averse* and *risk seeker*. As we said in the previous Section, the former are indifferent in choosing between two risky prospects, but with the same expected value; *risk-averse* ones, among prospects with the same expected value, prefer the less risky (for them you have $u(\mathbb{E}[X]) \geq \mathbb{E}[u(X)]$, then they make use of a concave utility function); those *risk seeker*, finally, share the opposite attitude (and hence make use of a convex u).

For example, suppose that our DM is asked to choose between accepting 100 € or playing the following game. The DM flips a coin, if head occur she gets 200 €, otherwise nothing. The expected value of both prospects is the same, 100 €, but DMs with different concepts of risk will act differently in deciding what to do. A risk neutral DM will be indifferent in the choice of prospect, maybe she will decide to toss the coin at first to decide whether accepting money or playing the lottery. A risk averse DM will instead accept the money (as a risk seeker will try to win the best possible prize).

To be more precise we restate the problem in the following way. Let X be a random variable taking values 0 or 200, both with probability $1/2$. So we have that the expected value of the lottery is given by $\mathbb{E}(X) = 100$. Now the three conditions of neutrality, aversion, and propensity to risk can be restated as follows. A DM is

Risk Neutral: when her utility function u is such that $u(\mathbb{E}(X)) = \mathbb{E}(u(X))$, for example she may consider $u(x) = x$;
Risk Averse: when $u(\mathbb{E}(X)) > \mathbb{E}(u(X))$, if she chooses a concave utility function, for example $u(x) = \log(x)$;

Risk Seeker: when $u(\mathbb{E}(X)) < \mathbb{E}(u(X))$, or when her utility function is convex.

Let's consider a utility function $u \in C^2(\mathbb{R})$, the space of differentiable functions with continuous derivatives until order 2. It is intuitive to notice, for first, that when considering monetary incomes or any kind of profitable goods as prospects, the satisfaction in receiving them shall grow as the amount of the prospect grows. This is the reason for which it is generally assumed (as we do in this work) that the utility function must be nondecreasing. Namely, for a given utility u, its first derivative u' is supposed to be greater or equal than zero. Moreover the sign of the second derivative plays an important role, since it determines univocally the behavior of the DM towards risk, placing her in one of the three categories of risk listed above. So we can state the following

THEOREM 3.7. *A DM with nondecreasing utility u is risk averse if and only if u is concave, risk seeker if and only if u is convex.*

The proof of this Theorem mainly follows from a direct application of Jensen's inequality.

Let now focus our attention on comparisons between risk and risk aversion. Risk aversion is the attitude to avoid uncertainty as well as to insure oneself from unpredictable events. So, up to von Neumann-Morgenstern principles of utility maximization, a risk averse DM will behave like follows.

Suppose that our DM has an initial capital x, that is a deterministic (positive) amount of money. Suppose that she is also risk averse, so that her utility function u will be nondecreasing and concave. The DM is now going to make an investment whose uncertain profit consists of a random variable Z (that is not necessarily nonnegative). In this terms, she will be ready to pay a premium π in order to replace Z with its expected value $\mathbb{E}(Z)$. This premium will depend on both the initial capital x and the law of the random profit Z, so $\pi = \pi(x, F_Z)$. So we can uniquely define it as the value satisfying

$$u(x + \mathbb{E}(Z) - \pi(x, F_Z)) = \mathbb{E}(u(x + Z)). \tag{3.6}$$

DEFINITION 3.8. *Given an initial capital x and a random profit Z, the quantity $\pi(x, F_Z)$ satisfying equation* (3.6) *is called* risk premium.

For the sake of simplicity, let $\mathbb{E}(Z) = 0$ and indicate with σ_Z^2 the variance of Z. By Taylor expansion we have

$$u(x - \pi) = u(x) - \pi u'(x) + \mathcal{O}(\pi^2) \tag{3.7}$$

$$\begin{aligned} \mathbb{E}(u(x + Z)) &= \mathbb{E}(u(x) + Zu'(x) + \frac{1}{2}Z^2u''(x) + \mathcal{O}(Z^3)) \\ &= u(x) + \frac{1}{2}\sigma_Z^2 u''(x) + o(\sigma_Z^2). \end{aligned} \tag{3.8}$$

Thus, rearranging the terms in (3.7) and (3.8) we obtain

$$\pi(x, F_Z) = \frac{1}{2}\sigma_Z^2 A(x) + o(\sigma_Z^2), \tag{3.9}$$

where

$$A(x) = -\frac{u''(x)}{u'(x)} = -\frac{d}{dx}\log u'(x). \tag{3.10}$$

DEFINITION 3.9. *The function $A(x)$ introduced in* (3.10) *is called* absolute local measure of risk aversion*, and represents the local propensity to insure at point x under the utility function u.*

If Z is not actuarially neutral, namely if $\mathbb{E}(Z) \neq 0$, the expression for the risk premium will take the form

$$\pi(x, F_Z) = \frac{1}{2}\sigma_Z^2 A(x + \mathbb{E}(Z)) + o(\sigma_Z^2). \tag{3.11}$$

If the profit of the investment is expressed by a multiplicative utility, the proportional risk premium π^* will be defined as the value that satisfies

$$u(x\mathbb{E}(Z) - x\pi^*(x, F_Z)) = \mathbb{E}(u(xZ)). \tag{3.12}$$

In case that $\mathbb{E}(Z) = 0$, the expression for π^* will be

$$\pi^*(x, F_Z) = \frac{1}{2}\sigma_Z^2 R(x) + o(\sigma_Z^2), \tag{3.13}$$

where $R(x) = xA(x)$.

DEFINITION 3.10. *The function $R(x)$ satisfying* (3.13) *is called* relative local measure of risk aversion.

One can notice that there is a relationship linking the two risk premiums, that is expressed by

$$\pi(x, F_{xZ}) = x\pi^*(x, F_Z).$$

Notice furthermore that the above representations of risk premium are "local" representations, since they describe DM's behavior towards small (infinitesimal) risks. The following Theorem shows that there is an analogy between local and global behaviors in terms of risk aversion.

THEOREM 3.11 (Arrow, Pratt). *Let u_1, u_2 be two utility functions with absolute local measures of risk aversion A_1, A_2 and risk premiums π_1, π_2 respectively. Then, for any choice of x and Z, the following conditions are equivalent:*

(1) $A_1(x) \geq A_2(x)$;
(2) $\pi_1(x, F_Z) \geq \pi_2(x, F_Z)$;
(3) $u_1(\cdot) = k(u_2(\cdot))$, with k increasing and concave.

See [**8**] and [**110**] for further details.

DEFINITION 3.12. *If the hypothesis of Theorem 3.11 are satisfied, then a DM with utility function u_1 is said more risk averse than a DM preferring u_2.*

Connections between small risks and measures of risk aversion have been noticed for first by De Finetti (see [**37**]).

Another important property is given by the following result.

THEOREM 3.13. *The following conditions hold:* $A(x)$ *is decreasing in* x *if and only if* $\pi(x, F_Z)$ *is decreasing in* x *for all* Z. *Analogously,* $R(x)$ *is decreasing in* x *if and only if* $\pi^*(x, F_Z)$ *is decreasing in* x *for all* Z.

If the conditions of Theorem 3.13 are met we say that u exhibits decreasing (absolute or relative) risk aversion.

3. Multi-attribute utilities and related properties

Recent years witnessed numerous attempts to generalize various aspects of the these notions to the case of multivariate risk (see, for example, the works by Duncan [**48**], Karni [**78**], Kihlstrom and Mirman [**81**], and Stiglitz [**133**]). The univariate case is qualitatively different from the multivariate one: in the first case the ordinal preferences of all decision makers are identical, whereas in the latter the preference orderings may differ among them.

Let $\mathbf{x} = (x_1, \ldots, x_n)$ be the (deterministic) commodity vector of an investor facing a risk $\mathbf{Z}$, that is expressed, in turn, by a $n-$dimensional (random) vector $(Z_1, \ldots, Z_n)$. Let u any real-valued function which is in the equivalence class of von Neumann-Morgenstern utility functions consistent with the individual's preferences. We assume that u is strictly increasing in each component and that $\mathbb{E}u(\mathbf{x} + \mathbf{Z})$ is finite.

We define a family of risk premium functions $\boldsymbol{\pi}(\mathbf{x}, \mathbf{Z})$ in the following way. For a given risk vector $\mathbf{Z}$, with $\mathbb{E}(\mathbf{Z}) = 0$, the vector $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$ must satisfy

$$u(\mathbf{x} - \boldsymbol{\pi}) = \mathbb{E}u(\mathbf{x} + \mathbf{Z}). \tag{3.14}$$

Note that the risk premium is unique in the univariate case while in the multivariate case the existence of a vector $\boldsymbol{\pi}$ is granted but uniqueness does not necessarily hold. A simple example of this situation is given by the following utility function: $u(x_1, x_2) = x_1 x_2$. Equation (3.14) is satisfied if $\pi_1\pi_2 - \pi_1 x_2 - \pi_2 x_1 = \sigma_{12}$.

A matrix measure of *multivariate local risk aversion*, which is directly related to the multivariate risk premiums, can be given as follows. Consider the Taylor series expansion of both members of equation (3.14). At first consider $u_{ij}(\mathbf{x}) = \partial^2 u(\mathbf{x})/\partial x_i \partial x_j$ to be continuous; we obtain

$$u(\mathbf{x} - \boldsymbol{\pi}) = u(\mathbf{x}) - \sum_{i=1}^{n} \pi_i u_i(\mathbf{x}) + \frac{1}{2} \sum_{i,j=1}^{n} \pi_i \pi_j u_{ij}(\mathbf{x} - \theta\boldsymbol{\pi})$$

with $\theta \in [0, 1]$. Secondly, if Var$(\mathbf{Z}) = \Sigma = [\sigma_{ij}]$ exists,

$$\mathbb{E}u(\mathbf{x} + \mathbf{Z}) = u(\mathbf{x}) + \frac{1}{2} \sum_{i,j=1}^{n} \sigma_{ij} u_{ij}(\mathbf{x}) + o(tr\Sigma),$$

where $tr\Sigma = \sum_{i=1}^{n} \sigma_{ii}$. Combining these two approximation one gets to

$$\mathbf{u}'\boldsymbol{\pi} = -\frac{1}{2} tr\mathbf{U}\Sigma, \tag{3.15}$$

where the $n \times n$ Hessian matrix $\mathbf{U} = [u_{ij}(\mathbf{x})]$ and the $n-$vector $\mathbf{u} = (u_i(\mathbf{x}))$. Any approximate solution of (3.15) will be of the form

$$\widehat{\boldsymbol{\pi}} = \frac{1}{2} dg\mathbf{A}\Sigma \tag{3.16}$$

where

$$\mathbf{A} = [a^{ij}] = \left[-\frac{u_{ij}}{u_i}\right] = [diag\mathbf{u}]^{-1}\mathbf{U}. \tag{3.17}$$

Reducing the problem to one dimensional case, $\mathbf{A}$ coincides with the absolute measure A presented in (3.10), so we can call $\mathbf{A}$ *absolute risk aversion matrix.* The importance of off-diagonal elements is given by the following two results (proofs are in [**48**]).

PROPOSITION 3.14. *The matrix* $\mathbf{A}$ *is diagonal if and only if* u *is additive. In this case the commodities are mutually risk independent.*

PROPOSITION 3.15. *If there exists a nonnegative risk premium vector* $\boldsymbol{\pi}$ *for all two-point gambles* $\mathbf{Z}$*, then* u *is concave. The viceversa also holds true.*

In this direction, it is clear that interactions among acts play a fundamental role in assessing risk and defining aversion to it. Comparisons among risk between risk averse DMs can be made by the comparison of their risk premiums, but one can see that this model is suitable only when dealing with small risks. Kihlstrom and Mirman, in fact, showed in [**81**] that under these hypothesis, a DM that is more risk averse than another DM in one direction will be more risk averse in any direction. This is due to the strong hypothesis that two utility functions represent the same preference ordering.

Extensions to the Arrow-Pratt concept of risk aversion to the multivariate case are somehow problematic. A more precise concept of multivariate aversion is that one of *correlation aversion* (CAV), introduced by Epstein and Tanny in [**53**]. For the sake of simplicity we restrict, for the moment, to the $2-$dimensional case.

Consider two vectors of outcomes (x_1, x_2) and (y_1, y_2), with $x_1 < y_1$, $x_2 < y_2$, and two lotteries L_1, L_2 such that

$$L_1 = \begin{cases} (x_1, x_2), & \text{w.p. } 1/2; \\ (y_1, y_2), & \text{w.p. } 1/2; \end{cases} \qquad L_2 = \begin{cases} (x_1, y_2), & \text{w.p. } 1/2; \\ (y_1, x_2), & \text{w.p. } 1/2; \end{cases}$$

In this respect, we can say that $\mathbf{x}$ is the vector of "bad" outcomes, while $\mathbf{y}$ is the "good" one (see [**45, 51**]). So lottery L_2 associates bad with good outcomes while L_1 is some kind of "all or nothing". Since the marginal outcomes of the two lotteries are the same, intuitively a DM preferring L_2

to L_1 will manifest a form of bivariate risk aversion. The above preference holds for all $\mathbf{x} \leq \mathbf{y}$ if and only if DM's utility function u satisfies

$$u(x_1, x_2) + u(y_1, y_2) \leq u(x_1, y_2) + u(y_1, x_2), \tag{3.18}$$

condition that corresponds to *submodularity* of the function u (compare with equation (1.10) in Definition 1.11). Furthermore, if u is twice differentiable, condition (3.18) can be simply rewritten as

$$u_{12} \leq 0. \tag{3.19}$$

Notice that, in dimension 2, such a formulation follows the one of multivariate local risk aversion introduced above. A deeper analysis of risk attitudes has been made in this direction, based on the sign of higher order derivatives of the utility function. This led to the introduction of properties like *prudence* and *temperance* (and *cross-prudence*, *cross-temperance* for multivariate case) and other general properties, that can be found, for example, in [**45, 46, 51, 140**].

4. Non-Expected Utility Theory

Decision Makers, intended to be rational in the sense of Theorem 3.4, usually try to maximize their expected utility according to the so-called *expected utility maximization principle*. But in the real world, assuming that everybody is rational is pure utopia. In the last years, in fact, many objections to this principle have been formulated, by means of some paradoxical examples. First of all Allais, in 1953, proposed a paradox in contrast to the theory proposed by von Neumann and Morgenstern; Ellsberg (1961) doubted the axiomatic formulation of Savage (1954), subsequently giving rise to an alternative expected utility theory based on generalized Choquet integrals (1965).

Actually, the very first example of this contradiction was proposed by Daniel Bernoulli over the 18^{th} century, becoming over the years the well-known *St. Petersburg paradox*. This example presents a casino that offers a lottery for a single player, described as follows.

At each stage a fair coin is tossed. The pot starts at 2 dollars and is doubled every time a head appears. The first time a tail appears, the game ends and the player wins whatever is in the pot. Thus the player wins 2 dollars if a tail appears on the first toss, 4 dollars if a head appears on the first toss and a tail on the second, and so on. In short, the player wins 2^k dollars, where k equals number of tosses. These events may appear with probability, respectively, $1/2, 1/4, \ldots, 2^{-k}$. To evaluate the expected value of the gamble, on just need to evaluate the quantity

$$E = \sum_{i=1}^{\infty} 2^i \cdot \left(\frac{1}{2}\right)^i = \sum_{i=1}^{\infty} 1 = \infty. \tag{3.20}$$

Assuming that the game can continue as much as the gambler wants, and that both gambler and casino have infinite amounts of money, the game result to have infinite expected utility, as the expected win seems not finite. Under this viewpoint any gambler should be willing to pay any amount of money to have the chance of participating to such a game. But this situation is obviously unfeasible. The paradox is then in the discrepancy between what people seem willing to pay to enter the game and its infinite expected value.

Another interesting example is the one of Allais's paradox (1953), that describes inconsistencies in choices when people are deciding between options in two gambling games, one of which involves a certain outcome.

Gamble A consists in a choice between A_1 and A_2, with

- A_1: 1 million € with certainty;
- A_2: 1% chance of zero, 89% chance of 1 million €, and 10% chance of 5 million €.

Gamble B is the following

- B_1: 89% chance of zero, 11% chance of 1 million €;
- B_2: 90% chance of zero and 10% chance of 5 million €.

Standard economic theory predicts that a person with consistent preferences will chose B_1 in the second gamble if prefers A_1 in the first (or B_2 if A_2). The expected value for A_1 and for B_1 are, respectively, smaller than the ones for A_2 and B_2, but chances to get zero are diminished (or completely eliminated). However, experimental evidence shows that real people commonly choose the inconsistent combinations (A_1, B_2) and (A_2, B_1). Kahneman and Tversky attributed this violation to expected utility principles to a "certainty effect", as they explain in [**77**], introducing the formulation of the so called "Prospect Theory". This new concept was mainly based on two principles, such as the effect of context in which decisions are assumed and the investor's aversion to losses. According to the former one has to consider that an investor usually makes different choices for a same decisional problem depending on the context in which such problem is presented; the latter makes the investor more risk seeker than she would have been risk averse in case of gain.

Ellsberg paradox (1961) can be illustrated by an urn game. An urn contains 90 balls, 30 of which are red, and the remaining 60 are divided into black and yellow balls, with unknown proportion. Subjects playing the game are asked for their preferences over two gambles. In both gambles one ball is drawn from the urn and players have to guess its color.

Gamble C is given by

- C_1: the ball is red;
- C_2: the ball is black.

Gamble D is divided into

- D_1: the ball is red or yellow;
- D_2: the ball is black or yellow.

Ellsberg found that many people prefer to bet on gambles C_1 and D_2, violating the "sure thing principle", which requires that if C_1 is preferred to C_2, than D_1 should be preferred to D_2. Ellsberg attributed this inconsistency to *ambiguity aversion* in the face of Knightian uncertainty. As defined by Knight in 1921 (see [**84**]), it describes fundamental uncertainty and unknowable probabilities. Knightian risk describes probabilities that can be quantified because they capture observable, repeatable events, which can be measured or which are given as prior information about proportion. With this example Ellsberg showed that the expected utility model fails in situations in which uncertain events are associated with probabilities that cannot be quantified, subsequently giving rise to an alternative expected utility theory based on generalized Choquet integrals (1965).

Other alternative theories have been proposed since the early '80s, like the ones from Machina (1982), Karni (1985), Yaari (1987) and other authors.

CHAPTER 4

The target Based Model for utility

We introduce a formal description of the Target-Based approach to utility theory for the case of $n > 1$ attributes and point out the connections with aggregation-based extensions of capacities on finite sets. Although capacities have been used in literature to go beyond the classical principle of maximization of expected utility, we show how such measures emerge in a natural way in the frame of the target-based approach to classical utility theory, when considering the multi-attribute case. Our discussion provides economic interpretations of different concepts of the theory of capacities. In particular, we analyze the meaning of extensions of capacities based on n-dimensional copulas. The latter describes stochastic dependence for random vectors of interest in the problem. We also trace the connections between the case of $\{0, 1\}$-valued capacities and the analysis of "coherent" reliability systems.

1. Introduction to Target Based model

A rich literature has been devoted in the last decade to the *Target-Based Approach* (TBA) to utility functions and economic decisions (see [**19, 20, 28, 29, 138, 139**]). This literature is still growing, with a main focus on applied aspects (see, for example, [**13, 144, 145**]).

Even from a theoretical point of view, however, some issues of interest demand for further analysis. In this direction, the present Chapter will consider some aspects that emerge in the analysis of the multi-attribute case. Generally TBA can provide probabilistic interpretations of different notions of utility theory. Here we will in particular interpret in terms of stochastic dependence the differences among copula-based extensions of a same fuzzy measure.

In order to explain the basic concepts of the TBA it is, in any case, convenient to start by recalling the single-attribute case. Let $\Xi := \{X_\alpha\}_{\alpha \in A}$ be a family of real-valued random variables, that are distributed according to probability distribution functions F_α respectively. Each element $X_\alpha \in \Xi$ is seen as a *prospect* or a *lottery* and a Decision Maker is expected to conveniently select one element out of Ξ (or, equivalently, $\alpha \in A$). Let $U : \mathbb{R} \to \mathbb{R}$ be a (non-decreasing) utility function, that describes the Decision Maker's attitude toward risk. Thus, according to the *Expected Utility*

Principle (see [**142**]), the DM's choice is performed by maximizing the integral

$$\mathbb{E}\left[U\left(X_\alpha\right)\right] = \int_{\mathbb{R}} U(x)\, dF_\alpha(x). \tag{4.1}$$

In the Target-Based approach one in addition assumes U to be right-continuous and bounded so that, by means of normalization, it can be seen as a probability distribution function over the real line. This approach suggests looking at U as at the distribution function F_T of a random variable T. This variable will be considered as a *target*, stochastically independent of all the prospects X_α. If T is a (real-valued) random variable stochastically independent of X_α in fact, one has

$$\mathbb{E}(F_T(X_\alpha)) = \int \mathbb{P}(T \leq x) F_\alpha(dx) = \mathbb{P}(T \leq X_\alpha), \tag{4.2}$$

and then, by setting $U = F_T$, the Expected Utility Principle prescribes a choice of $\alpha \in A$ which maximizes the quantity $\mathbb{E}\left[U\left(X_\alpha\right)\right] = \mathbb{P}\left(T \leq X_\alpha\right)$.

The conceptual organization and formalization of basic ideas have been proposed at the end of nineties of last century by Castagnoli, Li Calzi, and Bordley. Some arguments, that can be regarded nowadays as related with the origins of TBA, had been around however in the economic literature since a long time (see [**19, 28**] and references therein).

After the publication of these papers, several developments appeared in the subsequent years concerning the appropriate way to generalize the TBA to the case of multi-attribute utility functions, see in particular [**20, 138, 139**]. As already mentioned such an approach, when applicable, offers probabilistic interpretations of notions of utility theory, and this is accomplished in terms of properties of the probability distribution of a random target. Such interpretations, in their turn, are easily understandable and practically useful. In particular, they can help a Decision Maker in the process of assessing her/his own utility function.

A natural extension of the concept of Target-Based utility from the case $n = 1$ to the case of $n > 1$ attributes is based on a specific principle of individual choice pointed out in [**20**]. In this Chapter, we formalize such a principle in terms of the concept of capacity and analyze a TBA multi-attribute utility as a pair (m, F) where m is a capacity over $N = \{1, \ldots, n\}$ and F is an n-dimensional probability distribution function. For our purposes it is convenient to use the Sklar decomposition of F in terms of its one-dimensional margins and of its connecting copula. In such a frame, some aspects of aggregation functions and of copula-based extensions of capacities emerge in a straightforward way.

More precisely, the Chapter will present the following structure. In the next section, we will introduce the appropriate notation and detail the basic aspects of the multi-criteria Target-Based approach. Starting from the arguments presented in [**20**], we show how every Target-Based n-criteria utility is basically determined by a couple of objects: an n-dimensional probability distribution function and a fuzzy measure over $N := \{1, \ldots, n\}$. This discussion will allow us to point out, in Section 3, that some of the results presented by Kolesárová et al. in [**86**] admit, in a completely direct way, probabilistic interpretations and applications in terms of the TBA. It will in particular turn out that n-dimensional copulas, that can be used for the extension of fuzzy measures, describe stochastic dependence among the components of random vectors relevant in the problem. Section 4 will be devoted to the special case of $\{0, 1\}-$valued capacities. We shall see how, under such a specific condition, our arguments are directly related to the field of reliability and of lattice polynomial functions. Some final remarks concerning the relations between the parameters of TBA utilities and economic attitudes of a Decision Maker will be presented in Section 5. The notation we used is motivated by our effort to set a bridge between the two different settings. The term "attribute", as used in the present Chapter, is substantially a synonymous of "criteria".

2. Multi-Attribute Target-Based Utilities

In this section we deal with the TBA form of utility functions with $n > 1$ attributes. As recalled in the introduction, in the single-attribute case, $n = 1$, a TBA utility is essentially a non-decreasing, right-continuous, bounded function that, after suitable normalization, is regarded as the distribution function of a scalar random variable T with the meaning of a target. Actually even more general, non-necessarily increasing, "utilities" can be considered in the TBA when possibility of stochastic dependence is admitted between the target and the prospect (see [**19**], see also [**38**]), but our interest here is limited to the case of independence between such two objects.

At a first glance, one could consider the functions $F(x_1, \ldots, x_n)$ as the appropriate objects for a straightforward generalization of the definition of the TBA utilities to the n-attributes case. A given F should be interpreted as the joint distribution function of a *target vector* $\mathbf{T} := (T_1, \ldots, T_n)$. But such a choice would be extremely restrictive, however. A more convincing definition, on the contrary, can be based on the following principle: in the cases when a single deterministic target t_i ($i = 1, \ldots, n$) has been assessed for any attribute i by the Decision maker, the utility $U_{m,\mathbf{t}}(\mathbf{x})$ corresponding to an outcome $\mathbf{x} := (x_1, \ldots, x_n)$ depends only on the subset of those targets that are met by $\mathbf{x}$ (as in [**20**], Definition 1). More precisely, we assume the existence of a set function $m : 2^N \to \mathbb{R}_+$ such that

$$U_{m,\mathbf{t}}(\mathbf{x}) = m(Q(\mathbf{t}, \mathbf{x})), \tag{4.3}$$

where $Q\,(\mathbf{t}, \mathbf{x})$ is the subset of N defined by

$$Q(\mathbf{t}, \mathbf{x}) := \{i \in N | t_i \le x_i\}. \tag{4.4}$$

It is natural to require that the function m is finite, non-negative, and non-decreasing, namely such that

$$0 = m(\emptyset) \leq m(I) \leq m(N) < \infty$$

Without loss of generality one can also assume that m is scaled, in such a way that

$$m(N) = 1. \tag{4.5}$$

In other words, we are dealing with a *capacity* or a *fuzzy measure* $m : 2^N \to [0, 1]$.

Rather than deterministic targets however, it is generally interesting to admit the possibility that the vector $\mathbf{T}$ of the targets is random, as it happens in the single-attribute case. Denoting by $F_{\mathbf{T}}$ the joint distribution function of $\mathbf{T}$, we replace the definition of a multi-attribute utility function given in (4.3) with the following more general

DEFINITION 4.1. *A multi-attribute target-based utility function, with capacity m and with a random target T has the form*

$$U_{m,F}(\mathbf{x}) = \sum_{I \subseteq N} m(I)\, \mathbb{P}\left(\bigcap_{i \in I}\{T_i \leq x_i\} \cap \bigcap_{i \notin I}\{T_i > x_i\}\right). \tag{4.6}$$

It is clear that $U_{m,F}(\mathbf{x}) = U_{m,\mathbf{t}}(\mathbf{x})$ when the probability distribution described by $F_{\mathbf{T}}$ is degenerate over the point $\mathbf{t} \in \mathbb{R}^{\mathbf{n}}$. On the other hand the special choice $U_{m,F}(\mathbf{x}) = F_{\mathbf{T}}(\mathbf{x})$, mentioned above, is obtained by imposing the condition (4.5) together with

$$m(I) = 0 \text{ for all } I \subset N \tag{4.7}$$

This position corresponds then to a Decision Maker who is only satisfied when all the n targets are achieved.

The class of n-attributes utilities is of course much wider than the one constituted by the functions of the form (4.6). The latter class is however wide enough and the choice of a utility function within it is rather flexible, since a single function is determined by the pair $(m(\cdot), F_{\mathbf{T}})$. Sufficient or necessary conditions, under which a utility function is of the form (4.6), have been studied by Bordley and Kirkwood in [**20**]. Several situations, where such utilities can emerge as natural, have also been discussed.

For our purposes, the following notation will be useful. We denote by $M_m : [0, 1] \to \mathbb{R}$ the set-function obtained by letting, for $I \in 2^N$,

$$M_m(I) := \sum_{J \subseteq I} (-1)^{|I \setminus J|}\, m(J) \tag{4.8}$$

where $|I|$ indicates the cardinality of the set I. The function $M_m(\cdot)$ is the *Möbius Transform* of $m(\cdot)$ and, as a formula of the *inverse Möbius Transform*, we also have $m(I) = \sum_{J \subseteq I} M_m(J)$ (see e.g. [**115**]). For $\mathbf{x} \in \mathbb{R}^n$ and $I \subseteq N$, we set

$$\mathbf{x}_I := \{u_1, \ldots, u_n\} \quad \text{where } u_j = \begin{cases} x_j & j \in I, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.9}$$

If $F(\mathbf{x})$ is a probability distribution function over $\mathbb{R}^n$, $F^{(I)}(x_{j_1},\dots,x_{j_{|I|}}) = F(\mathbf{x}_I)$ will be its $|I|$-dimensional marginal. Now we denote by $G_i(\cdot)$ the marginal distribution of F for $i = 1,\dots,n$ and we assume it to be continuous and strictly increasing. Furthermore we will denote by C the *connecting copula* of F:

$$C(\mathbf{y}) := F(G_1^{-1}(y_1),\dots,G_n^{-1}(y_n)). \tag{4.10}$$

Using a notation similar to (4.9), for $\mathbf{y} \in [0,1]^n$ we set

$$\mathbf{y}_I := \{v_1,\dots,v_n\} \quad \text{where } v_j = \begin{cases} y_j & j \in I, \\ 1 & \text{otherwise.} \end{cases}$$

In this way for the connecting copula $C_F^{(I)}$ of $F^{(I)}$ we can write

$$C_F^{(I)}(y_{j_1},\dots,y_{j_{|I|}}) = C(\mathbf{y}_I). \tag{4.11}$$

The following result can be seen as an analogue of several results presented in different settings (see in particular [**86**] and [**93**]).

PROPOSITION 4.2. *The utility function $U_{m,F}$ can also be written in the equivalent form*

$$U_{m,F}(\mathbf{x}) = \sum_{I\subseteq N} M_m(I)\mathbb{P}(\mathbf{T} \le \mathbf{x}_I). \tag{4.12}$$

PROOF. The proof amounts to a direct application of the inclusion-exclusion principle. Set $A_i = \{T_i \le x_i\}$ and we denote its complement by A_i^c; we also set $A_I = \cap_{i\in I} A_i$ and $\hat{A}_I = \cap_{i\notin I} A_i^c$. Then Equation (4.6) can be rewritten as

$$U_{m,F}(\mathbf{x}) = \sum_{I\subseteq N} m(I)\mathbb{P}(A_I \cap \hat{A}_I).$$

By a direct application of the inclusion-exclusion principle we have

$$U_{m,F}(\mathbf{x}) = \sum_{I\subseteq N} m(I) \sum_{J\subseteq N\setminus I} (-1)^{|J|}\mathbb{P}(A_I \cap A_J),$$

then

$$U_{m,F}(\mathbf{x}) = \sum_{I\subseteq N}\sum_{H\subseteq I}(-1)^{|H|} m(H)\mathbb{P}(A_I) = \sum_{I\subseteq N} M_m(I)\mathbb{P}(A_I),$$

which is the right hand side of (4.12). □

We now consider the function $U_{m,F}(G_1^{-1}(y_1),\dots,G_n^{-1}(y_n))$. In view of (4.10) we see that such a function depends on F only through the connecting copula C and it will be denoted by $\widehat{U}_{m,C}$. Furthermore, the quantities $G_1(x_1),\dots,G_n(x_n)$ can be given the meaning of utilities, thus $\widehat{U}_{m,C}$ becomes the *aggregation function* of the marginal utilities $y_1,\dots,y_n$.

COROLLARY 4.3. *In the case in which the one-dimensional distributions $G_1(x_1), \dots, G_n(x_n)$ of F are continuous and strictly increasing, one can also write*

$$\widehat{U}_{m,C}(\mathbf{y}) = \sum_{I \subseteq N} M_m(I) C(\mathbf{y}_I). \tag{4.13}$$

For any fixed pair (m, F), we now turn to considering the expected utility corresponding to the choice of a *prospect* $\mathbf{X} := (X_1, \dots, X_n)$ distributed according to $F_{\mathbf{X}}$:

$$\begin{aligned} \mathbb{E}_{\mathbf{X}} \left(U_{m,F}(\mathbf{X}) \right) &= \int_{\mathbb{R}^n} U_{m,F}(\mathbf{x}) \, dF_{\mathbf{X}}(\mathbf{x}) \\ &= \sum_{I \subseteq N} M_m(I) \mathbb{P}(\mathbf{T}_I \leq \mathbf{X}_I). \end{aligned} \tag{4.14}$$

By taking into account (4.14) and by interchanging the integration order, we can also write

$$\begin{aligned} \mathbb{E}_{\mathbf{X}}(U_{m,F}(\mathbf{X})) &= \mathbb{E}_{\mathbf{X}} \left[\mathbb{E}_{\mathbf{T}}(U_{m,\mathbf{T}}(\mathbf{X})) \right] \\ &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} M_m(I(\mathbf{t}, \mathbf{x})) \, dF_{\mathbf{X}}(\mathbf{x}) \right] dF_{\mathbf{T}}(\mathbf{t}). \end{aligned} \tag{4.15}$$

See also the logic scheme of Figure 4.1.

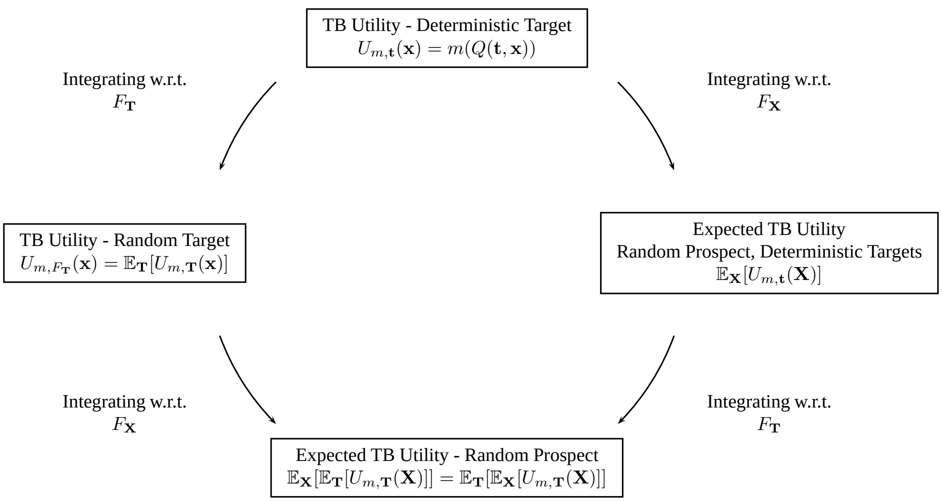

FIGURE 4.1. TB Utility Scheme

The formula (4.14) points out that, when evaluating the choice of a prospect $\mathbf{X}$, the random vector of interest is $\mathbf{D} = \mathbf{T} - \mathbf{X}$. Let us assume that the marginal distribution function of D_i, denoted by $H_i(\xi)$, is continuous and strictly increasing in $\xi = 0$ for $i = 1, \dots, n$, and put $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_n)$ with

$$\gamma_i = H_i(0). \tag{4.16}$$

Similarly to (4.11), let furthermore denote by $C^{(I)}_{F_D}$ the connecting copula of the marginal distribution corresponding to the coordinates subset $I \subseteq N$. Then (4.14) becomes

$$\widetilde{U}_{m,F}(\boldsymbol{\gamma}) := \mathbb{E}_{\mathbf{X}}\left(U_{m,F}(\mathbf{X})\right) = \sum_{I \subseteq N} M_m(I) C^{(I)}_{F_D}(\boldsymbol{\gamma})$$
$$= \sum_{I \subseteq N} M_m(I) C_{F_D}(\boldsymbol{\gamma}_I). \tag{4.17}$$

This formula highlights that, concerning the joint distribution of $\mathbf{D}$, we only need to specify the vector γ and $C_{F_D} = C^{(N)}_{F_D}$, the connecting copula of $\mathbf{D}$. From C_{F_D}, we can derive in fact the family of all marginal copulas $C^{(I)}_{F_D}$ by means of the formula (4.11) above.

3. Multi-Attribute TBA and Extensions of Fuzzy Measures

Let a capacity $m(\cdot)$ over 2^N and a n-dimensional copula $C : [0,1]^n \to [0,1]$ be given. For $\mathbf{y} \in [0,1]^n$, we can consider the *aggregation function*

$$V_{m,C}(\mathbf{y}) = \sum_{I \subseteq N} M_m(I) C(\mathbf{y}_I), \tag{4.18}$$

where $M_m(\cdot)$ denotes the Möbius transform of $m(\cdot)$ and $C(\mathbf{y}_I)$ is the connecting copula of $F^{(I)}$, see (4.11). We remind

DEFINITION 4.4. *An aggregation function $A : [0,1]^n \to [0,1]$ is a function non-decreasing in each component and that satisfies the boundary conditions $A(\mathbf{0}) = 0$ and $A(\mathbf{1}) = 1$.*

(see e.g. [**24**]). By the usual identification of $\{0,1\}^n$ with 2^N (where a subset $I \subseteq N$ is identified with its indicator function) one has, for $\overline{\mathbf{y}} \in \{0,1\}^n$ and for any copula C,

$$V_{m,C}(\overline{\mathbf{y}}) = m(I), \tag{4.19}$$

where $I = \{i | y_i = 1\}$. Thus any aggregation function of the form (4.18) can be seen as the extension to $[0,1]^n$ of the capacity $m(\cdot)$ defined over $\{0,1\}^n$. Extensions of a capacity over $\{0,1\}^n$ have been of interest in the fuzzy sets literature. Several properties of such extensions have been in particular studied by Kolesárová et al in [**86**]. In that paper the authors consider extensions of the form (4.18), where C is replaced by a more general aggregation function A. As corollaries of their general results, it follows that - in the special cases when A coincides with a copula C - $V_{m,C}$ is actually an aggregation function, and special properties of it are analyzed therein.

It is in particular noticed that, when C is the *product copula* one obtains the *Owen extension* and, when C is the *copula of comonotonicity*, namely

$$C(u_1, \ldots, u_n) = \min\{u_1, \ldots, u_n\}, \tag{4.20}$$

then one obtains the *Lovász extension*, or the Choquet integral of $\mathbf{y}$.

In the present framework, it is useful to give the aggregation function in (4.18) the form of a Riemann-Stiltjes in integral over $[0,1]^n$ as follows.

THEOREM 4.5. *Let m be a capacity over 2^N and C an n-dimensional copula. For $\mathbf{y} \in [0,1]^n$ one has*

$$V_{m,C}(\mathbf{y}) = \int_{[0,1]^n} m[Q(\mathbf{z},\mathbf{y})]\, dC(\mathbf{z}) \tag{4.21}$$

where $Q(\mathbf{z},\mathbf{y})$ is the set defined as in (4.4)*.*

PROOF. Let $I \subseteq N$. By definition of Q we have that $Q(\mathbf{z},\mathbf{y}) = I$ holds if and only if $z_i \leq y_i$ for $i \in I$ and $z_i > y_i$ for $i \notin I$. Hence

$$\begin{aligned} m[Q(\mathbf{z},\mathbf{y})] &= \sum_{I \subseteq N} m(I) \prod_{j \in I} \mathbb{1}_{\{z_j \leq y_j\}} \prod_{j \notin I} \mathbb{1}_{\{z_j > y_j\}} \\ &= \sum_{I \subseteq N} M_m(I) \prod_{j \in I} \mathbb{1}_{\{z_j \leq y_j\}}. \end{aligned}$$

By integrating this function over $[0,1]^n$ w.r.t. the probability measure associated to C, one has

$$\begin{aligned} \int_{[0,1]^n} m[Q(\mathbf{z},\mathbf{y})]\, dC(\mathbf{z}) &= \int_{[0,1]^n} \sum_{I \subseteq N} M_m(I) \prod_{j \in I} \mathbb{1}_{\{z_j \leq y_j\}}\, dC(\mathbf{z}) \\ &= \sum_{I \subseteq N} M_m(I) \int_{[0,1]^n} \prod_{j \in I} \mathbb{1}_{\{z_j \leq y_j\}}\, dC(\mathbf{z}) \\ &= \sum_{I \subseteq N} M_m(I) C(\mathbf{y}_I). \end{aligned} \tag{4.22}$$

□

REMARK 4.6. *Theorem 4.5 shows in which sense $V_{m,C}$ can be seen as a generalization of the Choquet integral. In fact $V_{m,C}$ reduces to a Choquet integral when C is the copula of comonotonicity.*

REMARK 4.7. *Consider now the case when $C_{\overline{\mathbf{z}}}$ is the probability distribution function degenerate over $\overline{\mathbf{z}} \in [0,1]^n$. In this case, as shown by* (4.21)*, $V_{m,C_{\overline{\mathbf{z}}}}$ reduces to*

$$V_{m,C_{\overline{\mathbf{z}}}}(\mathbf{y}) = m[Q(\overline{\mathbf{z}},\mathbf{y})]. \tag{4.23}$$

One can notice that, for any copula C, $V_{m,C_{\overline{\mathbf{z}}}}(\mathbf{y}) = V_{m,C}(\mathbf{w})$, where $\mathbf{w} \in \{0,1\}^n$ is defined by

$$w_i = \begin{cases} 1 & \textit{if} \quad z_i \leq y_i, \\ 0 & \textit{if} \quad z_i > y_i. \end{cases}$$

Notice also that Equation (4.23) *is just a different way to read the principle that led us to the Definition* (4.1) *of a TB multi-attribute utility function.*

As seen in the previous section, aggregation functions of the form (4.18) emerge in a natural way in the frame of TBA utilities. In such a frame the copula C takes a specific meaning as the copula that describes stochastic dependence properties of random vectors relevant in the decision problem at hand.

Let us consider the expected utility, associated to a multi-attribute prospect $\mathbf{X}$, of the target-based utility with target $\mathbf{T}$. As shown by expression (4.17), such expected utility has the form (4.18), when it is seen as a function of the quantities $\gamma_i = H_i(0)$, $i = 1, \dots, n$, introduced in (4.16). In such a case, C has then the meaning of the connecting copula of the vector $\mathbf{D} \equiv (T_1 - X_1, \dots, T_n - X_n)$.

Let furthermore $G_1, \dots, G_n$, the one-dimensional marginal distributions of $T_1, \dots, T_n$, be assumed continuous and strictly increasing and let C denote, this time, the connecting copula of $(T_1, \dots, T_n)$. Under these hypothesis $V_{m,C}(y_1, \dots, y_n)$ takes the meaning of an aggregation function $\widehat{U}_{m,C}(y_1, \dots, y_n)$ of the marginal utilities $y_1, \dots, y_n$, as (4.13) shows.

We then see that both the functions $\widehat{U}_{m,C_T}(\cdot)$ and $\widetilde{U}_{m,C_D}(\cdot)$, defined over $[0, 1]^n$, have the same formal expression (4.18) and are thus two different extensions of the capacity m. Starting from a same TBA utility function as in (4.13), they get different economic meanings. Both of them are definite integrals over $\mathbb{R}^n$, however. In particular, for $\widehat{U}_{m,C_T}$ and $\widetilde{U}_{m,F}$ we can obtain, as a corollary of Theorem 4.5,

PROPOSITION 4.8. *The aggregation functions $\widehat{U}_{m,C_T}$ and $\widetilde{U}_{m,F}$ are respectively given by*

$$\widehat{U}_{m,C_T}(\mathbf{y}) = \int_{[0,1]^n} m[Q(\mathbf{z}, \mathbf{y})] \, dC_T(\mathbf{z}). \tag{4.24}$$

$$\widetilde{U}_{m,F}(\boldsymbol{\gamma}) = \int_{[0,1]^n} m[Q(\mathbf{z}, \boldsymbol{\gamma})] \, dC_D(\mathbf{z}). \tag{4.25}$$

PROOF. As to the integral corresponding to $\widehat{U}_{m,C_T}(\mathbf{y})$, recall that, for $\mathbf{t} \in \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^n$, we had set $Q(\mathbf{t}, \mathbf{x}) := \{i | t_i \leq x_i\}$. By using formula (4.15), where F is the distribution function of the target vector $\mathbf{T}$, one has

$$U_{m,F}(\mathbf{x}) = \mathbb{E}_{\mathbf{T}}(U_{m,\mathbf{T}}(\mathbf{x})) = \mathbb{E}_{\mathbf{T}}[m(Q(\mathbf{T}, \mathbf{x}))] = \int_{\mathbb{R}^n} m[Q(\mathbf{t}, \mathbf{x})] \, dF_T(\mathbf{t}).$$

Notice now that $Q(\mathbf{t}, \mathbf{x})$ and $Q(\mathbf{z}, \mathbf{y})$ are exactly the same set, since $\{i | t_i \leq x_i\} = \{i | G^{-1}(t_i) \leq G^{-1}(x_i)\} = \{i | z_i \leq y_i\}$. Thus, recalling that $\mathbf{x} =$

$(G_1^{-1}(y_1), \ldots, G_n^{-1}(y_n))$, one has

$$\begin{aligned}\widehat{U}_{m,F}(\mathbf{y}) &= \int_{\mathbb{R}^n} m[Q(\mathbf{t}, G_1^{-1}(y_1), \ldots, G_n^{-1}(y_n))]\, dF_{\mathbf{T}}(\mathbf{t}) \\ &= \int_{I^n} m[Q(G_1^{-1}(z_1), \ldots, G_n^{-1}(z_n), G_1^{-1}(y_1), \ldots, G_n^{-1}(y_n))]\, dC(\mathbf{z}) \\ &= \int_{I^n} m[Q(\mathbf{z}, \mathbf{y})]\, dC(\mathbf{z}) = V_{m,C}(\mathbf{y}). \qquad (4.26)\end{aligned}$$

A similar argument can be used to prove (4.25). □

REMARK 4.9. *In the present frame, the Choquet integral admits the following economic interpretation. The choice of the copula of comonotonicity stands for the choice of a n-dimensional target, where all the random coordinates are just deterministic transformations of one and a same random variable. In this case $\widehat{U}_{m,C_T}(\mathbf{y})$ reduces to a Choquet integral.*

4. Reliability-structured TBA utilities

A very special class of capacities $m(\cdot)$ emerges as an immediate generalization of the case in (4.7) and is of interest in the frame of TBA utilities. For a brief introduction to the topic of reliability of systems we refer to Appendix A and to [**10**].

DEFINITION 4.10 (See [**20**], Definition 4). *A Target-Based utility function has a reliability structure when the capacity $m(\cdot)$ satisfies the condition*

$$m(I) \in \{0, 1\} \quad \textit{for all} \;\; I \in N.$$

Any such $m(\cdot)$ can then be seen as the *structure function* of a *coherent reliability system* S or, more generally, of a *semi-coherent* one (for further details see [**10**] and [**112**]).

We concentrate our attention on the case when both the coordinates $(T_1, \ldots, T_n)$ of the target and the coordinates $(X_1, \ldots, X_n)$ of the prospect are non-negative random variables that can then be interpreted as the vectors of the lifetimes of the components of S. The above reliability-based interpretation applies in a completely natural way, in this case.

For $\boldsymbol{\xi} \in \mathbb{R}_+^n$, we denote by $\tau(\boldsymbol{\xi})$ the lifetime of S when $\xi_1, \ldots, \xi_n$ are the values taken by such lifetimes, respectively. Then, as pointed out in [**96**], $\tau(\cdot)$ is a lattice polynomial function. Then (see [**15**]) it can be written both in a *disjunctive* and in a *conjunctive* form as a combination of the operators $\wedge$ and $\vee$ (see also [**96**], Proposition 2). When, in particular, the system admitting m as its structure function is coherent, such forms can be based on the *path sets* and the *cut sets* of the system (see again [**10**] and [**112**]).

The random variable $\tau(\mathbf{T})$ is the lifetime of S when the lifetimes of the components coincide with the coordinates of the DM's target and $\tau(\mathbf{X})$ is the lifetime of the system when the lifetimes of components coincide with the coordinates of a (random) prospect $\mathbf{X}$. Under such positions, the utility function $U_{m,F}(x_1,\ldots,x_n)$ can be read as a probability. More exactly

$$U_{m,F}(x_1,\ldots,x_n)=\mathbb{P}(\tau(\mathbf{T})\leq\tau(\mathbf{x})), \tag{4.27}$$

and the expected utility in (4.14) becomes

$$\mathbb{E}\left(U_{m,F}(\mathbf{X})\right)=\mathbb{P}(\tau(\mathbf{T})\leq\tau(\mathbf{X})). \tag{4.28}$$

We can then summarize as follows our conclusions. Consider a reliability-structured multi-attribute Target-Based utility $U_{m,F}(x_1,\ldots,x_n)$ with F the joint distribution function of n non-negative random variables and let $x_i\geq 0,\, i=1,\ldots,n$. Denote furthermore by

$$G_{\tau(\mathbf{T})}(\xi):=\mathbb{P}(\tau(\mathbf{T})\leq\xi)$$

the marginal distribution function of the lifetime $\tau(\mathbf{T})$. Then we have

PROPOSITION 4.11.

$$U_{m,F}(x_1,\ldots,x_n)=G_{\tau(\mathbf{T})}(\tau(\mathbf{x})). \tag{4.29}$$

This result shows that, in the reliability-structured case, a multi-attribute Target-Based utility $U_{m,F}$ reduces to a single-attribute Target-Based with a prospect $\tau(\mathbf{X})$ and a target $\tau(\mathbf{T})$. In particular the operator τ is a *mean* (see e.g. [**70**]): for $x>0$, $\tau(x,\ldots,x)=x$. Thus we obtain from (4.27) that the probability distribution function of $\tau(\mathbf{T})$ is given by

$$G_{\tau(\mathbf{T})}(\xi)=U_{m,F}(\xi,\ldots,\xi). \tag{4.30}$$

For a different but strictly related expression of $G_{\tau(\mathbf{T})}(\xi)$ see [**47**].

The formula (4.29) can be used in the two different directions: one can analyze questions about systems' reliability by using tools in the theory of aggregation operators and of extensions of capacities or, viceversa, different aspects of aggregation operators can be interpreted in terms of reliability practice, when the capacities are $\{0,1\}$-valued. In particular, the aggregation function $\widehat{U}_{m,C}$ in (4.13) can be given special interpretations in the present setting. From a technical point of view, in a reliability-structured frame, $G_1,\ldots,G_n$ are the one-dimensional marginal distributions of the components' lifetimes $T_1,\ldots,T_n$ of a system and C denotes the connecting copula of $\mathbf{T}$. By taking into account the equation (4.29) we obtain, for $\mathbf{y}\in[0,1]^n$,

$$\widehat{U}_{m,C}(\mathbf{y})=G_{\tau(\mathbf{T})}(\tau(G_1^{-1}(y_1),\ldots,G_n^{-1}(y_n))). \tag{4.31}$$

Notice that the operator $\tau(\mathbf{x})$ appearing in (4.27) and (4.31) is only determined by the capacity m, whereas the probability law $G_{\tau(\mathbf{T})}$ also depends on the copula C of F. In any case $\widehat{U}_{m,C}(\mathbf{y})$ is an integral, w.r.t. the capacity m, and the function to be integrated depends on C.

We also notice that, from a purely mathematical viewpoint, m can be paired with any copula C. Any capacity m, for instance, can be paired with the comonotonicity copula to obtain that $\widehat{U}_{m,C}(y)$ is a Choquet Integral. We also notice in this respect that, in such case, $\widehat{U}_{m,C}(y)$ is a lattice polynomial as well. From the economic point of view, on the contrary, imposing conditions describing the attitudes towards risk by part of a Decision Maker, creates some constraints on the choice of the pair (m, C). See also the brief discussion in the next section.

4.1. Symmetric Reliability-Structured Cases. Here we consider special conditions of invariance with respect to permutations. First we look at the very restrictive, but important, case of *symmetric,* reliability-structured capacities. The reliability systems admitting permutation-invariant structure functions ϕ are those of the type k-out-of-n. More precisely, a system is k-out-of-n when, for $\mathbf{x} \in \{0,1\}^n$, its structure function has the form

$$\phi_{k:n}(x_1, \dots, x_n) = \begin{cases} 1 & \text{if } \sum_i x_i \geq k, \\ 0 & \text{if } \sum_i x_i < k. \end{cases} \tag{4.32}$$

This is then the case of a system which is able to work as far as at least k of its components are working or, in other words, which fail at the instant of the $(n-k+1)-$th failure among its components. In (4.32), the structure function is seen as a function over N. Equivalently, when $\phi_{k:n}$ is seen as a set function, the value $\phi_{k:n}(I)$ is 0 or 1, only depending on the cardinality of I being larger or smaller than k.

PROPOSITION 4.12. *In the case of a k-out-of-n capacity $m = \phi_{k:n}$, we have*

$$U_{m,F}(\mathbf{x}) = \sum_{I \subseteq N, |I| \geq k} (-1)^{|I|-k} \binom{|I|-1}{|I|-k} \mathbb{P}\left(\mathbf{T} \leq \mathbf{x}_I\right).$$

PROOF. Recall Equation (4.12) and notice that, for $m = \phi_{k:n}$, the coefficients of the Möbius transform are given by

$$M_m(I) = \begin{cases} (-1)^{|I|-k} \binom{|I|-1}{|I|-k} & |I| \geq k, \\ 0 & \text{otherwise.} \end{cases}$$

□

It is clear that, in the case of a k-out-of-n systems, we have that

$$\tau(\boldsymbol{\xi}) = \xi_{(n-k+1)},$$

where $\xi_{(1)}, \dots, \xi_{(n)}$ denote the order statistics of $\xi_1, \dots, \xi_n$ and the formula (4.29) takes the special form

$$U_{\phi_{k:n},F}(\mathbf{x}) = G_{T_{(n-k+1)}}(x_{(n-k+1)}).$$

From (4.30), we in particular obtain the probability law of $T_{(n-k+1)}$:

$$G_{T_{(n-k+1)}}(\xi) := \mathbb{P}(T_{(n-k+1)} \leq \xi) = U_{\phi_{k:n},F}(\xi, \dots, \xi). \tag{4.33}$$

A different remarkable case of reliability-structured TBA utilities is obtained by imposing the condition of permutation-invariance over the joint distribution F, rather than over the capacity m. This is the case when

$T_1, \ldots, T_n$, the coordinates of the target $\mathbf{T}$, are assumed to be non-negative *exchangeable* random variables, namely the joint distribution $F(x_1, \ldots, x_n)$ is assumed invariant under permutations of its arguments $x_1, \ldots, x_n$. In this case the concept of *signature* of the system enters in the expression of the utility function $U_{m,F}$.

Given the structure function $\phi : \{0,1\}^n \to \{0,1\}$ of a semi-coherent system, the signature $\mathbf{s}^{(\phi)} = \mathbf{s} = (s_1, \ldots, s_n)$ is a probability distribution over $N = \{1, \ldots, n\}$ (as a basic reference, see e.g. [**117**]). For $j = 1, \ldots, n$, consider the events

$$E_j := \left(\tau(\mathbf{T}) = T_{(j)}\right),$$

with $T_1, \ldots, T_n$ denoting again the lifetimes of the components and $\tau(\mathbf{T})$ the lifetime of the system. When $T_1, \ldots, T_n$ are such that

$$\mathbb{P}\left(T_{i'} = T_{i''}, \text{ for some } i', i''\right) = 0, \tag{4.34}$$

$E_1, \ldots, E_n$ are exhaustive and pair-wise disjoint, and we have

$$\sum_{j=1}^{n} \mathbb{P}(E_j) = 1.$$

The components $s_1, \ldots, s_n$ of the signature are defined as $s_j = \mathbb{P}(E_j)$. It is easy to prove that, when $T_1, \ldots, T_n$ are exchangeable, the following properties hold:

- **a):** $\mathbf{s}^{(\phi)}$ does not depend on the joint probability distribution of the targets $T_1, \ldots, T_n$;
- **b):** For $\xi > 0$ and $j = 1, \ldots, n$, the event $(T_{(j)} \leq \xi)$ is stochastically independent of $E_1, \ldots, E_n$.

By the formula of total probabilities we then can write, for any $\xi > 0$,

$$\begin{aligned}
\mathbb{P}(\tau(\mathbf{T}) \leq \xi) &= \sum_{j=1}^{n} \mathbb{P}(E_j)\mathbb{P}(\tau(\mathbf{T}) \leq \xi | E_j) \\
&= \sum_{j=1}^{n} \mathbb{P}(\tau(\mathbf{T}) = T_{(j)})\mathbb{P}(\tau(\mathbf{T}) \leq \xi | \tau(\mathbf{T}) = T_{(j)}) \\
&= \sum_{j=1}^{n} s_j^{(\phi)} \mathbb{P}(T_{(j)} \leq \xi).
\end{aligned} \tag{4.35}$$

By the property **a)** the signature $\mathbf{s}^{(\phi)}$ is a combinatorial invariant of the system. See in particular [**97**] for the relations between the signature $\mathbf{s}^{(\phi)}$ and the "reliability function" of the system in case of i.i.d. components. For a discussion about the relations between $\mathbf{s}^{(\phi)}$ and symmetry properties see also [**129**]).

In view of (4.35) the signature $\mathbf{s}^{(\phi)}$ has a role in the representation of the utility function $U_{\phi,F}$ when F is exchangeable. By (4.30) we obtain

PROPOSITION 4.13. *Let F be an exchangeable joint distribution function over $\mathbb{R}^n_+$, satisfying the condition* (4.34). *For any reliability-structured capacity $m : 2^N \to \{0,1\}$ and for $\mathbf{x} \in \mathbb{R}^n_+$, one has*

$$U_{m,F}(\mathbf{x}) = \sum_{j=1}^{n} s_j^{(m)} \mathbb{P}(T_{(j)} \leq \tau(\mathbf{x})). \tag{4.36}$$

The terms $s_j^{(m)}$ and $\tau(\mathbf{x})$ in (4.36) are determined by m, whereas F determines the probability law of $T_{(j)}$, for $j = 1, \dots, n$. The formula (4.35) is a special case of (4.36): for $\mathbf{x} = (\xi, \dots, \xi)$ we obtain once more

$$\begin{aligned} G_{\tau(\mathbf{T})}(\xi) = U_{m,F}(\xi, \dots, \xi) &= \sum_{j=1}^{n} s_j^{(m)} \mathbb{P}(T_{(j)} \leq \xi) \\ &= \sum_{j=1}^{n} s_j^{(m)} U_{\phi_{k:n},F}(\xi, \dots, \xi). \end{aligned}$$

5. TBA utilities and attitudes toward goods and risk

Here we think of a Decision Maker who describes her/his attitudes towards n goods $\mathcal{G}_1, \dots, \mathcal{G}_n$ through a capacity m and defines her/his utility by choosing a target $\mathbf{T}$ with joint distribution function F. Thus $U_{m,F}(\mathbf{x})$ evaluates the satisfaction of the DM in receiving the quantity x_1 for the good $\mathcal{G}_1$, x_2 for the good $\mathcal{G}_2$ and so on. Different properties with economic meaning of a multi-attribute utility function can take a special form in the TBA case and in the reliability-structured TBA case, more in particular. One should analyze how can different properties be influenced by the choice of the parameters m, F or, in other terms, which constraints on the pair (m, F) are induced by fixing the attitudes of the DM. In this Section, we concentrate our attention on the basic concepts of supermodularity and submodularity (see [**136, 137**]) and present some related comments.

For a function $U : \mathbb{R}^n \to \mathbb{R}$ and for $\mathbf{x}', \mathbf{x}'' \in \mathbb{R}^n$, set

$$\nu^U(\mathbf{x}', \mathbf{x}'') := U(\mathbf{x}' \vee \mathbf{x}'') + U(\mathbf{x}' \wedge \mathbf{x}'') - U(\mathbf{x}') - U(\mathbf{x}''). \tag{4.37}$$

DEFINITION 4.14. *The function U is supermodular when $\nu^U(\mathbf{x}', \mathbf{x}'') \geq 0$ for all $\mathbf{x}', \mathbf{x}'' \in \mathbb{R}^n$, and submodular when $\nu^U(\mathbf{x}', \mathbf{x}'') \leq 0$. If U is both supermodular and submodular, then it is called modular.*

Under the condition that the function U is twice differentiable, an equivalent formulation in terms of the second order derivatives of U can be given. In particular the condition of supermodularity is given by

$$\partial^2 U(\mathbf{x}) / \partial x_i \partial x_j \geq 0 \tag{4.38}$$

for all $\mathbf{x} \in \mathbb{R}^n$ and $i \neq j$, $i, j = 1, \dots, n$.

For a utility function, it is well-known that supermodularity describes the case of *complementary* goods (see [**50, 118, 137**]), while submodularity is associated to *substitutable* goods. Two or more goods are called *complementary* if "they have little or no value by themselves, but they are more valuable when combined together", while they are called *substitutable* when "each of them satisfies the same need of the DM that the other good fulfills". In these settings we can say that a collection of goods are complements (and each pair is said to be complementary) if they have a real-valued supermodular utility function (Bulow et al. [**23**] use the term *strategic complements* to describe any two *activities* i and j for which formula (4.38) holds).

As a related interpretation, the properties of *supermodularity*, *submodularity*, and *modularity* of a multi-attribute utility U respectively describe, in an analytic language, the properties of *correlation affinity*, *correlation aversion*, and *correlation neutrality* (see e.g. [**136**] and [**137**]). In particular the concept of submodularity gives rise to a specific definition of *greater correlation* between two joint probability distributions (see Definition 4 in [**53**]).

Let us now come to TB utilities and to related problems of prospects choosing. We are essentially interested in decision problems where the following objects are considered to be fixed: the capacity m, the marginal distributions $G_1, \ldots, G_n$ of the targets' components $T_1, \ldots, T_n$, and the marginal distributions $G_{X_1}, \ldots, G_{X_n}$ for the components of the prospect. Since we have assumed stochastic independence between $\mathbf{X}$ and $\mathbf{T}$, the marginal probability distribution function $H_i(\cdot)$ of $D_i = T_i - X_i$ is suitably obtained by convolution from G_i and G_{X_i}. Then, at least in principle, the vector $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)$ is known, where $\gamma_i = H_i(0)$. The DM is supposed to declare the copula C of the target vector $\mathbf{T}$ and, on this basis, to select a copula for the random prospect $\mathbf{X}$. The choice of a prospect then amounts to the choice of a copula $C_{\mathbf{D}}$ for the vector $\mathbf{D} = \mathbf{T} - \mathbf{X}$.

For a TB utility function U, the expression in (4.37) becomes

$$\nu^U(\mathbf{x}', \mathbf{x}'') = \sum_{I \subseteq N} M_m(I) \nu^F(\mathbf{x}'_{\mathbf{I}}, \mathbf{x}''_{\mathbf{I}}) \tag{4.39}$$

for any pair of vectors $\mathbf{x}', \mathbf{x}'' \in \mathbb{R}^n$. The notation $\mathbf{x}'_{\mathbf{I}}, \mathbf{x}''_{\mathbf{I}}$ is as used in (4.9). Then the conditions of supermodularity, or submodularity, become

$$\sum_{I \subseteq N} M_m(I) \nu^F(\mathbf{x}'_{\mathbf{I}}, \mathbf{x}''_{\mathbf{I}}) \gtreqless 0. \tag{4.40}$$

Let the DM manifest correlation aversion or correlation affinity. Namely she/he wants to use a submodular, or supermodular, utility function. Of course correlation aversion/affinity concerns attitudes toward dependence among the coordinates of the prospect. On the other hand, for fixed m, the properties of supermodularity and submodularity are expressed through the choice of the connecting copula C for the target $\mathbf{T}$. Such properties are generally determined by the interplay between m and F. In conclusion, we are interested in conditions on the pair (m, F) for which condition (4.40) holds. In this direction we now discuss some special cases.

First of all we consider the case in which the capacity m is *totally monotone*. We remind that a capacity m is said totally monotone if its Möbius transform $M(I)$ is positive for all $I \subseteq N$ (see [**66**]). Since all the multivariate distribution functions are supermodular, we immediately see from (4.40) that if m is totally monotone, the utility function $U_{m,F}$ is supermodular whatever the distribution function F of the target is. So, in this special case, the condition of supermodularity is completely determined by the capacity m.

A further interesting case is met when the capacity m is additive: in this situation the interplay among variables has no effect on the overall amount of the utility $U_{m,F}$. In fact, the formula for $U_{m,F}$ reduces to

$$U_{m,F}(\mathbf{x}) = \sum_{i=1}^{n} m_i \mathbb{P}(T_i \leq x_i),$$

with $m_1 + \ldots + m_n = 1$. The expression in the r.h.s. represents an Ordered Weighted Average (see [**66**]) of the marginal distributions of the targets T_i. It is immediate to notice that $U_{m,F}(\mathbf{x})$ is modular for any choice of F. Furthermore it does not depend on the copula C of F. We notice that, in this case, the expected value of the utility $\mathbb{E}[U_{m,F}(\mathbf{X})]$ (see formulas (4.14) and (4.17)) for a fixed prospect $\mathbf{X}$ becomes $\widetilde{U}_{m,F}(\boldsymbol{\gamma}) = \sum_{i=1}^{n} m_i \gamma_i$.

Another likely situation is that in which the DM only considers interactions among small groups of goods, say k at most. In other words the DM is not interested in how they behave when considered in groups of cardinality larger than k. This condition leads to the choice of a $k-$additive capacity (see e.g. [**65**]). More in details

DEFINITION 4.15. *A capacity m is said k-additive if the coefficients of its Möbius transform M_m satisfy the condition $M_m(I) = 0$ for all I such that $|I| > k$, and $M_m(I) \neq 0$ for at least one element I with $|I| = k$.*

The assumption of k-additivity generally simplifies the study of the utility function. Under this hypothesis condition (4.40) reduces to

$$\nu^U(\mathbf{x}', \mathbf{x}'') = \sum_{I:|I|=2,\ldots,k} M_m(I) \nu^F(\mathbf{x_I}', \mathbf{x_I}'') \gtreqless 0.$$

We notice, in any case, that the possible validity of the conditions of submodularity and supermodularity generally depends on both the capacity m and the distribution F. In particular, in the case $k = 2$, a sufficient condition for supermodularity (submodularity) reads $M_m(\{i, j\}) \geq 0 \, (\leq 0)$, for all $i \neq j$.

Also of interest is the special case of reliability-structured utility functions, that we have considered in the previous section. First we notice that m being $\{0, 1\}$-valued has a direct economic interpretation: like a *binary* system, that can be *up* or *down* according to the current state (*up* or *down*) of each of its n components, so the DM is completely *satisfied* or completely *unhappy* according to which is the subset of targets that have been achieved. Cases where such utilities can be of economic relevance are discussed in [**20**]. Also, the special forms of TB utilities with m describing series systems or parallel systems are discussed there: these are the cases when m is the minimal or the maximal capacity, respectively, and correspond to the two extreme cases of *perfect* complementarity and *perfect* substitutability. In such cases we encounter supermodularity and submodularity, respectively, independently of the form of F. In all the other cases the condition of supermodularity, or submodularity respectively, reads

$$G_{\tau(\boldsymbol{T})}(\tau(\mathbf{x}' \vee \mathbf{x}'')) + G_{\tau(\boldsymbol{T})}(\tau(\mathbf{x}' \wedge \mathbf{x}'')) \\ -G_{\tau(\boldsymbol{T})}(\tau(\mathbf{x}')) - G_{\tau(\boldsymbol{T})}(\tau(\mathbf{x}'')) \gtreqless 0. \tag{4.41}$$

The validity of such a condition depends on the behavior of both the capacity m and the distribution function F of the targets. Notice that, when $\mathbf{T}$ is exchangeable, $G_{\tau(\boldsymbol{T})}$ is of the form (4.35), then condition (4.41) becomes

$$\sum_{j=1}^{n} s_j^{(\phi)} \cdot \big[G_{(j)}(\tau(\mathbf{x}' \vee \mathbf{x}'')) + G_{(j)}(\tau(\mathbf{x}' \wedge \mathbf{x}'')) \\ -G_{(j)}(\tau(\mathbf{x}')) - G_{(j)}(\tau(\mathbf{x}''))\big] \gtreqless 0,$$

where $G_{(j)}(x) = \mathbb{P}(T_{(j)} \leq x)$.

Still concerning the properties of supermodularity/submodularity, a very clear situation is met in the special case $n = 2$. We first notice that, in this case, formula (4.12) becomes

$$U_{m,F}(x_1, x_2) = M_1\, \mathbb{P}(T_1 \leq x_1) + M_2\, \mathbb{P}(T_2 \leq x_2) \\ + M_{1,2}\, \mathbb{P}(T_1 \leq x_1, T_2 \leq x_2), \tag{4.42}$$

where we have used, for m and M_m the shorter notation $m_1 = m(\{1\})$, $M_1 = M_m(\{1\})$, and so on. As a strongly simplifying feature of the present case, the utility function $U_{m,F}$ in (4.42) is, in any case, supermodular or submodular. In fact condition (4.39) reads

$$\nu^U(\mathbf{x}', \mathbf{x}'') = M_{1,2}\, \nu^F(\mathbf{x}', \mathbf{x}'').$$

Hence, since any joint distribution function F is supermodular, submodularity and supermodularity are respectively equivalent to the conditions

$$M_{1,2} \leq 0 \quad \text{and} \quad M_{1,2} \geq 0, \tag{4.43}$$

or, in terms of m,

$$m_1 + m_2 \geq 1 \quad \text{and} \quad m_1 + m_2 \leq 1. \tag{4.44}$$

Focus now attention, in particular, to the cases of *perfect complementarity* and *perfect substitutability*. The first one is equivalent to the condition $m_1 = m_2 = 0$ or, equivalently, $M_{1,2} = 1$, and describes the maximal possible affinity to correlation of the DM. Here the expression of the utility $U_{m,F}$ reduces to

$$U_{m,F}(x_1, x_2) = F(x_1, x_2), \tag{4.45}$$

which is exactly the joint distribution function of the two-dimensional target. In the opposite case, the utility reduces to $U_{m,F}(x_1, x_2) = G_1(x_1) + G_2(x_2) - F(x_1, x_2)$ or, analogously, $\widehat{U}_{m,C}(y_1, y_2) = C^*(y_1, y_2)$, where C^* stands for the dual of the copula C (for further details see [**106**]). All other cases can be grouped mainly into two sets, the strictly supermodular ones, with $m_1 + m_2 < 1$, and the submodular ones, with $m_1 + m_2 > 1$. Finally we notice a region of neutrality, along the diagonal corresponding to $m_1 + m_2 = 1$: this is the case of additivity of the capacity m, already discussed above. All these cases are summarized in Figure 4.2.

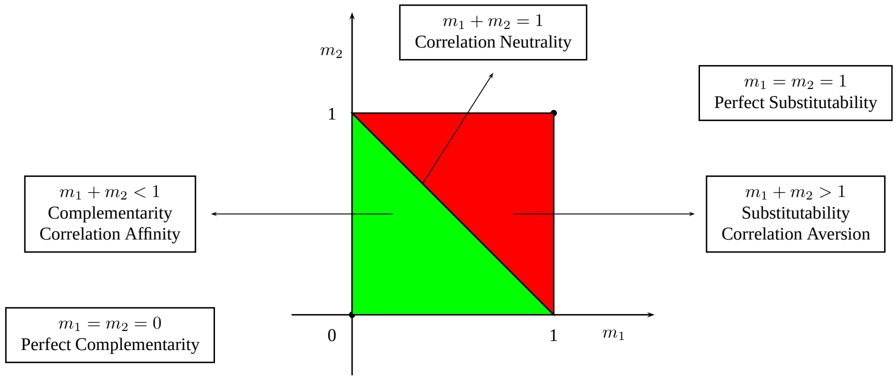

FIGURE 4.2. Scheme for complementarity and substitutability among two goods depending on their utility parameters m_1, m_2

We already noticed that, w.r.t. the capacity m, the aggregation function $\widehat{U}_{m,C}$ is an integral of m, depending on C, the connecting copula of F. For a fixed m, there is no restriction in the choice of C, from a purely mathematical point of view. We can see, on the contrary, that certain constraints on the pair (m, C) can arise from an economic point of view, depending on the attitudes of our Decision Maker. In other words, the type of integral of m, that the DM is led to consider as an aggregation $\widehat{U}_{m,C}$, depends on m itself once the attitudes of the DM have been fixed. As a simple example, let us consider the case of perfect complementarity in (4.45). In such case $\widehat{U}_{m,C}$ becomes $\widehat{U}_{m,C}(y_1, y_2) = C(y_1, y_2)$. Thus the aggregation function $\widehat{U}_{m,C}$ will grow with the growth of the copula C. This entails that a DM, who will manifest risk aversion besides correlation affinity, will choose the target which exhibits the greatest possible copula. Thus the most profitable choice is the maximal copula, $C(u, v) = u \wedge v$, namely the one of comonotonicity. Similar arguments can be developed for the study of the extreme opposite case, $m_1 = m_2 = 1$ $(M_{1,2} = -1)$.

6. Summary and concluding remarks

By introducing the target-based approach, Bordley and Li Calzi in [**19**] and Castagnoli and Li Calzi in [**28**] had in particular developed a new way to look at utility functions, and related extensions, in the field of decision problems under risk. In those papers, emphasis was given to the single-attribute case where, practically, there is no loss of generality in considering target-based utilities. As to the multi-attribute case, a treatment proposed a few years later (in [**20, 138, 139**]) had further added some new ideas to the field. In fact, the proposed extension is something different from the single-attribute definition. Actually, a direct generalization of the latter would lead one to consider much too special and restrictive forms of utilities, as we have remarked in the Introduction.

A principle of individual choice, clearly enucleated in [**20**], is at the basis of the given definition of multi-attribute target-based utilities. This principle is indeed quite natural and is related to the evaluation, by part of a Decision Maker, about the relative importance attributed to any possible subset of achieved targets. It emerges then that such an evaluation depends on the individual propensity toward the possible “coalitions” of attributes and that it is related with the concept of capacity.

Starting from the latter observation, we have formally considered a multi-attribute target-based utility (Definition 4.1) as a pair (m, F), where m is a capacity over $2^N = \{0, 1\}^n$ and F is a probability distribution function over $\mathbb{R}^n$. On this basis, we have pointed out that the theory of multi-attribute target-based utilities can hinge on a formal apparatus, provided by the field of fuzzy measures, extensions of fuzzy measures, and fuzzy, or universal, integrals. On the other hand, multi-attribute target-based utilities give rise to applications of the concepts and of results in this field. In particular, under special conditions, the arguments and results presented in [**86**] can have an interpretation useful to an heuristic view of the differences among various fuzzy integrals. As we have briefly recalled in Section 3, operators of the form

$$V_{m,A}(\mathbf{y}) = \sum_{I \subseteq N} M_m(I) A(\mathbf{y}_I) \tag{4.46}$$

have been analyzed in [**86**] as extensions of capacities m over 2^N. Generally speaking, the function A appearing in (4.46) is an aggregation function. In our frame, interest is concentrated on the special case when A is replaced by an n-dimensional copula C. The effect of such a particular condition is two-fold: on the one hand, it makes $V_{m,A} = V_{m,C}$ to have, itself, the properties of an aggregation function. On the other hand, it gives $V_{m,C}$ the meaning of an aggregation of marginal utilities; the special form of aggregation depends on the special type of stochastic dependence that is assumed among the coordinates of the target. An extreme condition of dependence with a special decisional meaning of its own, namely positive comonotonicity, lets such an aggregation coincide with a Choquet integral. We thus see the aggregation functions $\widehat{U}_{m,C}$ as a natural class of operators generalizing such integrals.

Concerning Choquet integrals, it is well known that they have been very widely studied and discussed in the past literature concerning utilities and decision under risk. In particular, in [**30**] and [**121**], it is shown how this concept allows one to build a quite general model of decision making under uncertainty, generalizing the Expected Utility model, in the frame of single-attribute decisions. We point out that its role in the present study appears under a rather different form: it is not used in fact to explain a general principle for decisions under uncertainty. It emerges as an extremely special case, just in the frame of Expected Utility. However its meaning in the TB Approach is peculiar of the multi-attribute case.

In multi-attribute decision problems under risk, the profile of a Decision Maker can be specified by taking into account different types of attitudes and forms of behavior, such as risk-aversion (or risk-affinity), correlation-aversion (or correlation-affinity), cross-prudence, etc. Generally these conditions are described in terms of qualitative properties of the utility functions (see e.g. [**45, 46, 51**]).

Let us come to the specific case of multi-attribute utility functions, that we had identified with the pairs (m, F). As a challenging program for future research, one should detail how the mentioned qualitative properties of utility functions determine (or are determined by) the form of m and F and reciprocal relations between them. For a DM with given attitudes toward risk, the choice of F - and then, in particular, of the copula C - is not completely free, but is influenced by the form of m itself. In the above Section 5, we have considered some significant special cases and sketched some conclusion in this direction. A more general analysis may result from future achievements about qualitative descriptions of target-based utilities.

Further research suggested by our work also concerns specific aspects of multivariate copulas. As shown by formula (4.17), the analysis of the present approach would benefit from new results concerning the connecting copula of the vector $\mathbf{D}$ obtained as the difference between the vectors $\mathbf{T}$ and $\mathbf{X}$. Here we have assumed stochastic independence.

More complex arguments would be involved in the cases when the possibility of some correlation between the vectors $\mathbf{T}$ and $\mathbf{X}$ is admitted. Some specific aspects in this direction, for the special case $n = 1$, have been dealt with in [**38**].

CHAPTER 5

Stochastic Precedence, Stochastic Orderings and connections to Utility Theory

The concept of *stochastic precedence* between two real-valued random variables has often emerged in different applied frameworks. It finds applications in various statistical contexts, like testing and sampling, reliability modeling, tests for distributional equality versus various alternatives (see, for example, [**5, 18, 118**]). Furthermore, this concept has been studied in the probabilistic context of Markov models for waiting times to occurrences of words in random sampling of letters from an alphabet (for references, see [**40, 41, 42, 43**]). Further applications can also arise in the fields of reliability and in the comparison of pool obtained by two opposite coalitions.

Motivations for our study arise, in particular, from the frame of Target-Based Approach in decisions under risk. In the previous Chapter we developed this model for multi-dimensional attributes, under the assumption of stochastic independence between Targets and Prospects. In this Chapter our analysis concerns the one-dimensional case, but with the assumption of stochastic dependence.

To our purposes, we introduce a slightly more general, and completely natural, concept of stochastic precedence and analyze its relations with the usual notions of stochastic ordering. Such a study leads us to introducing some special classes of bivariate copulas, namely the classes $\mathcal{L}_\gamma$. Properties of such classes are useful to describe the behavior of the Target-Based model under changes in the connecting copulas of the random variables, especially regarding their properties of symmetry and dependence. Examples are provided in this direction.

More precisely the structure of the Chapter is as follows. In Section 1 we introduce the concept of *generalized stochastic precedence* and the classes $\mathcal{L}_\gamma$. In Section 2, we analyze the main aspects of such classes and present a related characterization. Connections to measures of asymmetry of copulas are analyzed in Section 3, where we introduce a weak measure of non-exchangeability for the copulas in $\mathcal{L}_\gamma$. Some further basic properties of this class are detailed in Section 4, where a few examples are also presented. Finally, in Section 5, we trace connections to Target-Based utilities in the case of stochastic dependence between targets and prospects.

1. Basic Definitions

Let X_1, X_2 be two real random variables defined on a same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We will denote by F the joint distribution function and by G_1, G_2 their marginal distribution functions, respectively. For the sake of notational simplicity, we will initially concentrate our attention on the case when G_1, G_2 belong to the class $\mathcal{G}$ of all the probability distribution functions on the real line, that are continuous and strictly increasing in the domain where they are positive and smaller than one. As we shall see later, we can also consider more general cases, but the present restriction allows us to simplify the formulation and the proofs of our results. In order to account for some cases of interest with $\mathbb{P}(X_1 = X_2) > 0$, we will not assume that the distribution function F is absolutely continuous.

The random variable X_1 is said to *stochastically precede* X_2 if $\mathbb{P}(X_1 \leq X_2) \geq 1/2$, written $X_1 \preceq_{sp} X_2$. The interest of this concept for applications has been pointed out several times in the literature (see in particular [**6**], [**18**] and [**104**]). We recall the reader's attention on the fact that stochastic precedence does not define a stochastic order in that, for instance, it is not transitive. However it can be considered in some cases as an interesting condition, possibly alternative to the usual *stochastic ordering* $X_1 \preceq_{st} X_2$, defined by the inequality $G_1(t) \geq G_2(t)$, $\forall t \in \mathbb{R}$, see [**124**].

When X_1, X_2 are independent the implication $X_1 \preceq_{st} X_2 \Rightarrow X_1 \preceq_{sp} X_2$ holds (see [**6**]). It is also easy to find several other examples of bivariate probability models where the same implication holds. For instance the condition $X_1 \preceq_{st} X_2$ even entails $\mathbb{P}(X_1 \leq X_2) = 1$ when X_1, X_2 are *comonotonic* (see e.g. [**106**]), i.e. when $\mathbb{P}(X_2 = G_2^{-1}(G_1(X_1))) = 1$. On the other hand, cases of stochastic dependence can be found where the implication $X_1 \preceq_{st} X_2 \Rightarrow X_1 \preceq_{sp} X_2$ fails. A couple of examples will be presented in Section 4. See also Proposition 5.10. On the other hand the frame of words' occurrences produces, in a natural way, examples in the same direction, see e.g. [**40**].

In this framework we replace the notion $X_1 \preceq_{sp} X_2$ with the generalized concept defined as follows

DEFINITION 5.1 (Generalized Stochastic Precedence). *For any given* $\gamma \in [0, 1]$, *we say that* X_1 *stochastically precedes* X_2 *at level* γ *if the condition* $\mathbb{P}(X_1 \leq X_2) \geq \gamma$ *holds. This will be written* $X_1 \preceq_{sp}^{(\gamma)} X_2$.

Let $\mathcal{C}$ denote the class of all bivariate copulas (see also [**76, 106**]). Several arguments along the Chapter will be based on the concept of bivariate copula. The class of all bivariate copulas will be denoted by $\mathcal{C}$. We recall that the pair of random variables X_1, X_2, with distributions G_1, G_2, respectively, admits $C \in \mathcal{C}$ as its *connecting* copula whenever its joint distribution function is given by

$$F(x_1, x_2) = C(G_1(x_1), G_2(x_2)). \tag{5.1}$$

It is well known, by Sklar's Theorem 2.4 that the connecting copula is unique when G_1 and G_2 are continuous. We will use the notation

$$A := \{(x_1, x_2) \in \mathbb{R}^2 \,:\, x_1 \leq x_2\}, \tag{5.2}$$

so that we write

$$\mathbb{P}(X_1 \leq X_2) = \int_A dF(x_1, x_2) = \int_{\mathbb{R}^2} \mathbb{1}_A(x_1, x_2)\, dF(x_1, x_2). \tag{5.3}$$

For given $G_1, G_2 \in \mathcal{G}$ and $C \in \mathcal{C}$ we also set

$$\eta(C, G_1, G_2) := \mathbb{P}(X_1 \leq X_2), \tag{5.4}$$

where X_1 and X_2 are random variables with distributions G_1, G_2 respectively, and connecting copula C. Thus the condition $X_1 \preceq_{sp}^{(\gamma)} X_2$ can also be written $\eta(C, G_1, G_2) \geq \gamma$.

Suppose now that X_1, X_2 satisfy the condition $X_1 \preceq_{st} X_2$. As a main purpose of this Chapter, we give a lower bound for the probability $\mathbb{P}(X_1 \leq X_2)$ in terms of the stochastic dependence between X_1 and X_2 or, more precisely, in terms of conditions on the integral $\int_{A \cap [0,1]^2} dC$. More specifically we will analyze different aspects of the special classes of bivariate copulas, defined as follows.

DEFINITION 5.2. *For $\gamma \in [0, 1]$, we denote by $\mathcal{L}_\gamma$ the class of all copulas $C \in \mathcal{C}$ such that*

$$\eta(C, G_1, G_2) \geq \gamma \tag{5.5}$$

for all $G_1, G_2 \in \mathcal{G}$ with $G_1 \preceq_{st} G_2$.

Concerning the role of the concept of copula in our study, we point out the following simple facts. Consider the random variables $X_1' = \phi(X_1)$ and $X_2' = \phi(X_2)$ where $\phi : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function. Thus $X_1' \preceq_{st} X_2'$ if and only if $X_1 \preceq_{st} X_2$ and $X_1' \preceq_{sp}^{(\gamma)} X_2'$ if and only if $X_1 \preceq_{sp}^{(\gamma)} X_2$. At the same time the pair X_1', X_2' also admits the same connecting copula C.

2. A characterization of the class $\mathcal{L}_\gamma$

This Section will be devoted to providing a characterization of the class $\mathcal{L}_\gamma$ (see Theorem 5.7 and 5.8) along with related discussions. We start by detailing a few basic properties of the quantities $\eta(C, G_1, G_2)$, for $G_1, G_2 \in \mathcal{G}$ and $C \in \mathcal{C}$. In view of the condition $G_1, G_2 \in \mathcal{G}$ we can use the change of variables $u = G_1(x_1)$, $v = G_2(x_2)$. Thus we can rewrite the integral in (5.3) according to the following

PROPOSITION 5.3. *For given $G_1, G_2 \in \mathcal{G}$ and $C \in \mathcal{C}$, one has*

$$\eta(C, G_1, G_2) = \int_{[0,1]^2} \mathbb{1}_A(G_1^{-1}(u), G_2^{-1}(v))\, dC(u, v). \tag{5.6}$$

The use of the next Proposition is two-fold: it will be useful both for characterizing the class $\mathcal{L}_\gamma$ and establishing lower and upper bounds on the quantity $\eta(C, G_1, G_2)$.

PROPOSITION 5.4. *Let* $G_1, G_1', G_2, G_2' \in \mathcal{G}$. *Then*

$$G_2 \preceq_{st} G_2' \Rightarrow \eta(C, G_1, G_2) \leq \eta(C, G_1, G_2');$$
$$G_1 \preceq_{st} G_1' \Rightarrow \eta(C, G_1, G_2) \geq \eta(C, G_1', G_2).$$

PROOF. We prove only the first relation of Proposition 5.4, since the proof for the second one is analogous. By hypothesis, and since $G_1, G_2' \in \mathcal{G}$ for each $x \in (0, 1)$, one has

$$G_2^{-1}(x) \leq G_2'^{-1}(x).$$

Therefore

$$(G_1^{-1}(x), G_2^{-1}(x)) \in A \Rightarrow (G_1^{-1}(x), G_2'^{-1}(x)) \in A.$$

Hence, the proof can be concluded by recalling (5.6). □

From Proposition 5.4, in particular we get

$$\eta(C, G, G) \leq \eta(C, G', G) \quad \text{and} \quad \eta(C, G, G) \leq \eta(C, G, G''), \qquad (5.7)$$

for any choice of $G, G', G'' \in \mathcal{G}$ such that $G' \preceq_{st} G \preceq_{st} G''$.

A basic fact in the analysis of the classes $\mathcal{L}_\gamma$ is that the quantities of the form $\eta(C, G, G)$ only depend on the copula C. More formally we state the following result.

PROPOSITION 5.5. *For any pair of distribution functions* $G', G'' \in \mathcal{G}$, *one has*

$$\eta(C, G', G') = \eta(C, G'', G''). \qquad (5.8)$$

PROOF. Recalling (5.6) one obtains

$$\int_{I^2} \mathbb{1}_A(G'^{-1}(u), G'^{-1}(v))\, dC(u, v) = \int_{I^2} \mathbb{1}_A(G''^{-1}(u), G''^{-1}(v))\, dC(u, v)$$

because $\mathbb{1}_A(G'^{-1}(u), G'^{-1}(v)) = \mathbb{1}_A(G''^{-1}(u), G''^{-1}(v)) = \mathbb{1}_A(u, v)$, so equality in (5.8) is proved. □

As a consequence of Proposition 5.5 we can introduce the symbol

$$\eta(C) := \eta(C, G, G), \qquad (5.9)$$

and, by letting $G_1 = G_2 = G$ in (5.6), write

$$\eta(C) = \int_{A \cap [0,1]^2} dC \qquad (5.10)$$

for $G \in \mathcal{G}$. From Proposition 5.4 and from the inequalities (5.7), we obtain

PROPOSITION 5.6. *For* $G_1, G_2 \in \mathcal{G}$ *the following implication holds*

$$G_1 \preceq_{st} G_2 \Rightarrow \eta(C) \leq \eta(C, G_1, G_2).$$

We then see that the quantity $\eta(C)$ characterizes the class $\mathcal{L}_\gamma$, $0 \le \gamma \le 1$, in fact we can state the following

THEOREM 5.7. *$C \in \mathcal{L}_\gamma$ if and only if $\eta(C) \ge \gamma$.*

We thus have

$$\mathcal{L}_\gamma = \{C \in \mathcal{C} : \eta(C) \ge \gamma\} \tag{5.11}$$

and we can also write

$$\eta(C) = \inf_{G_1, G_2 \in \mathcal{G}} \{\eta(C, G_1, G_2) : G_1 \preceq_{st} G_2\}. \tag{5.12}$$

In other words the infimum in formula (5.12) is a minimum and it is attained when $G_1 = G_2$. We notice furthermore that the definition of $\eta(C, G_1, G_2)$ can be extended to the case when $G_1, G_2 \in D(\mathbb{R})$, the space of distribution functions on $\mathbb{R}$. The class $\mathcal{G}$ has however a special role in the present setting, as it is shown in the following result.

THEOREM 5.8. *Let $C \in \mathcal{C}$ and $G, H \in D(\mathbb{R})$ with $G \preceq_{st} H$. Then $\eta(C, G, H) \ge \eta(C)$.*

PROOF. Consider two sequences $(G_n : n \in \mathbb{N})$, $(H_n : n \in \mathbb{N})$ such that $G_n, H_n \in \mathcal{G}$ and $G_n \overset{w}{\to} G$, $H_n \overset{w}{\to} H$. Applying Theorem 2 in [**122**], we obtain that $C(G_n, H_n) \overset{w}{\to} C(G, H)$.

Consider now the new sequence $(\tilde{H}_n : n \in \mathbb{N})$, where we have posed $\tilde{H}_n(x) := \min\{G_n(x), H_n(x)\}$. Notice that $\tilde{H}_n \in \mathcal{G}$, moreover $G_n \preceq_{st} \tilde{H}_n$ and $\tilde{H}_n \overset{w}{\to} H$. This implies $C(G_n, \tilde{H}_n) \overset{w}{\to} C(G, H)$.

Now, by using the standard characterization of weak convergence on separable spaces (see [**14**] p. 67 Theorem 6.3),

$$\limsup_{n\to\infty} \int_B d\tilde{F}_n \le \int_B dF,$$

for any closed set $B \in \mathbb{R}^2$, where $F = C(G, H)$ and $\tilde{F}_n = C(G_n, \tilde{H}_n)$. Taking the closed set A defined in (5.2) one has

$$\eta(C) \le \limsup_{n\to\infty} \int_A d\tilde{F}_n \le \int_A dF = \eta(C, G, H). \tag{5.13}$$

□

REMARK 5.9. *Theorem 5.8 shows that the minimum of $\eta(C, G, H)$, for $G, H \in D(\mathbb{R})$, is attained at (C, G, G), for any $G \in \mathcal{G} \subset D(\mathbb{R})$. This result allows us to replace the class $\mathcal{G}$ with $D(\mathbb{R})$ in the expression of $\mathcal{L}_\gamma$ given in* (5.12)*. We notice furthermore that one can have $\eta(C, G', G') \neq \eta(C, G'', G'')$ when G', G'' are in $D(\mathbb{R})$.*

Concerning the classes $\mathcal{L}_\gamma$, we also define

$$\mathcal{B}_\gamma := \{C \in \mathcal{C} \mid \eta(C) = \gamma\}, \tag{5.14}$$

so that

$$\mathcal{L}_\gamma = \bigcup_{\gamma' \geq \gamma} \mathcal{B}_{\gamma'}.$$

We now show that the classes $\mathcal{B}_\gamma$, $\gamma \in [0,1]$, are all non empty. Several natural examples might be produced on this purpose. We fix attention on a simple example built in terms of the random variables X_1, $X_2^{(\gamma)}$ defined as follows. On the probability space $([0,1], \mathcal{B}[0,1], \lambda)$, where λ denotes the Lebesgue measure, we take $X_1(\omega) = \omega$, and

$$X_2^{(\gamma)}(\omega) = \begin{cases} \omega + 1 - \gamma & \text{if } \omega \in [0,\gamma], \\ \omega - \gamma & \text{if } \omega \in (\gamma, 1]. \end{cases} \tag{5.15}$$

As it happens for X_1, also the distribution of $X_2^{(\gamma)}$ is uniform in $[0,1]$ for any $\gamma \in [0,1]$ and the connecting copula of $X_1, X_2^{(\gamma)}$, that is then uniquely determined, will be denoted by C_γ.

PROPOSITION 5.10. *For any* $\gamma \in (0,1]$*, one has*

(i) $C_\gamma \in \mathcal{B}_\gamma$.
(ii) $C_\gamma(u,v) = \min\{u, v, \max\{u-\gamma, 0\} + \max\{v + \gamma - 1, 0\}\}$.

PROOF. (i) First we notice that $\mathbb{P}(X_1 \leq X_2^{(\gamma)}) = \gamma$. In fact

$$\begin{aligned} \mathbb{P}(X_1 \leq X_2^{(\gamma)}) &= \mathbb{P}(X_1 \leq X_1 + 1 - \gamma, X_1 \leq \gamma) \\ &\quad + \mathbb{P}(X_1 \leq X_1 - \gamma, X_1 > \gamma) = \gamma. \end{aligned}$$

Whence, $\eta(C_\gamma) = \mathbb{P}(X_1 \leq X_2^{(\gamma)}) = \gamma$, since both the distributions of $X_1, X_2^{(\gamma)}$ belong to $\mathcal{G}$.

(ii) For $x_1, x_2 \in [0,1]$ we can write

$$\begin{aligned} &F_{X_1, X_2^{(\gamma)}}(x_1, x_2) := \mathbb{P}(X_1 \leq x_1, X_2^{(\gamma)} \leq x_2) \\ &\quad = \mathbb{P}(X_1 \leq x_1, X_1 + 1 - \gamma \leq x_2,\ X_1 \leq \gamma) \\ &\quad + \mathbb{P}(X_1 \leq x_1, X_1 \leq x_2 + \gamma,\ X_1 > \gamma) \\ &\quad = \mathbb{P}(X_1 \leq \min\{x_1, x_2 + \gamma - 1, \gamma\}) + \mathbb{P}(\gamma < X_1 \leq \min\{x_1, x_2 + \gamma\}) \\ &\quad = \max\{\min\{x_1, x_2 + \gamma - 1, \gamma\}, 0\} + \max\{\min\{x_1, x_2 + \gamma\} - \gamma, 0\} \\ &\quad = \min\{x_1, x_2, \max\{x_1 - \gamma, 0\} + \max\{x_2 + \gamma - 1, 0\}\}. \end{aligned}$$

Since both the marginal distributions of X_1 and $X_2^{(\gamma)}$ are uniform, it follows that

$$C_\gamma(u,v) = \min\{u, v, \max\{u - \gamma, 0\} + \max\{v + \gamma - 1, 0\}\}.$$

□

The copulas C_γ have also been considered for different purposes in the literature, see e.g. [**107**] and [**127**]. We point out that the identity $\eta(C_\gamma) = \gamma$ (for $\gamma \in (0,1]$) could also have been obtained directly from formula (5.10).

In this special case the computation of $\mathbb{P}(X_1 \leq X_2)$ is however straightforward.

As an immediate consequence of Proposition 5.10 we have that $\mathcal{L}_{\gamma'}$ is strictly contained in $\mathcal{L}_\gamma$ for any $0 \leq \gamma < \gamma' \leq 1$. We notice furthermore that $\mathcal{L}_0 = \mathcal{C}$ and $\mathcal{L}_1 = \{C \in \mathcal{C} : \int_{A\cap[0,1]^2} dC = 1\} \neq \emptyset$.

Graphs of C_γ for different values of γ are provided in Figure 5.1.

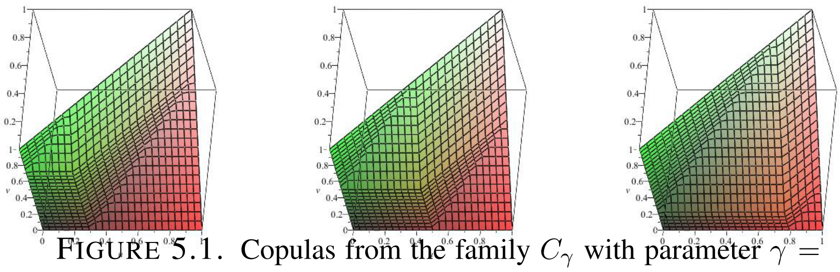

FIGURE 5.1. Copulas from the family C_γ with parameter $\gamma = 0.3,\ 0.5,\ 0.8$ respectively

3. A weak measure of non-exchangeability

For what follows it is now convenient also to consider the quantities $\xi(C, G_1, G_2)$ and $\xi(C)$ defined as follows:

$$\xi(C, G_1, G_2) := \mathbb{P}(X_1 = X_2), \tag{5.16}$$

$$\xi(C) := \xi(C, G, G), \tag{5.17}$$

where X_1 and X_2 are random variables with distributions $G_1, G_2 \in \mathcal{G}$ respectively and connecting copula C.

For a given bivariate model we have considered so far the quantities $\eta(C)$ with C denoting the connecting copula. In what follows we point out the relations among $\eta(C)$, $\eta(\widehat{C})$, $\eta(C^t)$ where $\widehat{C}$ and C^t denote the *survival copula* and the *transposed copula*, respectively. The transposed copula C^t is defined by

$$C^t(u, v) := C(v, u) \tag{5.18}$$

so that if C is the connecting copula of the pair (X_1, X_2), then C^t is the copula of the pair (X_2, X_1). Whence, if X_1 and X_2 have the same distribution $G \in \mathcal{G}$, then

$$\eta(C^t) = \mathbb{P}(X_2 \leq X_1).$$

On the other hand the notion of survival copula of the pair (X_1, X_2), which comes out as natural when considering pairs of non-negative random variables, is defined by the equation

$$\overline{F}_{X_1,X_2}(x_1, x_2) = \widehat{C}\left[\overline{G}_1(x_1), \overline{G}_2(x_2)\right], \tag{5.19}$$

with $\overline{G}_1$ and $\overline{G}_2$ respectively denoting the marginal survival functions:

$$\overline{G}_1(x_1) = \mathbb{P}(X_1 > x_1), \quad \overline{G}_2(x_2) = \mathbb{P}(X_2 > x_2).$$

The relationship between the survival copula $\widehat{C}$ of (X_1, X_2) and the connecting copula C is given by (see [**106**])

$$\widehat{C}(u,v) = u + v - 1 + C(1-u, 1-v). \tag{5.20}$$

The following result shows the relations tying the different quantities $\eta(C)$, $\eta(\widehat{C})$, $\eta(C^t)$.

PROPOSITION 5.11. *Let $C \in \mathcal{C}$. The following relation holds:*

$$\eta(\widehat{C}) = \eta(C^t) = 1 - \eta(C) + \xi(C). \tag{5.21}$$

PROOF. By the definition of η applied to $\widehat{C}$ one has

$$\begin{aligned}\eta(\widehat{C}) &= \int_{I^2} \mathbb{1}_A(u,v)\, d\widehat{C}(u,v) = \int_{I^2} \mathbb{1}_A(1-u', 1-v')\, dC(u',v') \\ &= 1 - \int_{I^2} \mathbb{1}_{A\setminus\partial A}(u',v')\, dC(u',v') = 1 - \eta(C) + \xi(C).\end{aligned}$$

Once again, by definition of η, we have

$$\begin{aligned}\eta(C^t) &= \int_{I^2} \mathbb{1}_A(u,v)\, dC^t(u,v) = \int_{I^2} \mathbb{1}_A(v',u')\, dC(u',v') \\ &= 1 - \int_{I^2} \mathbb{1}_{A\setminus\partial A}(u',v')\, dC(u',v') = 1 - \eta(C) + \xi(C),\end{aligned}$$

and finally $\eta(C^t) = \eta(\widehat{C})$. □

Fix now $G \in \mathcal{G}$ and let X_1, X_2 be random variables with a symmetric connecting copula C and both marginal distribution functions coinciding with G. Then their joint distribution function F_{X_1,X_2} is exchangeable and $\mathbb{P}(X_1 < X_2) = \mathbb{P}(X_2 < X_1) = (1 - \xi(C))/2$. Thus

$$\eta(C) = \mathbb{P}(X_1 < X_2) + \xi(C) = \frac{1 + \xi(C)}{2} \geq \frac{1}{2}. \tag{5.22}$$

We have $\eta(C) = 1/2$ when $\xi(C) = 2\eta(C) - 1 = 0$. As an immediate consequence of Theorem 5.7, we then get that any symmetric copula belongs to $\mathcal{L}_\gamma$ for any $\gamma \leq 1/2$, in other words when the copula is symmetric one has that the stochastic order implies the stochastic precedence.

On the other hand we are also interested in conditions under which the probability $\mathbb{P}(T \leq X)$ is "large enough", even if the marginal distributions of T and X are close each other. As a matter of fact, for random variables T and X with "close" marginal distributions, $\mathbb{P}(T \leq X)$ can be large only when the copula C is far from being symmetric. For this purpose it is opportune to recall the concept of *exchangeability* of random variables, introduced in 2.31: two random variables X_1 and X_2, with margins G_1 and G_2 and joint law $F_{1,2}$, are exchangeable if and only if $G_1 = G_2$ and $F_{1,2} = F_{2,1}$.

In Chapter 2 we also introduced the concept of *measure of non-exchangeability*, useful to understand the degree of non-exchangeability of couples of random variables or, analogously, the level of asymmetry of their connecting copula. Our aim is now to check if the index η can be considered as a suitable measure of asymmetry, or if, in any case, may give us information about copulas in this direction. To this purpose one can rather consider the quantity

$$\nu(C) := |\eta(C) - \eta(C^t)|, \tag{5.23}$$

for $C \in \mathcal{C}$. We are now going to show that the function ν defined above is a weak measure of non-exchangeability.

PROPOSITION 5.12. *The function $\nu : \mathcal{C} \in \mathbb{R}_+$ defined in 5.23 satisfies properties **B1,B3,B4,B5** of Definition* (2.34).

We give hints for the proof of Proposition 5.12.

B1: ν is bounded: $|\eta(C) - \eta(C^t)| \le \eta(C) + \eta(C^t) \le 2$;
B3-B4: $\nu(C) = \nu(C^t) = \nu(\widehat{C})$ by a direct application of (5.21);
B5: if (C_n) and C are in $\mathcal{C}$ and if C_n converges uniformly to C, then $\mu(C_n)$ converges to $\mu(C)$ as n tends to ∞, see Theorem 2 in [**122**].

For what concerns property **B2**, we shall need that $\nu(C) = 0$ if, and only if, C is symmetric. Of course if the copula C is symmetric, provided that $\xi(C) = 0$, we have $\nu(C) = 0$, but the opposite implication may fail. In this sense ν can be seen as a weak measure, because may lack of such a property. Notice that, for computational purposes, ν can also be written as

$$\nu(C) = |2\eta(C) - 1 - \xi(C)|. \tag{5.24}$$

In the special case of copulas C_γ (see Proposition 5.10) the equivalence holds, for $\gamma \in [0,1]$. In this case we have

$$\nu(C_\gamma) = \begin{cases} |2\gamma - 1| & \text{for } \gamma \in (0,1), \\ 0 & \text{for } \gamma = 0, 1. \end{cases} \qquad \xi(C_\gamma) = \begin{cases} 0 & \text{for } \gamma \in (0,1), \\ 1 & \text{for } \gamma = 0, 1. \end{cases}$$

The curve of the function ν is represented in Figure 5.2. Notice that, for the special cases $\gamma = 1/3$ and $\gamma = 2/3$, we have $\nu(C) = 1/3$, value that coincides with the one given by the measure d_∞ proposed in [**107**].

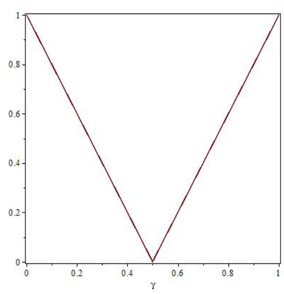

FIGURE 5.2. Graph of ν for the copulas C_γ

4. Further properties of $\mathcal{L}_\gamma$ and examples

We start this Section by analyzing further properties of the classes $\mathcal{L}_\gamma$ that can also shed light on the relations between stochastic precedence and stochastic orderings. First we notice that the previous Definition 5.2 has been formulated in terms of the usual stochastic ordering $\preceq_{st}$. However similar results can also be obtained for other important concepts of stochastic ordering that have been considered in the literature, such as the *hazard rate*, the *likelihood ratio*, and the other orderings (see Appendix B for further details about the topic, as well as [**124**]).

Let us fix, in fact, a stochastic ordering $\preceq_*$ different from $\preceq_{st}$. Definition 5.2 can be modified by replacing therein $\preceq_{st}$ with $\preceq_*$ and this operation leads us to a new class of copulas that we can denote by $\mathcal{L}_\gamma^{(*)}$. More precisely we set

$$\mathcal{L}_\gamma^{(*)} := \{C \in \mathcal{C} : \eta(C, G_1, G_2) \geq \gamma, \ \forall G_1, G_2 \in \mathcal{G} \ s.t. \ G_1 \preceq_* G_2\} \tag{5.25}$$

or equivalently

$$\mathcal{L}_\gamma^{(*)} = \{C \in \mathcal{C} : \eta^*(C) \geq \gamma\} \tag{5.26}$$

where

$$\eta^*(C) := \inf_{G_1, G_2 \in \mathcal{G}} \{\eta(C, G_1, G_2) : G_1 \preceq_* G_2\}. \tag{5.27}$$

For given $\gamma \in (0,1)$, one might wonder about possible relations between $\mathcal{L}_\gamma^{(*)}$ and $\mathcal{L}_\gamma$. Actually one has the following result, which will be formulated for binary relations (not necessarily stochastic orderings) over the space $D(\mathbb{R})$.

PROPOSITION 5.13. *Let $\preceq_*$ be a relation satisfying*

(a) *for any $G \in D(\mathbb{R})$ one has $G \preceq_* G$;*
(b) *for any $G_1, G_2 \in D(\mathbb{R})$ with $G_1 \preceq_* G_2$ one has $G_1 \preceq_{st} G_2$.*

Then $\mathcal{L}_\gamma = \mathcal{L}_\gamma^{(*)}$.

PROOF. In view of (b), one has that $\eta(C) \leq \eta^*(C)$. In fact both the quantities $\eta(C)$ and $\eta^*(C)$ are obtained as an infimum of the same functional and, compared with η, the quantity η^* is an infimum computed on a smaller set.

Due to (a), however, $\eta(C)$ and $\eta^*(C)$ are both obtained, in (5.12) and (5.27) respectively, as minima attained on a same point (G, G). We can then conclude that $\mathcal{L}_\gamma^{(*)} = \mathcal{L}_\gamma$. □

Concerning Proposition 5.13 we notice that, for example, the hazard rate and the likelihood ratio orderings, $\preceq_{hr}$ and $\preceq_{lr}$, both satisfy the conditions (a) and (b).

In applied problems it can be relevant to remark that imposing stochastic orderings stronger than $\preceq_{st}$ does not necessarily increase the level of stochastic precedence.

For the sake of notational simplicity we come back to considering the usual stochastic ordering $\preceq_{st}$ and the class $\mathcal{L}_\gamma$.

A basic property of the classes $\mathcal{L}_\gamma$ and $\mathcal{B}_\gamma$ is given by the following result.

PROPOSITION 5.14. *For $\gamma \in [0, 1]$, the classes $\mathcal{L}_\gamma$, $\mathcal{L}_\gamma^c = \mathcal{C} \setminus \mathcal{L}_\gamma$, and $\mathcal{B}_\gamma$ are convex.*

PROOF. We consider two bivariate copulas $C_1, C_2 \in \mathcal{L}_\gamma$ and a convex combination of them, i.e. take $\alpha \in (0, 1)$ and $C := \alpha C_1 + (1-\alpha)C_2$. We show that $C \in \mathcal{L}_\gamma$, indeed

$$\begin{aligned}\eta(C) = \int_A dC(u,v) = \alpha \int_A dC_1(u,v) + (1-\alpha)\int_A dC_2(u,v) \\ = \alpha\eta(C_1) + (1-\alpha)\eta(C_2).\end{aligned}$$

Since $\eta(C_1), \eta(C_2)$ are larger or equal than γ then $\eta(C) \geq \gamma$, whence $\mathcal{L}_\gamma$ is convex. Now one can use the same argument in order to show that $\mathcal{L}_\gamma^c$ and $\mathcal{B}_\gamma$ are convex as well. □

An immediate application of Proposition 5.14 concerns the case when, given a random parameter Θ, all the connecting copulas of the conditional distributions of (T, X), belong to a same class $\mathcal{L}_\gamma$. Proposition 5.14 in fact, guarantees that the copula of (T, X) belongs to $\mathcal{L}_\gamma$ as well.

Some aspects of the definitions and results given so far will be demonstrated here by presenting a few examples. We notice that, as shown by Proposition 5.10, the condition $\preceq_{st}$ does not imply $\preceq_{sp}^{(\gamma)}$, with $\gamma \in (0, 1)$. For the special case $\gamma = 1/2$ we now present an example of applied interest.

EXAMPLE 5.15.

Let X, Y be two non-negative random variables, where Y has an exponentially density $f_Y(y)$ with failure rate λ and where stochastic dependence between X and Y is described by a "load-sharing" dynamic model as follows: conditionally on $(Y = y)$, the failure rate of X amounts to $\alpha = 1$ for $t < y$ and to β for $t > y$. We assume $1 < \lambda < \beta < 1 + \lambda$. This position gives rise to a jointly absolutely continuous distribution for which we can consider

$$\mathbb{P}(X > x|Y = y) := \int_x^{+\infty} f_{X,Y}(t,y)dt,$$

$f_{X,Y}$ denoting the joint density of X, Y. As to the survival function of X, for any fixed value $x > 0$, we can argue as follows.

$$\overline{F}_X(x) := \mathbb{P}(X > x) = \int_0^{+\infty} \mathbb{P}(X > x|Y = y) f_Y(y)dy$$

The integral over $\mathbb{R}_+$ can be split in two parts, as follows. Over the interval $[0, x]$, we have

$$\int_0^x \mathbb{P}(X > x|Y = y)f_Y(y)dy =$$
$$\int_0^x \mathbb{P}(X > y|Y = y)\mathbb{P}(X > x|Y = y, X > y)f_Y(y)dy =$$
$$\int_0^x e^{-y}e^{-b(x-y)}f_Y(y)dy$$

while, over $[x, +\infty]$,

$$\int_x^{+\infty} \mathbb{P}(X > x|Y = y)f_Y(y)dy = \int_x^{+\infty} e^{-x}f_Y(y)dy.$$

Then we have, for all $x > 0$,

$$\begin{aligned}\overline{F}_X(x) &= e^{-bx}\frac{\lambda}{1+\lambda-b}[1 - e^{-(1+\lambda-b)x}] + e^{-(1+\lambda)x} \\ &= \left(1 - \frac{\lambda}{1+\lambda-\beta}\right)e^{-(1+\lambda)x} + \frac{\lambda}{1+\lambda-\beta}\,e^{-\beta x} \leq e^{-\lambda x}.\end{aligned}$$

We can then conclude that $X \preceq_{st} Y$. On the other hand the same position gives also rise to $\mathbb{P}(X \leq Y) = 1/(1+\lambda) < 1/2$.

The next example shows that for three random variables T, X', X'', the implication $T \preceq_{st} X' \preceq_{st} X'' \Rightarrow \mathbb{P}(T \leq X'') \leq \mathbb{P}(T \leq X')$ can fail when the connecting copulas of (T, X') and (T, X'') are different.

EXAMPLE 5.16.

Let $Y_1, \ldots, Y_5$ be i.i.d. random variables, with a continuous distribution and defined on a same probability space, and set

$$T = \min\{Y_1, Y_2\},\ X' = \max\{Y_1, Y_2\},\ X'' = \max(Y_3, Y_4, Y_5).$$

Thus $X' \preceq_{st} X''$, but $\mathbb{P}(T \leq X') = 1$ and $\mathbb{P}(T \leq X'') < 1$.

REMARK 5.17. *For some special types of copula C, the computation of $\eta(C, G_1, G_2)$ can be carried out directly, in terms of probabilistic arguments, provided the distributions G_1, G_2 belong to some appropriate class. This circumstance in particular manifests for the models considered in the subsequent examples. Let C be a copula satisfying such conditions. Then Proposition 5.4 can be used to obtain inequalities for $\eta(C, H_1, H_2)$ even if H_1, H_2 do not belong to $\mathcal{G}$ provided, e.g., that $H_1 \preceq_{st} G_1$, $G_2 \preceq_{st} H_2$ and $G_1, G_2 \in \mathcal{G}$.*

The next example will be devoted to bivariate gaussian models, i.e. to a relevant case of symmetric copulas.

EXAMPLE 5.18. ***Gaussian Copulas.***

The family of bivariate gaussian copulas (see e.g. [**106**]) is parameterized by the correlation coefficient $\rho \in (-1,1)$. The corresponding copula $C^{(\rho)}$ is absolutely continuous and symmetric, and $\eta(C^{(\rho)}) = 1/2$ and, thus, it does not depend on ρ. For fixed pairs of distributions G_1, G_2, on the contrary, the quantity $\eta(C^{(\rho)}, G_1, G_2)$ does actually depend on ρ, besides on G_1 and G_2. This class provides the most direct instance of the situation outlined in the above Remark 5.17. The value for $\eta(C^{(\rho)}, G_1, G_2)$ is in fact immediately obtained when G_1, G_2 are gaussian. Let X_1, X_2 denote gaussian random variables with connecting copula $C^{(\rho)}$ and parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$. Since the random variable $Z = X_1 - X_2$ is distributed according to $\mathcal{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)$ we can write

$$\eta(C^{(\rho)}, G_1, G_2) = \mathbb{P}(Z \leq 0) = \Phi\left(\frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}}\right). \qquad (5.28)$$

We recall that, when $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ for $i = 1, 2$, the necessary and sufficient condition for $X_1 \preceq_{st} X_2$ is $\mu_1 \leq \mu_2$ and $\sigma_1 = \sigma_2$ (see e.g. [**6**]). In other words, for G_1, G_2 gaussian, $G_1 \preceq_{st} G_2$ means $X_1 \preceq_{sp} X_2$ and $\sigma_1 = \sigma_2$. By using the formula in (5.28), with $\sigma_1 = \sigma_2 = \sigma$, we have

$$\eta(C^{(\rho)}, G_1, G_2) = \Phi\left(\frac{\mu_2 - \mu_1}{\sigma\sqrt{2(1-\rho)}}\right). \qquad (5.29)$$

Thus $G_1 \preceq_{st} G_2 \Rightarrow \eta(C^{(\rho)}, G_1, G_2) \geq 1/2$, as shown by Proposition 5.6 and Theorem 5.8. We notice that $\eta(C^{(\rho)}, G_1, G_2)$ is an increasing function of ρ.

Proposition 5.4 can be extended to obtain, say, that

$$\eta(C^{(\rho)}, G_1, G_2) \leq \eta(C^{(\rho)}, H_1, H_2),$$

when $H_1 \preceq_{st} G_1$ and $G_2 \preceq_{st} H_2$, for $G_1, G_2 \in \mathcal{G}$ and $H_1, H_2 \notin \mathcal{G}$. We then can give inequalities for $\eta(C^{(\rho)}, H_1, H_2)$ in terms of (5.28), provided H_1, H_2 are suitably comparable in the $\preceq_{st}$ sense with gaussian distributions.

In the cases when $\xi(C) > 0$, we should obviously distinguish between computations of $\mathbb{P}(X_1 \leq X_2)$ and $\mathbb{P}(X_1 < X_2)$, where C is the connecting copula of X_1, X_2. A remarkable case when this circumstance happens is considered in the following example.

EXAMPLE 5.19. ***Marshall-Olkin Models***

We consider the Marshall-Olkin copulas (see e.g [**76, 102, 106**]), namely those whose expression is the following:

$$\widehat{C}^{(\alpha_1,\alpha_2)}(u,v) := u\,v\,\min\{u^{-\alpha_1},\, v^{-\alpha_2}\}$$

for $0 < \alpha_i < 1$, $i = 1, 2$. We notice that the Marshall-Olkin copula has a singular part that is concentrated on the curve $u^{\alpha_1} = v^{\alpha_2}$ (see also Figure

5.3). Actually the measure of such a singular component is given by

$$\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2-\alpha_1\alpha_2}.$$

FIGURE 5.3. Marshall-Olkin Copula (left) and graph of $u^{\alpha_1} = v^{\alpha_2}$ (right). Special case $\alpha_1 = 0.4$, $\alpha_2 = 0.2$.

As for the computation of $\eta(\widehat{C}^{(\alpha_1,\alpha_2)})$ we use the expression in (5.10). By separately considering the curve $u^{\alpha_1} = v^{\alpha_2}$ and the domains where $\widehat{C}^{(\alpha_1,\alpha_2)}$ is absolutely continuous, we obtain

$$\eta(\widehat{C}^{(\alpha_1,\alpha_2)}) = \frac{1}{2-\alpha_1\wedge\alpha_2}\left(1-\frac{(\alpha_1-\alpha_1\wedge\alpha_2)(\alpha_1\wedge\alpha_2)}{\alpha_1-\alpha_2}\right).$$

Consider the copula

$$C^{(\alpha_1,\alpha_2)}(u,v) := \widehat{C}^{(\alpha_1,\alpha_2)}(1-u,1-v)+u+v-1.$$

We will see now that the value of $\eta(C^{(\alpha_1,\alpha_2)},G_1,G_2)$ directly follows from probabilistic arguments, provided G_1, G_2 are exponential distributions with appropriate parameters. Let in fact V, W and Z be three random variables independent and exponentially distributed with parameters $\mu_1 = 1/\alpha_1 - 1$, $\mu_2 = 1/\alpha_2 - 1$ and $\mu = 1$, respectively. The new random variables

$$X_1 := V\wedge Z, \qquad X_2 := W\wedge Z,$$

have survival copula $\widehat{C}^{(\alpha_1,\alpha_2)}$, connecting copula $C^{(\alpha_1,\alpha_2)}$, and exponential distributions $G_1^{(\alpha_1)}$ and $G_2^{(\alpha_2)}$, with parameters $1/\alpha_1$ and $1/\alpha_2$ respectively. We now proceed with the computation of

$$\eta(C^{(\alpha_1,\alpha_2)},G_1^{(\alpha_1)},G_2^{(\alpha_2)}) = \mathbb{P}(X_1\le X_2).$$

We can write

$$\begin{aligned}\xi(C^{(\alpha_1,\alpha_2)},G_1^{(\alpha_1)},G_2^{(\alpha_2)}) &= \mathbb{P}(X_1 = X_2) = \mathbb{P}(Z\le V\wedge W)\\ &= \frac{1}{\mu_1+\mu_2+1} = \frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2-\alpha_1\alpha_2},\end{aligned}$$

furthermore

$$\mathbb{P}(X_1 < X_2) = \mathbb{P}(V < W\wedge Z) = \frac{(1-\alpha_1)\,\alpha_2}{\alpha_1+\alpha_2-\alpha_1\alpha_2},$$

and finally we obtain

$$\mathbb{P}(X_1 \le X_2) = \frac{\alpha_2}{\alpha_1 + \alpha_2 - \alpha_1\alpha_2}.$$

Then

$$\eta(C^{(\alpha_1,\alpha_2)}, G_1^{(\alpha_1)}, G_2^{(\alpha_2)}) = \frac{\alpha_2}{\alpha_1 + \alpha_2 - \alpha_1\alpha_2}.$$

Finally, the evaluation of $\nu(C)$ is straightforward and we obtain

$$\nu(C^{(\alpha_1,\alpha_2)}) = \frac{\alpha_1 \wedge \alpha_2}{2 - \alpha_1 \wedge \alpha_2}.$$

We notice that, also in the present Marshall-Olkin case, the index ν defined in (5.23) perfectly fits with the definition of measure of non-exchangeability given in [**49**]. In fact one has that $\nu(C) = 0$ only in the case $\alpha_1 = \alpha_2 = 0$, that corresponds to $C(u,v) = uv$, the independence copula.

We now conclude this Section with an example showing an extreme case in the direction of Remark 5.17.

EXAMPLE 5.20. ***Copulas of order statistics***

Let A, B be two i.i.d. random variables with d.f. $G \in \mathcal{G}$ and denote by X_1, X_2 their order statistics, namely $X_1 = \min\{A,B\}, X_2 = \max\{A,B\}$. The distributions of X_1, X_2 depend on G and are respectively given by

$$F_1^{(G)}(x_1) = \mathbb{P}(\min\{X_1,X_2\} \le x_1) = 2G(x_1) - G(x_1)^2,$$
$$F_2^{(G)}(x_2) = \mathbb{P}(\max\{X_1,X_2\} \le x_2) = G(x_2)^2\,.$$

Let $Z := \{(u,v) \in I^2 \,:\, v \ge (1-(1-u)^{1/2})^2\}$. The connecting copula of (X_1, X_2), represented in Figure 5.4, is given by

$$K(u,v) = \begin{cases} 2(1-(1-u)^{1/2})v^{1/2} - (1-(1-u)^{1/2})^2 & \text{if } (u,v) \in Z, \\ v & \text{otherwise.} \end{cases}$$

We have, by definition,

$$\eta(K, F_1^{(G)}, F_2^{(G)}) = 1,$$

and it does not depend on G. We notice, on the other hand, that the computation of $\eta(K) = \eta(K,G,G)$, with $G \in \mathcal{G}$, is to be carried out explicitly, since the pair (G,G) can never appear as the pair of marginal distributions of order statistics. By recalling (5.6) one obtains

$$\eta(K) = \int_{[0,1]^2} \frac{\mathbb{1}_A(u,v)}{2\sqrt{v}\sqrt{1-u}}\,dv\,du = 2 - \frac{\pi}{2} < \frac{1}{2},$$
$$\nu(K) = |2\eta(K) - 1| = \pi - 3.$$

We can extend this example to the case when the connecting copula of A, B is a copula D different from the product copula Π, but still A and B are identically distributed according to a distribution function G. In this case the connecting copula K of X_1, X_2 depends on D, but again it does not depend on G (see [**105**] page 478).

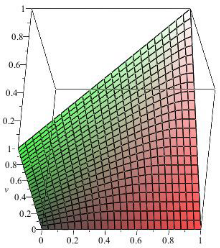

FIGURE 5.4. Ordered Statistic Copula K

5. The classes $\mathcal{L}_\gamma$ in the Target-Based Approach

In this Section we trace connections between our results about stochastic precedence, introduced in the previous sections, and the Target-Based Approach to decision problems under risk.

So far we introduced the Target-Based Model of utility and studied many of its properties, especially in the multi-attribute case and in the case of independence between targets and prospects. Here we concentrate attention on the single-attribute case, where (T, X) is a pair of real-valued random variables. Furthermore, we are interested in the case where there is dependence between T and X.

It is clear that the objects of central interest in the TBA are, for a fixed target T, the probabilities $\mathbb{P}(T \leq X)$ and the analysis developed in the previous sections can reveal of interest. Here we assume the existence of regular conditional distributions and, in particular, for any prospect X we assume that we can determine the function $\upsilon_T^{(X)}(x) := \mathbb{P}(T \leq x | X = x)$. Hence we can write

$$\mathbb{P}(T \leq X) = \int_{\mathbb{R}} \upsilon_T^{(X)}(x)\, dF_X(x).$$

Before continuing it is useful to remind the special case when X and T are stochastically independent. In this case we can write

$$\mathbb{P}(T \leq X) = \int_{\mathbb{R}} \upsilon_T^{(X)}(x)\, dF_X(x) = \int_{\mathbb{R}} F_T(x)\, dF_X(x).$$

In such a case, as we already remarked in Chapter 4, $\mathbb{P}(T \leq X)$ can be seen as the expected value of a utility: by considering $U = F_T$ as the utility function, we have (see formula (4.2))

$$\mathbb{E}(U(X)) = \int_{\mathbb{R}} U(x)\, dF_X(x) = \int_{\mathbb{R}} F_T(x)\, dF_X(x) = \mathbb{P}(T \leq X).$$

Under the condition of independence, any bounded and right-continuous utility function can thus be seen as the distribution function of a target T, and vice-versa. Such an hypothesis represented a balance point in the study of Target-Based model illustrated in Chapter 4. In this sense, our model can be seen as an extension of classical models for utility, although it adapts to the expected utility principle. TBA however becomes, in a sense, more general than the expected utility approach by allowing for stochastic dependence between targets and prospects. In fact the TBA considers more general decision rules, if we admit the possibility of some correlation between X and T. In this case, $\upsilon_T^{(X)}(x)$ does not coincide anymore with the distribution function $F_T(x)$ of the target. We refer to [**19, 28**] for further discussion in this sense.

We now briefly summarize the arguments of previous sections in the perspective of a decision problem where, for a fixed target T, we aim to rank two different prospects X_1, X_2, with marginal distributions G_{X_1}, G_{X_2}, and with connecting copulas C_1, C_2, corresponding to the pairs (T, X_1) and (T, X_2), respectively.

In the case of independence, a prospect X_2 should be obviously preferred to a prospect X_1 if $X_1 \preceq_{st} X_2$. In the case of dependence, on the contrary, this comparison is not sufficient anymore. In fact the choice of a prospect X should be based not only on the corresponding distribution F_X, but also on the connecting copula of the pair (T, X).

For fixed C, the quantity $\eta(C, G_T, G_X) = \mathbb{P}(T \leq X)$ is equal to the quantity $\eta(C)$ for all pairs such that $G_T = G_X = G$, with G belonging to the class $\mathcal{G}$ (See Proposition 5.5) while, for $G_T \neq G_X$, the implication $T \preceq_{st} X \Rightarrow \mathbb{P}(T \leq X) \geq \gamma$ does not necessarily hold (see Proposition 5.10 and Example 5.15).

For two different prospects X_1, X_2, Proposition 5.4 guarantees that, when $C_1 = C_2 = C$, the condition $G_T \preceq_{st} G_{X_1} \preceq_{st} G_{X_2}$ implies

$$\eta(C, G_T, G_{X_1}) = \mathbb{P}(T \leq X_1) \leq \eta(C, G_T, G_{X_2}) = \mathbb{P}(T \leq X_2).$$

As shown by Example 5.16, when $C_1 \neq C_2$, we can have both the conditions $\eta(C_1, G_T, G_{X_1}) > \eta(C_2, G_T, G_{X_2})$ and $G_T \preceq_{st} G_{X_1} \preceq_{st} G_{X_2}$ $(G_{X_1} \neq G_{X_2})$.

Concerning the quantities $\eta(C_1, G_T, G_{X_1})$ and $\eta(C_2, G_T, G_{X_2})$, Theorems 5.7 and 5.8 show that, for $G_T \preceq_{st} G_{X_i}$ $(i = 1, 2)$,

$$\mathbb{P}(T \leq X_i) = \eta(C_i, G_T, G_{X_i}) \geq \eta(C_i).$$

Finally, let us consider the case when the only available information about C_1 and C_2 is that $\eta(C_i) \geq \gamma_i$ (i.e. that C_i belongs to the class $\mathcal{L}_{\gamma_i}$). Then a rough and conservative choice between X_1 and X_2 suggests to select X_i with the larger value of γ_i, provided $G_{X_1} \preceq_{st} G_{X_2}$ or that X_1, X_2 are nearly identically distributed.

All these apparently paradoxical results suggest that the criteria for selection of random variables based only on stochastic orderings are not suitable enough for decision-making problems, such as those described by the TBA, when dependance among variables is present. We have shown, in fact, that the usual stochastic orderings can give results in disagreement with the expected utility concepts expressed by TBA. Furthermore we explicitly provided examples in which the choice of a prospect which is "better" in the stochastic sense may give worse results in the utility context.

In order to describe his preferences to the best, a DM adopting the Target-Based model will then also need to take into account properties of dependence of the random variables involved in his choices, trough the study of their connecting copulas. To this purpose a deeper analysis of the copulas of the classes $\mathcal{L}_\gamma$ is to be performed, especially for what concerns the properties of dependence and asymmetry.

Conclusions and Future Work

In this work we showed the importance of the target-based model in decision making and utility theory. We presented an extension of multi-attribute target-based model, representing preferences according to the von-Neumann Morgenstern utility theory, although built by means of non-additive measures. This model provides, in fact, an analysis of the joint behavior of targets and prospects, describing them in terms of their joint probability distributions, by means of properties of copulas, and by (non-additive) importance weights defined in terms of capacities. On this basis, we have pointed out that the theory of multi-attribute target-based utilities can hinge on a formal apparatus, provided by the field of fuzzy measures, extensions of fuzzy measures, and fuzzy, or universal, integrals.

Further improvements can be made to this model, from one side, by deeply investigating the role of capacities in establishing the importance of groups of prospects. On the other side, properties of risk aversion in high dimensions have to be mastered, through the analysis of the connecting copulas of targets and prospects. An overall interaction between copulas and capacities is to be studied in deep, by taking into account the features that these objects jointly assume in our model.

In this work we also presented an extension of the concept of stochastic precedence and provided comparison with the usual concepts of stochastic orders, in terms of properties of copulas. We provided some examples in this direction and found link to the target-based model of utility.

Extensions of this topic can be made trough a more accurate analysis of properties of copulas, especially regarding dependence and asymmetry. Connections with the existing concept of measures of concordance and measures of asymmetry can be improved for this purpose.

APPENDIX A

A brief introduction to Reliability of Systems

Reliability is defined as the probability that a device will perform its intended function during a specified period of time under stated conditions. In this brief note we will consider reliability of a system for a fixed moment of time, so that the state of the system is assumed to depend only on the state of its components. We will distinguish between only two states: a functioning state and a failed one. Let's refer to a variable ϕ to indicate the state of the whole system, made up of n components, $\{1, \ldots, n\}$. To indicate the state of a single component, say the $i-$th component, we use a binary indicator variable x_i that may assume two values: $x_i = 1$ if component i is functioning, $x_i = 0$ if component i is failed. The value of ϕ, in turn, can be 0 or 1 if the system is failed or working. The function $\phi(\mathbf{x})$ is called the *structure function* of the system, where $\mathbf{x} = (x_1, \ldots, x_n)$ is the vector of its components.

The most common examples of systems built in this way are the ones of series system and parallel system. The series system has structure function given by

$$\phi(\mathbf{x}) = \prod_{i=1}^{n} x_i = \min(x_1, \ldots, x_n),$$

and represents a system that can only work if all components are working. Parallel system represents the very opposite case, in which the system works if at least one of its component is functioning. Its structure function is given by

$$\phi(\mathbf{x}) = \coprod_{i=1}^{n} x_i = 1 - \prod_{i=1}^{n}(1 - x_i) = \max(x_1, \ldots, x_n).$$

These two are examples of *symmetric systems*, in which the state of the system only depends on the number of working components, regardless of what they are. They are particular cases of the k-out-of-n system presented in (4.32), in which the system works if at least k components out of n work. Notice that the systems introduced above are expressed by means of *lattice polynomial functions*, roughly speaking by functions only made by simple logical operators like $\min$ and $\max$ (for a better explanation and some properties about lattice polynomial functions see, for example, [**47**]).

We now list some definitions that will be useful later on.

DEFINITION A.1. *Given a structure ϕ, its dual ϕ^* is given by*

$$\phi^*(\mathbf{x}) = 1 - \phi(\mathbf{1} - \mathbf{x}),$$

where $\mathbf{1} - \mathbf{x} := (1 - x_1, \ldots, 1 - x_n)$.

It is easy to check that the dual of a series system is a parallel one and vice-versa, while the dual of a k-out-of-n structure is a $(n-k+1)$-out-of-n structure.

DEFINITION A.2. *The* $i-th$ *component is* irrelevant *to the structure* ϕ *if* ϕ *is constant in* x_i*, i.e. if* $\phi(1_i, \mathbf{x}) = \phi(0_i, \mathbf{x})$ *for all* $\phi(\cdot_i, \mathbf{x})$. *Namely*

$$\phi(\cdot_i, \mathbf{x}) = (x_1, \ldots, x_{i-1}, \cdot_i, x_{i+1}, \ldots, x_n).$$

Otherwise i *will be called* relevant.

DEFINITION A.3. *A system of components is* coherent *if its structure function* ϕ *is increasing and each component is relevant. If* ϕ *is only non-decreasing the system will be called* semi-coherent.

The property of monotonicity is important for physical systems, for which there is no opportunity that while improving the performance of a component, the system may tend to deteriorate. Coherent systems also enjoy the following boundary property:

$$\prod_{i=1}^{n} x_i \leq \phi(\mathbf{x}) \leq \coprod_{i=1}^{n} x_i.$$

Alternative ways to represent a coherent structure can be given by means of its working/failing states, as follows. Let $\mathbf{x}$ indicate the states of the sets of components $CN = \{1, \ldots, n\}$. Then we define $CN_0(\mathbf{x}) = \{i | x_i = 0\}$ and $CN_1(\mathbf{x}) = \{i | x_i = 1\}$. Assume that the structure (CN, ϕ) is coherent.

DEFINITION A.4. *A* path vector *is a vector* $\mathbf{x}$ *such that* $\phi(\mathbf{x}) = 1$ *and* $CN_1(\mathbf{x})$ *is the corresponding* path set. *A* minimal path vector *is a vector* $\mathbf{x}$ *such that* $\phi(\mathbf{x}) = 1$ *and, for any* $\mathbf{y} < \mathbf{x}$, $\phi(\mathbf{y}) = 0$. *The corresponding* minimal path set *is* $CN_1(\mathbf{x})$.

A cut vector *is a vector* $\mathbf{x}$ *such that* $\phi(\mathbf{x}) = 0$ *and* $CN_0(\mathbf{x})$ *is the corresponding* cut set. *A* minimal cut vector *is a vector* $\mathbf{x}$ *such that* $\phi(\mathbf{x}) = 0$ *and, for any* $\mathbf{y} > \mathbf{x}$, $\phi(\mathbf{y}) = 1$. *The corresponding* minimal cut set *is* $CN_0(\mathbf{x})$.

If we denote by P_j the $j-$th minimal path set of ϕ, we may define

$$\rho_j(\mathbf{x}) = \prod_{x_i \in P_j} x_i$$

as the $j-th$ *minimal path series structure*, which takes values 1 if all components in the minimal path set function, 0 otherwise ($j = 1, \ldots, p$, where p is the number of minimal path sets of ϕ). Then we can write the representation of ϕ through its path sets as

$$\phi(\mathbf{x}) = \coprod_{j=1}^{p} \rho_j(\mathbf{x}) = \coprod_{j=1}^{p} \prod_{x_i \in P_j} x_i.$$

A similar result can be obtained in view of the cut sets of ϕ,

$$\phi(\mathbf{x}) = \prod_{j=1}^{k} \kappa_j(\mathbf{x}) = \prod_{j=1}^{k} \prod_{x_i \in K_j} x_i,$$

where K_j is the $j-$th minimal cut set of ϕ, $j = 1, \dots, k$, and κ_j is the $j-$th *minimal parallel cut structure*.

We are now ready to introduce the concept of reliability of a system.

DEFINITION A.5. *Assume that the states of the components of a system ϕ are random variables $X_1, \dots, X_n$, with*

$$\mathbb{P}(X_i = 1) = p_i = \mathbb{E}[X_i],$$

for $i = 1, \dots, n$. We refer to p_i as the reliability *of i. The reliability of the system is similarly defined by*

$$\mathbb{P}(\phi(\mathbf{X}) = 1) = h = \mathbb{E}[\phi(\mathbf{X})].$$

The reliability of the examples mentioned above can be easily evaluated. Series systems, as well as parallel and k-out-of-n systems, are symmetric, so every component has the same reliability, say p. We have

(1) $\phi(\mathbf{X}) = p^n$ for series system;
(2) $\phi(\mathbf{X}) = 1 - (1-p)^n$ for parallel systems;
(3) $\phi(\mathbf{X}) = \sum_{i=k}^{n} \binom{n}{i} p^i (1-p)^{n-i}$ for k-out-of-n systems.

We give a final remark about lower and upper bounds for reliability. Let E_r be the event that all the components in minimal path set P_r work. Then

$$\mathbb{P}(E_r) = \prod_{i \in P_r} p_i.$$

System success corresponds to the event $E = \cup_{r=1}^{p} E_r$, if the system has p minimal path sets. Then

$$h = \mathbb{P}\left(\bigcup_{r=1}^{p} E_r \right).$$

Let

$$S_k = \sum_{1 \le i_1 < \ldots < i_k \le p} \mathbb{P}(E_{i_1} \cap \ldots \cap E_{i_k}),$$

then, by means of the inclusion-exclusion principle, we have

$$h = \sum_{k=1}^{p} (-1)^{k-1} S_k,$$

and

$$h \le S_1, \; h \ge S_1 - S_2, \; h \le S_1 - S_2 + S_3,$$

and so on. This method provides, hence, successive upper and lower bounds on system reliability, which converge to the exact system reliability.

For further properties of systems and a deeper study of their reliability we refer to [**10**].

APPENDIX B

Some notions of Stochastic Orderings

Here we briefly introduce the main stochastic orders with a few properties useful in this paper.

DEFINITION B.1. *Let X and Y be two random variables such that*

$$\mathbb{P}(X > z) \leq \mathbb{P}(Y > z)$$

for all $z \in \mathbb{R}$. Then X is said to be smaller than Y in the usual stochastic order*, and it will be written $X \preceq_{st} Y$.*

Roughly speaking, X is less likely than Y to take large values, when "large" means for values bigger than any fixed $z \in \mathbb{R}$. Characterization of stochastic ordering can be given, as the following two results state.

THEOREM B.2. *Two random variables X and Y satisfy $X \preceq_{st} Y$ if, and only if, there exist two random variables $\hat{X}$ and $\hat{Y}$, defined on a same probability space, such that*

$$\hat{X} =_{st} X, \ \hat{Y} =_{st} Y, \ \text{and } \mathbb{P}(\hat{X} \leq \hat{Y}) = 1.$$

Another way to read the previous Theorem is the following

THEOREM B.3. *Two random variables X and Y satisfy $X \preceq_{st} Y$ if, and only if, there exist a random variable Z and two functions ψ_1 and ψ_2 such that $\psi_1(z) \leq \psi_2(z)$ for all z and $X =_{st} \psi_1(Z)$ and $Y =_{st} \psi_2(Z)$.*

For proofs of these Theorems and some properties of stochastic order we refer to [**124**]. Consider now the following

DEFINITION B.4. *If X is a non-negative variable with an absolutely continuous distribution function F, then the* hazard rate *of X at $t \geq 0$ is defined as*

$$r(t) = \frac{d}{dt}(-\log(\overline{F}(t))) = \frac{f(t)}{\overline{F}(t)},$$

where $\overline{F}(t) = 1 - F(t)$ is the survival function and $f(t) = \partial_t F(t)$ is the corresponding density function.

The hazard rate is a very important instrument in reliability theory, since many properties of systems follow from its definition (we refer to [**10**] for further information). Moreover, a new type of ordering can be built upon it.

DEFINITION B.5. *Let X and Y be two non-negative random variables with hazard rates, respectively, $r(t)$ and $q(t)$, $t \geq 0$. Then X is smaller than*

Y in the hazard rate order *(denoted by $X \preceq_{hr} Y$) if, and only if, $r(t) \geq q(t)$ for all $t \geq 0$.*

An equivalent condition is the following: if F and G are the distribution functions of X and Y respectively, then $X \preceq_{hr} Y$ if, and only if, $\overline{F}(t)/\overline{G}(t)$ is a decreasing function of t. The link between hazard rate and stochastic order is determined by the following

THEOREM B.6. *If X and Y are two random variables such that $X \preceq_{hr} Y$, then $X \preceq_{st} Y$.*

Consider now the property of *monotone likelihood ratio*, a property regarding the ratio of two probability density functions. As usual for monotonic relationships, the likelihood ratio's monotonicity comes in handy in statistics, particularly when using maximum-likelihood estimation. In our context, it gives rise to a corresponding ordering, that can be stated as follows.

DEFINITION B.7. *Two random variables X and Y, with density functions f and g respectively, have decreasing likelihood ratio if $f(t)/g(t)$ decreases over the union of the supports of X and Y. In this case we say that X is smaller than Y in the* likelihood ratio order*, written $X \preceq_{lr} Y$.*

The connection between likelihood ratio and the other two orderings is given by the following result.

THEOREM B.8. *If X and Y are two random variables such that $X \preceq_{lr} Y$, then $X \preceq_{hr} Y$.*

It is then clear that this ordering is stronger than the other two orderings presented here, in fact we have

$$X \preceq_{lr} Y \Rightarrow X \preceq_{hr} Y \Rightarrow X \preceq_{st} Y.$$

Many other orderings are present in literature, for knowledge we refer again to [**10**].

List of Figures

Bibliography

[1] A. E. Abbas and J. E. Matheson. Normative target-based decision making. *Managerial and Decision Economics*, 26(6):373–385, 2005.

[2] S. Aki and K. Hirano. Sooner and later waiting time problems for runs in markov dependent bivariate trials. *Annals of the Institute of Statistical Mathematics*, 51:17–29, 1999.

[3] S. Angilella, S. Greco, F. Lamantia, and B. Matarazzo. Assessing non-additive utility for multicriteria decision aid. *European Journal of Operational Research*, 158(3):734–744, 2004.

[4] D. Antzoulakos and A. Philippou. On waiting time problems associated with runs in markov dependent trials. *Annals of the Institute of Statistical Mathematics*, 51:323–330, 1999.

[5] M. A. Arcones, P. H. Kvam, and F. J. Samaniego. Nonparametric estimation of a distribution subject to a stochastic precedence constraint. *Journal of the American Statistical Association*, 97(457):170–182, 2002.

[6] M. A. Arcones, P. H. Kvam, and F. J. Samaniego. Nonparametric estimation of a distribution subject to a stochastic precedence constraint. *J. Amer. Statist. Assoc.*, 97(457):170–182, 2002.

[7] K. Arrow. The theory of discrimination. *Discrimination in labor markets*, 3(10), 1973.

[8] K. Arrow. *Essays in the theory of risk-bearing*. North-Holland, 1976.

[9] J. Avérous, C. Genest, and S. C Kochar. On the dependence structure of order statistics. *Journal of multivariate analysis*, 94(1):159–171, 2005.

[10] R. Barlow and F. Proschan. *Statistical theory of reliability and life testing: probability models*. International series in decision processes. Holt, Rinehart and Winston, 1975.

[11] A. Basu. *Measure Theory And Probability*. Prentice-Hall Of India Pvt. Limited, 2004.

[12] G. Beliakov, T. Calvo, and A. Pradera. *Aggregation Functions: A Guide for Practitioners*. Studies in fuzziness and soft computing. Springer-Verlag, 2007.

[13] W. Bi and L. Zhang. Target oriented multiattribute group decision making approach and application in hydroelectric project evaluation. *Systems Engineering Procedia*, 5:454–460, 2012.

[14] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. Wiley, 2009.

[15] G. Birkhoff. *Lattice Theory*. Number v. 25,pt. 2 in American Mathematical Society colloquium publications. American Mathematical Society, 1967.

[16] G. Blom and D. Thorburn. How many random digits are required until given sequences are obtained? *J. Appl. Probab.*, 19(3):518–531, 1982.

[17] G. Blom and D. Thorburn. Correction: "How many random digits are required until given sequences are obtained?" [J. Appl. Probab. **19** (1982), no. 3, 518–531. *J. Appl. Probab.*, 22(2):485, 1985.

[18] P. J. Boland, H. Singh, and B. Cukic. The stochastic precedence ordering with applications in sampling and testing. *J. Appl. Probab.*, 41(1):73–82, 2004.

[19] R. Bordley and M. LiCalzi. Decision analysis using targets instead of utility functions. *Decisions in Economics and Finance*, 23(1):53–74, 2000.

[20] R. F. Bordley and C. W. Kirkwood. Multiattribute preference analysis with performance targets. *Operations Research*, 52(6):823–835, 2004.

[21] P. Brémaud. *Markov chains*, volume 31 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999. Gibbs fields, Monte Carlo simulation, and queues.

[22] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer, 1999.

[23] J. I. Bulow, J. D. Geanakoplos, and P. D. Klemperer. Multimarket oligopoly: Strategic substitutes and complements. *The Journal of Political Economy*, pages 488–511, 1985.

[24] T. Calvo, G. Mayor, and R. Mesiar. *Aggregation operators: new trends and applications*, volume 97. Springer, 2002.

[25] M. Cardin. Aggregation functionals on complete lattices. 2011.

[26] M. Cardin and M. Manzi. Aggregation functions: an approach using copulae. Technical report, 2007.

[27] M. Cardin and M. Manzi. Aggregation functions: A multivariate approach using copulae. *Available at SSRN 1084803*, 2008.

[28] E. Castagnoli and M. LiCalzi. Expected utility without utility. *Theory and Decision*, 41(3):281–301, 1996.

[29] E. Castagnoli and M. LiCalzi. Benchmarking real-valued acts. *Games and Economic Behavior*, 57(2):236–253, 2006.

[30] A. Chateauneuf and M. Cohen. *Cardinal Extensions of the EU Model Based on the Choquet Integral*, pages 401–433. ISTE, 2010.

[31] A. Chateauneuf and J.-Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of mobius inversion. *Mathematical Social Sciences*, 17(3):263 – 283, 1989.

[32] R. Chen and A. Zame. On fair coin-tossing games. *J. Multivariate Anal.*, 9(1):150–156, 1979.

[33] R. W. Chen, A. Zame, and B. Rosenberg. On the first occurrence of strings. *Electron. J. Combin.*, 16(1):1–16, 2009.

[34] G. Choquet. Theory of capacities. In *Annales de linstitut Fourier*, volume 5, page 54, 1953.

[35] L. M. De Campos and M. Jorge. Characterization and comparison of sugeno and choquet integrals. *Fuzzy Sets and Systems*, 52(1):61–67, 1992.

[36] B. De Finetti. *Sur la condition d'"Equivalence partielle."*.

[37] B. De Finetti. *Sulla preferibilità*. Pubblicazioni delle Facoltà di scienze e di ingegneria dell'università di Trieste, serie B. CEDAM, 1952.

[38] E. De Santis, F. Fantozzi, and F. Spizzichino. Relations between stochastic precedence and stochastic orderings, preprint. 2014.

[39] E. De Santis and F. Spizzichino. Change-point models and conditionally pure birth processes: An inequality on the stochastic intensity. *Journal of Applied Probability*, 41(4):pp. 939–952, 2004.

[40] E. De Santis and F. Spizzichino. First occurrence of a word among the elements of a finite dictionary in random sequences of letters. *Electron. J. Probab.*, 17:1–9, 2012.

[41] E. De Santis and F. Spizzichino. Stochastic comparisons between first-passage times for markov chains. *arXiv:1210.1116 [math.PR]*, pages 1–18, 2012.

[42] E. De Santis and F. Spizzichino. Waiting for ABRACADABRA. Occurrences of words and leading numbers. Emmer, Michele (ed.), Imagine Math. Between culture and mathematics. Milano: Springer. 175-185, 2012.

[43] E. De Santis and F. Spizzichino. Random evolution of degradation and occurrences of words in random sequences of letters. *Applied Reliability Engineering and Risk Analysis: Probabilistic Models and Statistical Inference*, page 205, 2013.

[44] D. Denneberg. *Non-additive measure and integral*, volume 27. Springer, 1994.

[45] M. Denuit, L. Eeckhoudt, and B. Rey. Some consequences of correlation aversion in decision science. *Annals of Operations Research*, 176(1):259–269, 2010.

[46] M. Denuit, L. Eeckhoudt, I. Tsetlin, and R. Winkler. Multivariate concave and convex stochastic dominance. 2010.

[47] A. Dukhovny and J. L. Marichal. Reliability of systems with dependent components based on lattice polynomial description. *Stochastic Models*, 28(1):167–184, 2012.

[48] G. T. Duncan. A matrix measure of multivariate local risk aversion. *Econometrica*, 45(4):pp. 895–903, 1977.

[49] F. Durante, E. P. Klement, C. Sempi, and M. Úbeda-Flores. Measures of non-exchangeability for bivariate random vectors. *Statist. Papers*, 51(3):687–699, 2010.

[50] F. Edgeworth. *The Pure Theory of Taxation*. Economic Journal. 1897.

[51] L. Eeckhoudt, B. Rey, and H. Schlesinger. A good sign for multivariate risk taking. *Management Science*, 53(1):117–124, 2007.

[52] D. Ellsberg. Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669, 1961.

[53] L. G. Epstein and S. M. Tanny. Increasing generalized correlation: a definition and some economic consequences. *Canadian Journal of Economics*, pages 16–34, 1980.

[54] F. Fantozzi and F. Spizzichino. Multi-attribute target-based utilities and extensions of fuzzy measures. *Fuzzy Sets and Systems*, (0):–, 2014.

[55] W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley and Sons Inc., New York, 1968.

[56] F. Ferreira and A. Pacheco. Level-crossing ordering of semi-Markov processes and Markov chains. *J. Appl. Probab.*, 42(4):989–1002, 2005.

[57] P. Fishburn. Subjective expected utility: A review of normative theories. *Theory and Decision*, 13(2):139–199, 1981.

[58] R. Fountain. An ordering among generalized closeness criteria. In *Distributions With Given Marginals and Statistical Modelling*, pages 73–79. Springer, 2002.

[59] I. Gilboa. Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics*, 16(1):65 – 88, 1987.

[60] I. Gilboa and D. Schmeidler. Canonical Representation of Set Functions. *Mathematical Methods of Operations Research*, Vol.20:pp.197–212, Feb. 1995.

[61] J. Glaz, M. Kulldorff, V. Pozdnyakov, and J. M. Steele. Gambling teams and waiting times for patterns in two-state Markov chains. *J. Appl. Probab.*, 43(1):127–140, 2006.

[62] M. Grabisch. Fuzzy integral in multicriteria decision making. *Fuzzy Sets and Systems*, 69(3):279 – 298, 1995. Fuzzy Information Processing.

[63] M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89(3):445 – 456, 1996.

[64] M. Grabisch. Alternative representations of discrete fuzzy measures for decision making. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5(05):587–607, 1997.

[65] M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy sets and systems*, 92(2):167–189, 1997.

[66] M. Grabisch. The interaction and mobius representations of fuzzy measures on finite spaces, k-additive measures: a survey. In M. Grabisch, T. Murofushi, M. Sugeno, and J. Kacprzyk, editors, *Fuzzy Measures and Integrals - Theory and Applications*, pages 70–93. Physica Verlag, 2000.

[67] M. Grabisch and C. Labreuche. Bi-capacities-i: definition, möbius transform and interaction. *Fuzzy Sets and Systems*, 151(2):211–236, April 2005.

[68] M. Grabisch and C. Labreuche. Bi-capacities-ii: the choquet integral. *Fuzzy Sets and Systems*, 151(2):237–259, April 2005.

[69] M. Grabisch and C. Labreuche. A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, 175(1):247–290, Mar. 2010.

[70] M. Grabisch, J. L. Marichal, R. Mesiar, and E. Pap. Aggregation functions: Means. *Information Sciences*, 181(1):1 – 22, 2011.

[71] M. Grabisch and M. Roubens. An axiomatic approach to the concept of interaction among players in cooperative games. *International Journal of Game Theory*, 28(4):547–565, 1999.

[72] P. Hammer, S. Rudeanu, and R. Bellmann. *Boolean Methods in Operations Research and Related Areas*. A-Konometrie Und Unternehmensforschung Econometrics and Oper. SPRINGER Publishing Company, 1968.

[73] Y.-C. Hung, R. W. Chen, A. Zame, and M.-R. Chen. A note on the first occurrence of strings. *Electron. J. Combin.*, 17(1):Note 5, 8, 2010.

[74] A. Irle and J. Gani. The detection of words and an ordering for Markov chains. *J. Appl. Probab.*, 38A:66–77, 2001.

[75] P. Jaworski, F. Durante, W. Härdle, and T. Rychlik. *Copula Theory and Its Applications: Proceedings of the Workshop Held in Warsaw, 25-26 September 2009*. Lecture Notes in Statistics / Lecture Notes in Statistics - Proceedings. Springer, 2010.

[76] H. Joe. *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. 1997.

[77] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):pp. 263–292, 1979.

[78] E. Karni. On multivariate risk aversion. *Econometrica: Journal of the Econometric Society*, pages 1391–1401, 1979.

[79] E. Karni. Increasing risk with state-dependent preferences. *Journal of Economic Theory*, 35(1):172–177, 1985.

[80] R. L. Keeney and H. Raiffa. *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press, 1993.

[81] R. E. Kihlstrom and L. J. Mirman. Risk aversion with many commodities. *Journal of Economic Theory*, 8(3):361–388, July 1974.

[82] E. P. Klement and R. Mesiar. How non-symmetric can a copula be? *Commentationes Mathematicae Universitatis Carolinae*, 47(1):141–148, 2006.

[83] E. P. Klement, R. Mesiar, and E. Pap. A universal integral as common frame for choquet and sugeno integral. *Fuzzy Systems, IEEE Transactions on*, 18(1):178–187, 2010.

[84] F. H. Knight. *Risk, uncertainty and profit*. Courier Dover Publications, 2012.

[85] M. M. Köksalan, J. Wallenius, and S. Zionts. *Multiple criteria decision making: from early history to the 21st century*. World Scientific, 2011.

[86] A. Kolesárová, A. Stupnanová, and J. Beganová. Aggregation-based extensions of fuzzy measures. *Fuzzy Sets and Systems*, 194:1 – 14, 2012.

[87] E. L. Lehmann. Some concepts of dependence. *The Annals of Mathematical Statistics*, 37(5):1137–1153, 10 1966.

[88] E. Lehrer. A new integral for capacities. *Economic Theory*, 39(1):157–176, 2009.

[89] S.-Y. R. Li. A martingale approach to the study of occurrence of sequence patterns in repeated experiments. *Ann. Probab.*, 8(6):1171–1176, 1980.

[90] L. Lovász. Submodular functions and convexity. In A. Bachem, B. Korte, and M. Grötschel, editors, *Mathematical Programming The State of the Art*, pages 235–257. Springer Berlin Heidelberg, 1983.

[91] M. J. Machina. ”expected utility” analysis without the independence axiom. *Econometrica: Journal of the Econometric Society*, pages 277–323, 1982.
[92] M. J. Machina. Decision-making in the presence of risk. *Science*, 236(4801):537–543, 1987.
[93] J. L. Marichal. *Aggregation Operators for Multicriteria Decision Aid*. PhD thesis, Institute of Mathematics, University of Liège, Liège, Belgium, 1998.
[94] J.-L. Marichal. An axiomatic approach of the discrete choquet integral as a tool to aggregate interacting criteria. *Fuzzy Systems, IEEE Transactions on*, 8(6):800–807, 2000.
[95] J.-L. Marichal. Behavioral analysis of aggregation in multicriteria decision aid. In *Preferences and decisions under incomplete knowledge*, pages 153–178. Springer, 2000.
[96] J. L. Marichal. Cumulative distribution functions and moments of lattice polynomials. *Statistics & probability letters*, 76(12):1273–1279, 2006.
[97] J. L. Marichal and P. Mathonet. Computing system signatures through reliability functions. *Statistics & Probability Letters*, 83(3):710 – 717, 2013.
[98] M. Marinacci. Decomposition and representation of coalitional games. *Mathematics of Operations Research*, 21(4):1000–1015, 1996.
[99] A. W. Marshall and I. Olkin. A generalized bivariate exponential distribution. Technical report, DTIC Document, 1966.
[100] A. W. Marshall and I. Olkin. A multivariate exponential distribution. *Journal of the American Statistical Association*, 62(317):30–44, 1967.
[101] P. Miranda, M. Grabisch, and P. Gil. Axiomatic structure of k-additive capacities. *Mathematical Social Sciences*, 49(2):153–178, March 2005.
[102] P. Muliere and M. Scarsini. Characterization of a Marshall-Olkin type class of distributions. *Ann. Inst. Statist. Math.*, 39(2):429–441, 1987.
[103] T. Murofushi and S. Soneda. Techniques for reading fuzzy measures (iii): interaction index. In *In 9th Fuzzy System Symposium*, pages 693–696, 1993.
[104] J. Navarro and R. Rubio. Comparisons of coherent systems using stochastic precedence. *TEST*, 19:469–486, 2010.
[105] J. Navarro and F. Spizzichino. On the relationships between copulas of order statistics and marginal distributions. *Statistics & probability letters*, 80(5):473–479, 2010.
[106] R. Nelsen. *An Introduction to Copulas*. Springer Series in Statistics. Springer, 2006.
[107] R. Nelsen. Extremes of nonexchangeability. *Statistical Papers*, 48(4):695–695, 2007.
[108] G. Owen. Multilinear extensions of games. *Management Science*, 18(5-part-2):64–79, 1972.
[109] J. W. Pratt. Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society*, pages 122–136, 1964.
[110] J. W. Pratt and R. J. Zeckhauser. Proper risk aversion. *Econometrica: Journal of the Econometric Society*, pages 143–154, 1987.
[111] J. Quiggin. A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343, 1982.
[112] K. G. Ramamurthy. *Coherent structures and simple games*, volume 6 of *Theory and Decision Library. Series C: Game Theory, Mathematical Programming and Operations Research*. Kluwer Academic Publishers Group, Dordrecht, 1990.
[113] S. Robin and J. J. Daudin. Exact distribution of word occurrences in a random sequence of letters. *J. Appl. Probab.*, 36(1):179–193, 1999.
[114] S. Robin and J.-J. Daudin. Exact distribution of the distances between any occurrences of a set of words. *Ann. Inst. Statist. Math.*, 53(4):895–905, 2001.

[115] G. C. Rota. On the foundations of combinatorial theory i. theory of möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2(4):340–368, 1964.

[116] M. Rothschild and J. E. Stiglitz. Increasing risk: I. a definition. *Journal of Economic Theory*, 2(3):225 – 243, 1970.

[117] F. J. Samaniego. *System signatures and their applications in engineering reliability*, volume 110. Springer, 2007.

[118] P. Samuelson. *Foundations of Economic Analysis*. Economic Studies. Harvard University Press, 1983.

[119] L. J. Savage. *The foundations of statistics*. Courier Dover Publications, 2012.

[120] M. Scarsini. Copulae of capacities on product spaces. *Lecture Notes-Monograph Series*, pages 307–318, 1996.

[121] D. Schmeidler. Integral representation without additivity. *Proceedings of the American mathematical society*, 97(2):255–261, 1986.

[122] C. Sempi. Convergence of copulas: critical remarks. *Rad. Mat.*, 12(2):241–249, 2004.

[123] G. Shafer. *A mathematical theory of evidence*, volume 1. Princeton university press Princeton, 1976.

[124] M. Shaked and J. G. Shanthikumar. *Stochastic orders*. Springer Series in Statistics. Springer, New York, 2007.

[125] L. S. Shapley. A value for n-person games. 1952.

[126] R. Shukla and R. C. Srivastava. The statistical analysis of direct repeats in nucleic acid sequences. *J. Appl. Probab.*, 22(1):15–24, 1985.

[127] K. Siburg and P. Stoimenov. Symmetry of functions and exchangeability of random variables. *Statistical Papers*, 52(1):1–15, 2011.

[128] I. Singer. Extensions of functions of 0-1 variables and applications to combinatorial optimization. *Numerical Functional Analysis and Optimization*, 7(1):23–62, 1985.

[129] F. Spizzichino. The role of signature and symmetrization for systems with non-exchangeable components. *Advances in Mathematical Modeling for Reliability, IOS Press, Amsterdam*, pages 138–148, 2008.

[130] S. Sriboonchita, W. Wong, S. Dhompongsa, and H. Nguyen. *Stochastic Dominance and Applications to Finance, Risk and Economics*. Taylor & Francis, 2010.

[131] V. T. Stefanov and A. G. Pakes. Explicit distributional results in pattern formation. *Ann. Appl. Probab.*, 7(3):666–678, 1997.

[132] V. T. Stefanov, S. Robin, and S. Schbath. Waiting times for clumps of patterns and for structured motifs in random sequences. *Discrete Appl. Math.*, 155(6-7):868–880, 2007.

[133] J. E. Stiglitz. Behavior towards risk with many commodities. *Econometrica: Journal of the Econometric Society*, pages 660–667, 1969.

[134] V. Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, pages 423–439, 1965.

[135] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.

[136] D. M. Topkis. Equilibrium points in nonzero-sum n-person submodular games. *SIAM Journal on Control and Optimization*, 17(6):773–787, 1979.

[137] D. M. Topkis. *Supermodularity and complementarity*. Princeton University Press, 1998.

[138] I. Tsetlin and R. L. Winkler. On equivalent target-oriented formulations for multiattribute utility. *Decision Analysis*, 3(2):94–99, 2006.

[139] I. Tsetlin and R. L. Winkler. Decision making with multiattribute performance targets: The impact of changes in performance and target distributions. *Operations research*, 55(2):226–233, 2007.

[140] I. Tsetlin and R. L. Winkler. Multiattribute utility satisfying a preference for combining good with bad. *Manage. Sci.*, 55(12):1942–1952, Dec. 2009.

[141] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323, 1992.

[142] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior (Commemorative Edition)*. Princeton Classic Editions. Princeton University Press, 2007.

[143] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[144] H.-B. Yan, V.-N. Huynh, T. Ma, and Y. Nakamori. Non-additive multi-attribute fuzzy target-oriented decision analysis. *Information Sciences*, 240:21 – 44, 2013.

[145] H.-B. Yan, V.-N. Huynh, T. Murai, and Y. Nakamori. Kansei evaluation based on prioritized multi-attribute fuzzy target-oriented decision analysis. *Information Sciences*, 178(21):4080 – 4093, 2008.