On the cycle-transitive comparison of artificially coupled random variables

B. De Baets a,*, H. De Meyer b

a Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure links 653, B-9000 Gent, Belgium
b Department of Applied Mathematics and Computer Science, Ghent University, Krijgslaan 281 S9, B-9000 Gent, Belgium

Received 14 July 2006; received in revised form 15 December 2006; accepted 29 May 2007
Available online 4 July 2007

Abstract

Given a collection of random variables, we build a probabilistic relation that, in the case of continuous random variables, expresses for each couple of random variables the probability that the first one takes a greater value than the second one. In order to compute this probability, the random variables are artificially coupled by means of a fixed commutative copula. The main result of this paper pertains to the transitivity of this probabilistic relation. Provided the commutative copula satisfies some additional condition, this transitivity can be described elegantly within the cycle-transitivity framework. It ranges between two known types of transitivity: T_L-transitivity and partial stochastic transitivity.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Comparison of random variables; Copulas; Cycle-transitivity; Frank copula; Probabilistic relation; Transitivity

1. Introduction

Many methods can be established for the comparison of the components (random variables, r.v.) of a random vector (X_1, ..., X_n), as there are many ways to extract useful information from the joint cumulative distribution function (c.d.f.) \( F_{X_1,...,X_n} \) that characterizes the random vector.

A first simplification consists in restricting the comparison strategy to methods that aim at comparing the r.v. two by two. In science it is common practice to regard the joint interaction between three or more entities as the result of their pairwise interactions solely. In probability theory, it means that a method for comparing r.v. should only use the information contained in the bivariate c.d.f. \( F_{X_i,X_j} \). Therefore, from the point of view of model building, one can very well ignore the existence of a multivariate c.d.f. and just describe mutual dependencies between the r.v. by means of the bivariate c.d.f. Of course one should be aware that not all choices of bivariate c.d.f. are compatible with a multivariate c.d.f. The problem of characterizing those ensembles of bivariate c.d.f. that can be identified with the marginal bivariate c.d.f. of a single multivariate c.d.f., is

* Corresponding author. Tel.: +32 9 264 59 41; fax: +32 9 264 62 20.
E-mail addresses: bernard.debaets@ugent.be (B. De Baets), hans.demeyer@ugent.be (H. De Meyer).

0888-613X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijar.2007.05.010
known as the compatibility problem [22]. In this paper, we are not primarily concerned with the compatibility and therefore will refer to the couplings between pairs of r.v. as artificial couplings.

A second simplifying step often made is to bypass the information contained in the bivariate c.d.f. to devise a comparison method that entirely relies on the one-dimensional marginal c.d.f. In this case there is even not a compatibility problem, as for any set of univariate c.d.f. $F_{X_i}$, the product $F_{X_1}F_{X_2} \cdots F_{X_n}$ is a valid joint c.d.f., namely the one expressing the independence of the r.v. The comparison method reduces to the comparison of one-dimensional c.d.f. There are many ways to compare such functions, and by far the simplest one is the method that builds a partial order on the set of r.v. using the principle of first order stochastic dominance. It states that a r.v. $X$ is weakly preferred to a r.v. $Y$ if for all $u \in \mathbb{R}$ it holds that $F_X(u) \leq F_Y(u)$. Other types of stochastic dominance, such as second order stochastic dominance, also appear in various applications [20]. At the extreme end of the chain of simplifications, are the methods that compare r.v. by means of a characteristic or a function of some characteristics derived from the one-dimensional marginal c.d.f. The simplest example is the weak order induced by the expected values of the r.v.

In this paper, we ignore the latter simplifying steps and fully utilize the information contained in the (artificial) bivariate c.d.f. We extend and further analyze a method that we have recently put forward to compare independent random variables in a pairwise manner [8]. More precisely, a so-called probabilistic relation $Q$ is generated, which can be regarded as a graded preference relation, expressing intensities of preference [16]. For discrete r.v., these intensities of preference can in turn be regarded as winning probabilities in a dice game, each r.v. being associated to a (possibly unfair) hypothetical dice with an arbitrary number of faces each containing an arbitrary number of eyes [10].

In our setting, the cornerstone for computing the probabilistic relation $Q$ is the knowledge of the bivariate c.d.f. $F_{X_i,X_j}$ for all couples of r.v. $(X_i,X_j)$. In general, the function $C_{ij}$ that joins the one-dimensional marginal c.d.f. $F_{X_i}$ and $F_{X_j}$ into the bivariate marginal c.d.f. $F_{X_i,X_j}$, i.e. $F_{X_i,X_j} = C_{ij}(F_{X_i},F_{X_j})$ is known as a (commutative) copula [22]. Note that the copulas should not be the same for all pairs of r.v. For a collection of independent r.v. however, they are all equal to the ordinary product $T_{P}(x,y) = xy$ and the pairwise comparison expressed in terms of a probabilistic relation therefore relies upon the knowledge of the one-dimensional marginal c.d.f. solely, as is the case in stochastic dominance methods. Our comparison method, however, is not equivalent to any known kind of stochastic dominance, but should rather be regarded as a graded variant of it.

In this paper, we investigate the case where all copulas used to artificially couple the r.v. are the same but not necessarily equal to the product, neither to the greatest copula $T_{M}(x,y) = \min(x,y)$, the minimum operator, nor to the smallest copula $T_{L}(x,y) = \max(x+y-1,0)$, also known as the Łukasiewicz t-norm. These three special cases have been considered recently by the present authors and in particular it has been revealed that the probabilistic relations generated by these couplings possess transitivity properties that can be nicely characterized. The main theorem of this paper extends these transitivity results to a particular class of copulas, which contains the family of Frank copulas.

Transitivity is a simple, yet powerful property of relations. It plays a decisive role in many fields, such as graph theory, clustering techniques, preference modelling, etc. In preference modelling, for instance, rationality considerations often lead to the demand of transitivity. The concept of transitivity is unique for crisp relations, but for probabilistic relations there is a whole range of transitivity properties. Often, one tries to capture the transitivity in the form of a type of stochastic transitivity, but in our previous work it appeared that the latter framework is too narrow to cover the transitivity properties of the probabilistic relations generated already in the case of independent r.v. On the other hand, also the framework of $T$-transitivity, with $T$ a t-norm, and well known from the theory of fuzzy relations, proved insufficient to deal with the transitivity of probabilistic relations. Instead, we have developed a new framework, called cycle-transitivity framework, that allows to characterize the types of transitivity that will arise in the present investigation.

In the next section, we briefly summarize the main concepts to be used, such as copulas and cycle-transitivity. In Section 3, we recall the method used to compare r.v. and the way a probabilistic relation is generated from it. Section 4 contains the main result of this paper in the form of a theorem. Finally, Section 5 is concerned with the family of Frank copulas.
2. Transitivity frameworks

2.1. Transitivity of fuzzy relations

A binary relation $R$ on a universe $A$ is called transitive if for any $(a, b, c) \in A^3$ it holds that

$$((a, b) \in R \land (b, c) \in R) \Rightarrow (a, c) \in R.$$  \hspace{1cm} (1)

However, many other equivalent formulations may be devised, such as

$$R(a, b) \geq x \land R(b, c) \geq x \Rightarrow R(a, c) \geq x,$$  \hspace{1cm} (2)

for any $x \in [0, 1]$. Alternatively, transitivity can also be expressed in the following functional form:

$$\min(R(a, b), R(b, c)) \leq R(a, c).$$  \hspace{1cm} (3)

Note that on $\{0, 1\}^2$ the minimum operation is nothing else but Boolean conjunction.

In the setting of fuzzy set theory, formulation Equation (3) has led to the popular notion of $T$-transitivity, where a t-norm $T[18]$ is used as generalization of the Boolean conjunction. Recall that a fuzzy relation $R$ on $A$ is an $A^2 \rightarrow [0, 1]$ mapping that expresses the degree of relationship between elements of $A$.

**Definition 1.** Let $T$ be a t-norm. A fuzzy relation $R$ on $A$ is called $T$-transitive if for any $(a, b, c) \in A^3$ it holds that

$$T(R(a, b), R(b, c)) \leq R(a, c).$$  \hspace{1cm} (4)

The three main continuous t-norms are the algebraic product $T_P$ and the Łukasiewicz t-norm $T_L$.

In [4] we have extended the above definition to other binary operations frequently occurring in the field of probability theory and statistics, namely quasi-copulas and copulas. For an excellent monograph on copulas and related operations, we refer to [22].

**Definition 2.** A binary operation $C : [0, 1]^2 \rightarrow [0, 1]$ is called a copula if it satisfies:

(i) Neutral element 1: $(\forall x \in [0, 1])(C(x, 1) = C(1, x) = x)$.

(ii) Absorbing element 0: $(\forall x \in [0, 1])(C(x, 0) = C(0, x) = 0)$.

(iii) Monotonicity: $C$ is increasing in each variable.

(iv) Moderate growth: $(\forall(x_1, x_2, y_1, y_2) \in [0, 1]^4)$

$$(x_1 \leq x_2 \land y_1 \leq y_2) \Rightarrow (C(x_1, y_1) + C(x_2, y_2) \geq C(x_1, y_2) + C(x_2, y_1)).$$

Note that condition (iii) can be omitted (as it follows from (ii) and (iv)). It is well known that a copula is a t-norm if and only if it is associative; conversely, a t-norm $T$ is a copula if and only if it is 1-Lipschitz, i.e. $(\forall(x_1, x_2, y_1, y_2) \in [0, 1]^4)$

$$(|T(x_1, y_1) - T(x_2, y_2)| \leq |x_1 - x_2| + |y_1 - y_2|).$$

Finally, note that for any copula $C$ it holds that $T_L \leq C \leq T_M$.

In analogy to $T$-transitivity, we suggest the following definition.

**Definition 3.** Let $C$ be a copula. A fuzzy relation $R$ on $A$ is called $C$-transitive if for any $(a, b, c) \in A^3$ it holds that

$$C(R(a, b), R(b, c)) \leq R(a, c).$$  \hspace{1cm} (5)

If $C$ is a copula, then the binary operation $\tilde{C}$ defined by

$$\tilde{C}(x, y) = x + y - C(x, y)$$  \hspace{1cm} (6)

is called the dual of $C$, while the binary operation $C^*$ defined by
As limit cases, one obtains

\[ \tilde{C}(x, y) = x + y - 1 + C(1 - x, 1 - y) \]  

is a copula, called the survival copula of \( C \). A copula is called stable if it coincides with its survival copula \([19]\).

A very important family of stable copulas is the family of Frank copulas \((T^F_\lambda)_{\lambda \in [0, \infty]} [15]\). This family is also well known in fuzzy set theory as the family of Frank t-norms. We will use the parametrization commonly used in the latter field, well knowing that in probability theory a different parametrization is used. For \( \lambda \in [0, 1[ \cup ]1, \infty[ \), the Frank copula \( T^F_\lambda \) is defined by

\[ T^F_\lambda(x, y) = \log_\lambda \left( 1 + \frac{(\lambda^x - 1)(\lambda^y - 1)}{\lambda - 1} \right). \]  

As limit cases, one obtains \( T_M (\lambda \to 0) \), \( T_P (\lambda \to 1) \) and \( T_L (\lambda \to \infty) \).

### 2.2. Transitivity of probabilistic relations

Another class of \( A^2 \to [0, 1] \) mappings are the so-called probabilistic relations \( Q \) satisfying

\[ Q(a, b) + Q(b, a) = 1, \]  

for any \((a, b) \in A^2\). For such relations, it holds in particular that \( Q(a, a) = 1/2 \). Transitivity properties for probabilistic relations rather have the logical flavour of Eq. (2). There exist various kinds of stochastic transitivity for probabilistic relations \([1,21]\).

**Definition 4.** Let \( g \) be a commutative increasing \([1/2, 1]^2 \to [1/2, 1] \) mapping. A probabilistic relation \( Q \) on \( A \) is called stochastic transitive w.r.t. \( g \), or shortly \( g \)-stochastic transitive, if for any \((a, b, c) \in A^3\) it holds that

\[ (Q(a, b) \geq 1/2 \land Q(b, c) \geq 1/2) \Rightarrow Q(a, c) \geq g(Q(a, b), Q(b, c)). \]  

This definition includes many well-known types of stochastic transitivity. Indeed, \( g \)-stochastic transitivity is known as \([21]\):

(i) strong stochastic transitivity when \( g = \max \);
(ii) moderate stochastic transitivity when \( g = \min \);
(iii) weak stochastic transitivity when \( g = 1/2 \).

Although \( T \)-transitivity, with \( T \) a t-norm, or \( C \)-transitivity, with \( C \) a copula, on the one hand, and \( g \)-stochastic transitivity on the other hand are of a completely different nature and arise in different contexts, there have been several successful attempts to unify both concepts into a single general framework. One such framework, called \( FG \)-transitivity, has been developed by Switalski \([24,25]\). It formally generalizes \( g \)-stochastic transitivity in the sense that \( Q(a, c) \) is now bounded both from below and above by \([1/2, 1]^2 \to [0, 1] \) mappings.

### 2.3. Cycle-transitivity

The present authors recently introduced the cycle-transitivity framework \([4]\). In analogy to \( FG \)-transitivity, cycle-transitivity involves an upper and a lower bound function; however, instead of \( Q(a, c) \), it is the cyclic invariant sum \( Q(a, b) + Q(b, c) + Q(c, a) \) that is bounded. We will give a brief introduction to cycle-transitivity, as it will turn out to be, to our knowledge, the only appropriate framework for expressing the type of transitivity exhibited by the probabilistic relation generated from the pairwise comparison of a collection of r.v.

In the cycle-transitivity framework \([4]\), for a probabilistic relation \( Q = [q_{ij}] \), the quantities

\[ a_{ijk} = \min(q_{ij}, q_{jk}, q_{ki}), b_{ijk} = \text{med}(q_{ij}, q_{jk}, q_{ki}), c_{ijk} = \max(q_{ij}, q_{jk}, q_{ki}), \]  

(12)
are defined for all \((i,j,k)\). Obviously, \(x_{ijk} \leq y_{ijk} \leq z_{ijk}\). Also, the notation \(\Delta = \{(x,y,z) \in [0,1]^3 | x \leq y \leq z\}\) will be used.

**Definition 5.** A function \(U : \Delta \rightarrow \mathbb{R}\) is called an upper bound function if it satisfies:

(i) \(U(0,0,1) \geq 0\) and \(U(0,1,1) \geq 1\);

(ii) for any \((x,\beta,\gamma) \in \Delta:\)
\[
U(x,\beta,\gamma) + U(1-\gamma,1-\beta,1-x) \geq 1.
\]

The function \(L : \Delta \rightarrow \mathbb{R}\) defined by
\[
L(x,\beta,\gamma) = 1 - U(1-\gamma,1-\beta,1-x)
\]

is called the dual lower bound function of a given upper bound function \(U\).

**Definition 6.** A probabilistic relation \(Q = [q_{ij}]\) is called cycle-transitive w.r.t. an upper bound function \(U\), if for all \((i,j,k)\) it holds that:
\[
L(x_{ijk},\beta_{ijk},\gamma_{ijk}) \leq x_{ijk} + \beta_{ijk} + \gamma_{ijk} - 1 \leq U(x_{ijk},\beta_{ijk},\gamma_{ijk}),
\]

where \(L\) is the dual lower bound function of \(U\).

If (14) holds for some \((i,j,k)\), then due to the built-in duality, it also holds for all permutations of \((i,j,k)\). On the other hand, this duality implies that it is sufficient to verify only the right-hand inequality (or equivalently, only the left-hand inequality) for two permutations of \((i,j,k)\) that are not cyclic permutations of one another, e.g. \((i,j,k)\) and \((k,j,i)\). When the lower bound equals the upper bound, i.e. \(L(a,b,c) = U(a,b,c)\) for all \((a,b,c) \in \Delta\) (in which case the inequalities in (14) become equalities), we say that the function \(U\) is self-dual.

The above definition implies that if a probabilistic relation \(Q\) is cycle-transitive w.r.t. \(U_1\) and \(U_1(a,b,c) \leq U_2(a,b,c)\) for all \((a,b,c) \in \Delta\), then \(Q\) is cycle-transitive w.r.t. \(U_2\). It is clear that \(U_1 \leq U_2\) is not a necessary condition for the latter implication to hold. Two upper bound functions \(U_1\) and \(U_2\) will be called equivalent if for any \((x,\beta,\gamma) \in \Delta\) it holds that \(x + \beta + \gamma - 1 \leq U_1(x,\beta,\gamma)\) is equivalent to \(x + \beta + \gamma - 1 \leq U_2(x,\beta,\gamma)\).

Cycle-transitivity includes as special cases \(C\)-transitivity, with \(C\) a commutative copula, and all known types of \(g\)-stochastic transitivity.

**Proposition 1.** Let \(C\) be a commutative copula. A probabilistic relation \(Q\) is \(C\)-transitive if and only if it is cycle-transitive w.r.t. the upper bound function \(U_C\) defined by
\[
U_C(x,\beta,\gamma) = x + \beta - C(x,\beta).
\]

The operation in Eq. (15) is the dual of the copula \(C\). In particular, if \(C\) belongs to the family of Frank copulas, i.e. \(C = T^\lambda\) for some \(\lambda \in [0,\infty]\), then \(U_C(x,\beta,\gamma) = S^\lambda(x,\beta)\), with \(S^\lambda\) the corresponding Frank copula/t-conorm. The following special cases are of particular interest:

(i) \(T_M\)-transitivity: \(U_M(x,\beta,\gamma) = \max(x,\beta)\);

(ii) \(T_r\)-transitivity: \(U_P(x,\beta,\gamma) = x + \beta - x\beta\);

(iii) \(T_L\)-transitivity: \(U_L(x,\beta,\gamma) = \min(x + \beta,1)\). An equivalent upper bound function is given by \(U'_L(x,\beta,\gamma) = 1\).

**Proposition 2.** Let \(g\) be a commutative, increasing \([1/2,1]^2 \rightarrow [1/2,1]\) mapping such that \(g(1/2,x) \leq x\) for any \(x \in [1/2,1]\). A probabilistic relation \(Q\) on \(A\) is \(g\)-stochastic transitive if and only if it is cycle-transitive w.r.t. the upper bound function \(U_g\) defined by
\[
U_g(x,\beta,\gamma) = \begin{cases} 
\beta + \gamma - g(\beta,\gamma), & \text{if } \beta \geq 1/2 \land x < 1/2, \\
1/2, & \text{if } x \geq 1/2, \\
2, & \text{if } \beta < 1/2.
\end{cases}
\]
We obtain as special cases (only mentioning the expression for the case $\beta \geq 1/2$ and $\alpha < 1/2$):

(i) strong stochastic transitivity: $U_s(\alpha, \beta, \gamma) = \beta$ (a self-dual upper bound function);
(ii) moderate stochastic transitivity: $U_m(\alpha, \beta, \gamma) = \gamma$;
(iii) weak stochastic transitivity: $U_w(\alpha, \beta, \gamma) = \beta + \gamma - 1/2$.

In our former work, a type of transitivity that can neither be classified as a type of $C$-transitivity, nor as a type of $g$-stochastic transitivity, has proven to play a predominant role and this new type of transitivity has been called dice-transitivity [10].

**Definition 7.** Cycle-transitivity w.r.t. the upper bound function $U_D$ defined by

\[ U_D(\alpha, \beta, \gamma) = \beta + \gamma - \beta \gamma, \] (17)

is called dice-transitivity.

Since $U_p \leq U_D \leq U_L$ and also $U_{ms} \leq U_D$, dice-transitivity can be situated between $T_p$-transitivity and $T_L$-transitivity, and also between moderate stochastic transitivity and $T_{L}$-transitivity. We have also shown that dice-transitivity cannot be cast into the $FG$-transitivity framework [3].

Yet another form of transitivity is a slight weakening of moderate stochastic transitivity. A probabilistic relation $Q$ on $A$ is called partially stochastic transitive [13] if for any $(a, b, c) \in A^3$ it holds that

\[ (Q(a, b) > 1/2 \land Q(b, c) > 1/2) \Rightarrow Q(a, c) \geq \min(Q(a, b), Q(b, c)). \]

Also this type of transitivity can be expressed elegantly in the cycle-transitivity framework [7].

**Proposition 3.** Cycle-transitivity w.r.t. the upper bound function $U_{ps}$ defined by

\[ U_{ps}(\alpha, \beta, \gamma) = \gamma \] (18)

is equivalent to partial stochastic transitivity.

### 3. A method for comparing random variables

An immediate way of comparing two r.v. is to consider the probability that the first one takes a greater value than the second one. Proceeding along this line of thought, a random vector $(X_1, X_2, \ldots, X_m)$ generates a probabilistic relation.

**Definition 8.** Given a random vector $(X_1, X_2, \ldots, X_m)$, the binary relation $Q$ defined by

\[ Q(X_i, X_j) = \text{Prob}\{X_i > X_j\} + \frac{1}{2} \text{Prob}\{X_i = X_j\} \] (19)

is a probabilistic relation.

Note that for discrete r.v. $X_i$ and $X_j$, $Q(X_i, X_j)$ is not just the probability that $X_i$ takes a greater value than $X_j$, as half of the probability of a tie is also taken into account. In general, probabilistic relations are not only a convenient tool for expressing the result of the pairwise comparison of a set of alternatives [1], but they also appear in various fields such as game theory [12], voting theory [17,23] and psychological studies on preference and discrimination in (individual or collective) decision making methods [11].

For two discrete r.v. $X_i$ and $X_j$, $Q(X_i, X_j)$ can be computed as

\[ Q(X_i, X_j) = \sum_{k=1}^{L} p_{X_i, X_j}(k, l) + \frac{1}{2} \sum_{k=1}^{L} p_{X_i, X_j}(k, k), \] (20)

with $p_{X_i, X_j}$ the joint probability mass function (p.m.f.) of $(X_i, X_j)$. For two continuous r.v. $X_i$ and $X_j$, $Q(X_i, X_j)$ can be computed as

\[ Q(X_i, X_j) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{x} f_{X_i, X_j}(x, y)dy, \] (21)

with $f_{X_i, X_j}$ the joint probability density function (p.d.f.) of $(X_i, X_j)$.
The above comparison method takes into account the bivariate marginal c.d.f. and hence the pairwise dependence of the components of the random vector. The information contained in the probabilistic relation is therefore much richer than if, for instance, we would have based the comparison of \( X_i \) and \( X_j \) solely on their expected values. Despite the fact that the dependence structure is entirely captured by the multivariate c.d.f., the pairwise comparison is only apt to take into account pairwise dependence, as only bivariate c.d.f. are involved. Indeed, the bivariate c.d.f. do not fully disclose the dependence structure; the r.v. may even be pairwise independent while not mutually independent.

Since the copulas \( C_{ij} \) that couple the univariate marginal c.d.f. into the bivariate marginal c.d.f. can be different from another, the analysis of the probabilistic relation and in particular the identification of its transitivity properties appear rather cumbersome. It is nonetheless possible to state in general, without making any assumptions on the bivariate c.d.f., that the probabilistic relation \( Q \) generated by an arbitrary random vector always shows some minimal form of transitivity.

For fuzzy relations, \( T_m \)-transitivity is the type of transitivity closest to the transitivity of crisp relations as all \( \alpha \)-cuts of a \( T_m \)-transitive fuzzy relation are transitive crisp relations [26]. At the other end of the ‘transitivity scale’, one could very well imagine a situation of total absence of any type of transitivity in the sense that,\( Q(a,b) \) and \( Q(b,c) \), there is no restriction at all on the value of \( Q(a,c) \). In practice, however, fuzzy or probabilistic relations resulting from pairwise comparison methods, whether it concerns, e.g. physical or economical processes, or processes related to human reasoning and behaviour, often show some minimal form of transitivity. This ‘weakest’ type of transitivity apparently inherent to reality is \( T_1 \)-transitivity. It frequently arises in the context of similarity measurement [2,5], and is strongly connected to the triangle inequality in metric spaces [6]. Not surprisingly, the probabilistic relation \( Q \) generated by a random vector also shows this type of transitivity.

**Proposition 4.** The probabilistic relation \( Q \) generated by a random vector is \( T_L \)-transitive.

**Proof.** Let \((X, Y, Z)\) be any components of the given random vector. We first consider the case of continuous r.v. characterized by a 3-dimensional joint p.d.f. \( f_{X,Y,Z} \). Since the event that two or more r.v. take the same value has zero probability, we consider the following mutually exclusive events: (1) \( X > Y > Z \), (2) \( X > Z > Y \), (3) \( Y > X > Z \), (4) \( Y > Z > X \), (5) \( Z > X > Y \) and (6) \( Z > Y > X \) and we generically denote by \( f_i, i \in \{1, \ldots, 6\} \), the integral of \( f_{X,Y,Z} \) on the domain described by the double inequality numbered (i). For instance,

\[
f^1 = \iint_{x>y>z} f_{X,Y,Z}(x,y,z) \, dx \, dy \, dz.
\]

Obviously, \( \sum_{i=1}^{6} f_i = 1 \). Since \( Q(X,Y) = \text{Prob}\{X > Y\} \), one immediately obtains that \( Q(X,Y) = f^1 + f^2 + f^3 \). Similarly, \( Q(Y,Z) = f^4 + f^5 + f^6 \) and \( Q(Z,X) = f^7 + f^8 + f^9 \). It follows that \( Q(X,Y) + Q(Y,Z) + Q(Z,X) - 1 = f^1 + f^2 + f^3 + f^4 + f^5 + f^6 \leq 1 \), which proves that \( Q \) is \( T_1 \)-transitive.

Next, consider the case of discrete r.v. characterized by a 3-dimensional joint p.m.f. \( f_{X,Y,Z} \). In this case, it is necessary to keep the notation \( f \) for probability masses. Besides the six events described before, seven more events have to be distinguished, namely (7) \( X = Y > Z \), (8) \( X = Z > Y \), (9) \( Z > X = Y \), (10) \( Y > X = Z \), (11) \( Y = Z > X \), (12) \( X > Y = Z \) and (13) \( X = Y = Z \). We denote by \( f^i \) the sum of \( f_{X,Y,Z} \) on the domain described by the double inequality–equality numbered (i). Clearly, it holds that \( \sum_{i=1}^{13} f_i = 1 \). Since \( Q(X,Y) = \text{Prob}\{X > Y\} + \frac{1}{2} \text{Prob}\{X = Y\} \), it also holds that

\[
Q(X,Y) = f^1 + f^2 + f^3 + f^4 + f^5 + f^6 + f^7 + f^8 + f^9 + f^{10} + f^{12} + \frac{1}{2} f^{13}.
\]

Combining \( Q(X,Y) \) with similar expressions for \( Q(Y,Z) \) and \( Q(Z,X) \), we finally obtain that \( Q(X,Y) + Q(Y,Z) + Q(Z,X) - 1 = f^1 + f^2 + f^3 + f^4 + f^5 + f^6 + f^7 + f^8 + f^9 + f^{10} + f^{12} + \frac{1}{2} f^{13} \leq 1 \), which again shows that \( Q \) is \( T_L \)-transitive. \( \Box \)

Our further interest is to study the situation where (momentarily) abstraction is made that the r.v. are components of a random vector, and all bivariate c.d.f. are enforced to depend in the same way upon the univariate c.d.f., in other words, we consider the situation of all copulas being the same, well knowing that this might not be possible at all. In fact, this simplification is equivalent to considering instead of a random vector, a collection of r.v. and to artificially compare them, all in the same manner and based upon a same copula.
To get insight in what kind of transitivity properties one might expect in general, the present authors have previously unravelled three particular cases, namely the case of the product copula $T_P$, and the cases of the two extreme copulas, the minimum operator $T_M$ and the Łukasiewicz t-norm $T_L$, respectively related to a presumed but not-necessarily existing comonotonic and countermonotonic pairwise dependence of the r.v. [22]. From these studies the following results can be reported.

**Proposition 5 ([8,10]).** The probabilistic relation $Q$ generated by a collection of random variables pairwisely coupled by $T_P$ is dice-transitive, i.e. it is cycle-transitive w.r.t. to the upper bound function $U_D$. In particular, the probabilistic relation generated by a collection of independent r.v. is dice-transitive.

**Proposition 6 ([7,9]).** The probabilistic relation $Q$ generated by a collection of random variables pairwisely coupled by $T_M$ is cycle-transitive w.r.t. to the upper bound function $U$ given by

$$U(x, \beta, \gamma) = \min(\beta + \gamma, 1).$$

Cycle-transitivity w.r.t. the upper bound function $U$ is equivalent to cycle-transitivity w.r.t. $U_L$ or $U_0$.

**Proposition 7 ([7,9]).** The probabilistic relation $Q$ generated by a collection of random variables pairwisely coupled by $T_L$ is partially stochastic transitive, i.e. it is cycle-transitive w.r.t. to the upper bound function $U_{ps}$.

Let us recall that the proofs were first given for discrete uniformly distributed r.v. [9,10]. It allowed for an interpretation of the values $Q(X_i, X_j)$ as winning probabilities in a hypothetical dice game, or equivalently, as a method for the pairwise comparison of ordered lists of numbers. Subsequently, we have shown that as far as transitivity is concerned, this situation is generic and therefore characterizes the type of transitivity observed in general [7,8]. Note that this does not contradict the fact that if the r.v. possess distributions that belong to certain parametrized families, the generated probabilistic relation can ‘gain’ transitivity [8]. In the present paper, we will not deal with special families of marginal distributions.

### 4. Transitive comparison of artificially coupled random variables

In this section, we derive the main result of the paper. It concerns the transitivity of the probabilistic relation generated by a collection of random variables that are pairwisely (and artificially) coupled by a same commutative copula $C$.

In the extensive and intricate proof, we will consider as generic situation continuous r.v. that are uniformly distributed on finitely countable unions of finite intervals. This generalizes the situation of uniformly distributed discrete r.v. that was used as generic situation in the proofs for the special cases of $C$ being either $T_P$, $T_M$ or $T_L$.

In the following, let $X$, $Y$ and $Z$ be three continuous r.v. with p.d.f. given by

$$f_X(u) = \begin{cases} \frac{1}{|D_X|}, & \text{if } u \in D_X, \\ 0, & \text{elsewhere,} \end{cases}$$

$$f_Y(u) = \begin{cases} \frac{1}{|D_Y|}, & \text{if } u \in D_Y, \\ 0, & \text{elsewhere,} \end{cases}$$

$$f_Z(u) = \begin{cases} \frac{1}{|D_Z|}, & \text{if } u \in D_Z, \\ 0, & \text{elsewhere,} \end{cases}$$

where $D_X$, $D_Y$ and $D_Z$ are finitely countable unions of finite real intervals such that $D_X \cap D_Y \cap D_Z = \emptyset$; also, $|\cdot|$ denotes the total length. Note that $X$, $Y$ and $Z$ are uniformly distributed on disjoint subsets of the real line.

To compare $X$ and $Y$ we use the copula $C$. Since $X$ and $Y$ cannot take the same value, it follows that

$$Q(X, Y) = \Pr\{X > Y\} = \int_{x>y} dF_{X,Y}.$$
with \( F_{X,Y}(x,y) = C(F_X(x), F_Y(y)) \) the bivariate c.d.f. of \((X, Y)\). Let us introduce the following graphical representation of the probability \( Q(X, Y) \). We draw a 2-dimensional \((F_X(x), F_Y(y))\)-plot, where \( x \) and \( y \) run from \(-\infty\) to \(+\infty\). With each couple \((x, y)\) corresponds a point \((F_X(x), F_Y(y))\) in the unit square. As depicted in Fig. 1, the curve \((F_X(u), F_Y(u)), u \in ]-\infty, +\infty[\) that separates the regions where \( x > y \) and \( x < y \), has the shape of a staircase which departs either from the bottom line or the left vertical line and ends either at the top line or the right vertical line. In the following, we call this demarcation line in the \((F_X, F_Y)\)-plot the \((X, Y)\)-staircase, and we will make a distinction between ‘noses’ and ‘holes’ for the turning points on this staircase. The staircase shape follows from the fact that \( X \) and \( Y \) are uniformly distributed on disjoint subintervals of the real line.

The computation of probabilities proceeds as follows

\[
\text{Prob} \{ x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2 \} \\
= F_{X,Y}(x_1, y_1) + F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) \\
= C(s_1, t_1) + C(s_2, t_2) - C(s_1, t_2) - C(s_2, t_1),
\]

where \( s_1 = F_X(x_1), \) \( s_2 = F_X(x_2), \) \( t_1 = F_Y(y_1) \) and \( t_2 = F_Y(y_2). \) The points \((s_1, t_1)\) and \((s_2, t_2)\) determine a rectangular domain in the \((F_X, F_Y)\)-plot (the dashed rectangle in Fig. 1) and the probability is computed by means of the \( C\)-weighted area of the rectangle. We call the probability associated with this rectangle and copula \( C\), the \( C\)-weighted area of the rectangle. To compute the probability \( Q(X, Y) \) we must partition the region \( x > y \) of Fig. 1 into rectangular parts and sum up the probabilities associated with these rectangles, in other words, we compute the \( C\)-weighted area of the region \( x > y \). Note that since \( C \) is a copula, the \( C\)-weighted area is non-negative.

Assume now that \( Q(X, Y) \) has a fixed value \( \lambda \) with \( 0 \leq \lambda \leq 1 \). There are many uniform distributions of \( X \) and \( Y \) on unions of intervals that yield this probability \( \lambda \), or, in graphical terms, there are many \((X, Y)\)-staircases that delimit a region \( x > y \) with \( C\)-weighted area equal to \( \lambda \). We can distinguish two extreme cases: the \((X, Y)\)-staircase is either a single horizontal or a single vertical line. These extreme cases are depicted in Fig. 2.

In Fig. 2a the region \( x > y \) is the rectangle below the horizontal line and its \( C\)-weighted area is given by

\[
C(0, 0) + C(1, \lambda) - C(0, \lambda) - C(1, 0) = \lambda.
\]

Similarly, in the other extreme case Fig. 2b, the region \( x > y \) is the rectangle to the right of the vertical line with \( C\)-weighted area given by

\[
C(1 - \lambda, 0) + C(1, 1) - C(1 - \lambda, 1) - C(0, 1) = 1 - (1 - \lambda) = \lambda.
\]

Herewith, we have gathered the ingredients that allow us to state and prove the main theorem.

**Theorem 1.** Let \( C \) be a commutative copula such that for any \( n > 1 \) and for any \( 0 \leq x_1 \leq \cdots \leq x_n \leq 1 \) and \( 0 \leq y_1 \leq \cdots \leq y_n \leq 1 \), it holds that

\[
\sum_i C(x_i, y_i) - \sum_i C(x_{n-2i}, y_{n-2i-1}) - \sum_i C(x_{n-2i-1}, y_{n-2i}) \\
\leq C \left( x_n + \sum_i C(x_{n-2i}, y_{n-2i-1}) - \sum_i C(x_{n-2i-1}, y_{n-2i-1}), y_n + \sum_i C(x_{n-2i-1}, y_{n-2i-2}) - \sum_i C(x_{n-2i-1}, y_{n-2i}) \right),
\]

\[(22)\]

![Fig. 1. \((X, Y)\)-staircase delimiting the region \( x > y \) in the 2-dimensional \((F_X, F_Y)\)-plot.](image)
where the sums extend over all integer values that lead to meaningful indices of $x$ and $y$. Then the probabilistic relation $Q$ generated by a collection of random variables pairwisely coupled by $C$ is cycle-transitive w.r.t. to the upper bound function $U^C$ defined by:

$$U^C(a, b, c) = \max(b + C(1/b, \gamma), c + C(b, 1/\gamma)).$$

(23)

If $C$ is stable, then

$$U^C(a, b, c) = b + C(1/b, \gamma) = \gamma + C(b, 1/\gamma).$$

(24)

**Proof.** Let $X$, $Y$ and $Z$ be three uniformly distributed r.v. on domains $D_X$, $D_Y$ and $D_Z$, and pairwisely coupled by the commutative copula $C$. Assume that $Q(X, Z)$ and $Q(Z, Y)$ are given. We want to find the extreme values that can be attained by $Q(X, Y)$. Throughout the proof, we will use the shorter notation $q_{XY}$ instead of $Q(X, Y)$.

There exist many c.d.f. $F_X$ and $F_Z$ for which Prob{$X > Z$}, computed with $C$, has the given value $q_{XZ}$. In graphical terms, with any $(X, Z)$-staircase in the $(F_X, F_Z)$-plot such that the region $x > z$ has $C$-weighted area equal to $q_{XZ}$, we can associate appropriate c.d.f. $F_X$ and $F_Z$. Similarly, there exist many c.d.f. $F_Z$ and $F_Y$ such that the $(Z, Y)$-staircase in the $(F_Z, F_Y)$-plot delimits a region $z > y$ that has $C$-weighted area equal to $q_{ZY}$. Assume that such an $(X, Z)$-staircase and a $(Z, Y)$-staircase have been chosen. We then superpose the two plots on a single $(F_X, F_Y)$-plot as shown in Fig. 3.

In the $(F_X, F_Y)$-plot, the staircase that delimits the intersection of the regions $x > z$ and $z > y$ from the $(F_X, F_Z)$- and $(F_Y, F_Y)$-plots, is the region for which it definitely holds that $x > y$ and its $C$-weighted area is therefore a lower bound for $q_{XY}$. On the other hand, the staircase that delimits the union of the regions $x > z$ and $z > y$, is the largest region for which it can possibly hold that $x > y$, and therefore its $C$-weighted area is an

![Fig. 2](image-url) The two extreme cases of delimiting the region $x > y$ such that $Q(X, Y) = \lambda$.

![Fig. 3](image-url) An $(X, Z)$-staircase and a $(Z, Y)$-staircase superposed in an $(F_X, F_Y)$-plot.
upper bound for $q_{XY}$. Indeed, for any point outside the union of the two regions $x \geq z$ and $z \geq y$, it holds that $x < z$ and $z < y$ and therefore also that $x < y$.

As we are only interested in the upper bound for $q_{XY}$, from here onwards we only consider the union of the regions $x > z$ and $z > y$. We must determine for which $(X, Z)$- and $(Z, Y)$-staircases the $C$-weighted area of the union of the regions is maximal and then prove that there exist uniform distributions $F_X$, $F_Y$ and $F_Z$ for which this maximum is attained. To get an idea of what the maximum could be, let us consider the situations where the $(X, Z)$- and $(Z, Y)$-staircases are either horizontal or vertical lines. Four combinations should be distinguished.

Firstly, if the $(X, Z)$-staircase is a vertical line and the $(Z, Y)$-staircase is a horizontal line, such that $q_{XZ} = 1 - q_{ZX}$ and $q_{ZY} = 1 - q_{YZ}$ are fixed, we find the situation as depicted in Fig. 4. To obtain an upper bound for $q_{XY}$, we have to compute the $C$-weighted area of the union of the two regions. Partitioning this union as shown in Fig. 4, we obtain:

$$q_{xy} \leq [C(0, 0) + C(1 - q_{XZ}, q_{ZY}) - C(0, q_{ZY}) - C(1 - q_{XZ}, 0)] + q_{XZ}$$

$$= q_{XZ} + C(1 - q_{XZ}, q_{ZY})$$

$$= 1 - q_{XZ} + C(q_{XZ}, 1 - q_{ZY}),$$

from which it follows that

$$q_{XY} + q_{YZ} + q_{XZ} - 1 \leq q_{XY} + C(q_{XZ}, 1 - q_{ZY}). \quad (25)$$

Note that in this case, the upper bound for $q_{XY}$ can be attained for instance with the choice

$$D_x = |1 - q_{YZ}, 1 - q_{ZY} + q_{XZ}[\cup]2 + q_{XZ}, 3|,$$

$$D_y = |0, 1 - q_{ZY}[\cup]2 - q_{YZ} + q_{XZ}, 2 + q_{XZ}, |,$$

$$D_z = |1 - q_{YZ} + q_{XZ}, 2 - q_{YZ} + q_{XZ}, |.$$

Since $(q_{XY}, q_{YZ}, q_{XZ})$ can be any permutation of $(x, \beta, \gamma)$, the upper bound condition implies the condition

$$x + \beta + \gamma - 1 \leq \max(x + C(1 - x, \beta), x + C(1 - z, \gamma), \beta + C(1 - \beta, x), \beta + C(1 - \beta, \gamma),$$

$$\gamma + C(1 - \gamma, x), \gamma + C(1 - \gamma, \beta)).$$

Since $C$ is increasing, it holds that $C(1 - x, \beta) \leq C(1 - x, x)$, $C(1 - \beta, \beta) \leq C(1 - \beta, \gamma)$ and $C(1 - \gamma, x) \leq C(1 - \gamma, \beta)$, and since $C$ is 1-Lipschitz, it holds that $x + C(1 - x, \gamma) \leq \beta + C(1 - \beta, \gamma)$. The upper bound condition therefore reduces to

$$x + \beta + \gamma - 1 \leq \max(\beta + C(1 - x, \gamma), \gamma + C(1 - x, \beta)). \quad (26)$$

Note that the right-hand side of Eq. (26) is the upper bound stated in the theorem. As we already know that this upper bound can be attained, the proof will be complete if we can show that there does not exist an attainable higher bound. In fact, we will prove a somewhat stronger result, namely that no higher theoretical bound can be obtained, even making abstraction of the fact whether it is attainable or not. Secondly, with

![Fig. 4. Illustration of the computation of an upper bound for $q_{XY}$ from extreme $(X, Z)$- and $(Z, Y)$-staircases, the values of $q_{XZ}$ and $q_{YZ}$ being given.](image)
a horizontal \((X, Z)\)-staircase and a vertical \((Z, Y)\)-staircase, we obtain the situation depicted inFig. 5. The theoretical upper bound is now given by
\[
q_{XY} \leq \max(q_{XZ}, 1 - q_{ZY}) - C(0, 1 - q_{ZY}) - C(q_{XZ}, 0) + q_{ZY}
\]
\[
= q_{ZY} + C(q_{XZ}, 1 - q_{ZY}) - 1 = q_{YZ} + C(1 - q_{XZ} q_{YZ}),
\]
from which it follows that
\[
q_{XY} + q_{YZ} + q_{ZX} - 1 \leq q_{ZY} + C(1 - q_{XZ}, q_{YZ}).
\]
Note that the only difference with the previous situation is that \(q_{YZ}\) and \(q_{ZX}\) have changed roles. It can be shown that this theoretical bound, though it yields the same upper bound condition Eq. (26) as before, is not attainable in the setting of Fig. 5. Thirdly, let the \((X, Z)\)- and \((Z, Y)\)-staircases be both horizontal as shown in Fig. 6. It follows that
\[
q_{XY} \leq \max(q_{XZ}, q_{ZY}) = 1 - \min(q_{YZ}, q_{XZ}),
\]
whence
\[
q_{XY} + q_{YZ} + q_{ZX} - 1 \leq \max(q_{IZ}, q_{XZ}),
\]
which yields the upper bound condition
\[
\alpha + \beta + \gamma - 1 \leq \gamma.
\]
It can be shown that this upper bound can be attained, but as it is not greater than the upper bound in Eq. (26), we can ignore this situation. Finally, the reader can easily verify that if the \((X, Z)\)- and \((Z, Y)\)-staircases are vertical lines, then the same theoretical bound Eq. (28) is obtained and also this situation can therefore be ignored.

![Fig. 5. Illustration of the computation of an upper bound of \(q_{XY}\) from extreme \((X, Z)\)- and \((Z, Y)\)-staircases, the values of \(q_{XZ}\) and \(q_{ZY}\) being given.](image)

![Fig. 6. Illustration of the computation of an upper bound for \(q_{XY}\) from extreme \((X, Z)\)- and \((Z, Y)\)-staircases, the values of \(q_{XZ}\) and \(q_{ZY}\) being given.](image)
We now will prove that for given \( q_{XZ} \) and \( q_{ZY} \), the \( C \)-weighted area of the union of the two regions delimited by any \((X, Z)\)- and \((Z, Y)\)-staircases, cannot be greater than the \( C \)-weighted area of the union of the two regions delimited by the vertical and horizontal staircase, respectively. The proof goes in two steps.

Firstly, we show that it is sufficient to consider only those combinations of \((X, Z)\)- and \((Z, Y)\)-staircases that yield an \((X, Y)\)-staircase (delimiting the union of the regions \( x > z \) and \( z > y \)) of which the noses are taken alternately from the noses of the \((X, Z)\)-staircase and the \((Z, Y)\)-staircase. Indeed, assume for instance that two or more consecutive noses of the \((X, Y)\)-staircase are noses of the \((X, Z)\)-staircase, as illustrated in Fig. 7a (three noses on the staircase drawn with full lines, the other staircase being drawn with small dashes). The section from \( a \) to \( e \) of the \((X, Z)\)-staircase is modified into a new section that contains only one nose (section \( a-b-c-d-e \) drawn with large dashes) such that the \( C \)-weighted area of the \( x > z \) region under the modified \((X, Z)\)-staircase remains constant. It implies that the \( C \)-weighted area of the shaded region in the middle of Fig. 7b exactly compensates the sum of the \( C \)-weighted areas of the other two shaded domains. Consequently, the \( C \)-weighted area of the union of the two regions after transformation is at least as great as the \( C \)-weighted area of the union before transformation. Now, we carry out one more \( C \)-weighted area-preserving transformation on the \((Z, Y)\)-staircase, such that it contains only one hole where it lies below the transformed section of the \((X, Z)\)-staircase. This is illustrated in Fig. 7c. By means of well-chosen transformations, it is therefore always possible to obtain the maximal \( C \)-weighted area by means of intertwined \((X, Z)\)- and \((Z, Y)\)-staircases that alternately deliver the noses of the \((X, Y)\)-staircase, and forming a chain of rectangular boxes in the \((F_X, F_Y)\)-plot.

For the second step of the proof, we need to consider arbitrary intertwined staircases, or equivalently, arbitrary chains of rectangular boxes. To get insight in this general situation, let us first investigate the case of a chain of three rectangular boxes, and consider the middle box, as illustrated in Fig. 8. The \( C \)-weighted area of the region to the right and below the \((X, Z)\)-staircase (full line) is \( q_{XZ} \), while the \( C \)-weighted area of the region to the right and below the \((Z, Y)\)-staircase (dashed line) is \( q_{ZY} \). Let us compare the \( C \)-weighted area of the union of these regions to that of the union of the rectangular regions obtained by taking e.g. the horizontal line with ordinate \( 1 - q_{ZX} = q_{XZ} \) as the \((X, Z)\)-staircase and the vertical line with abscissa \( q_{YZ} \) as \((Z, Y)\)-staircase. Let us assume that \( y_1 \leq 1 - q_{ZX} \leq y_2 \) and \( x_1 \leq q_{YZ} \leq x_2 \). The shaded region labelled I which extends from \( x_1 \) to \( q_{YZ} \) and from \( 1 - q_{ZX} \) to \( y_1 \) is the region that is maximally added to the union of the two rectangular domains, whereas the shaded rectangular regions labelled II and III, extending from \( 0 \) to \( x_1 \) and from \( y_1 \) to \( 1 - q_{ZX} \), resp. from \( q_{YZ} \) to \( x_2 \) and from \( y_2 \) to \( 1 \), are the regions that are minimally subtracted from that union. It follows that the \( C \)-weighted area of the union of the two rectangular domains is a maximum if the \( C \)-weighted area of I is at most equal to the sum of the \( C \)-weighted areas of II and III, i.e. if

\[
C(q_{YZ}, y_2) + C(x_1, 1 - q_{ZX}) - C(x_1, y_2) - C(q_{YZ}, 1 - q_{ZX}) \\
\leq C(x_1, 1 - q_{ZX}) - C(x_1, y_1) + s + C(q_{YZ}, y_2) - q_{YZ} - C(x_2, y_2).
\]

Since \( y_2 - C(x_1, y_2) = 1 - q_{ZX} \) and \( x_2 - C(x_2, y_1) = q_{YZ} \), this condition is equivalent to the condition

\[
C(x_1, y_1) + C(x_2, y_2) - C(x_1, y_2) - C(x_2, y_1) \leq C(x_2 - C(x_2, y_1), y_2 - C(x_1, y_2))
\]

which gives the condition (30).

Fig. 7. Illustration of how staircases can be transformed into intertwined staircases forming rectangular boxes.
which is equivalent to Eq. (22) with \( n = 2 \) and is therefore fulfilled. We leave it to the reader to verify that if \( 1 - q_{ZX} < y_1 \) or \( q_{YZ} < x_1 \) the same condition is retrieved.

Let us finally compare the \( C \)-weighted area of the union of the regions \( x > z \) and \( z > y \) delimited by staircases to the \( C \)-weighted area of the union of the rectangular regions delimited by the vertical line with abscissa \( q_{ZX} \) as extreme \((X, Z)\)-staircase and the horizontal line with ordinate \( 1 - q_{YZ} \) as extreme \((Z, Y)\)-staircase. A completely similar derivation leads to the conclusion that the \( C \)-weighted area of the union of the regions delimited by the \((X, Z)\)- and \((Z, Y)\)-staircases is not greater than the \( C \)-weighted area of the union of the regions delimited by the vertical and horizontal lines, provided it holds that

\[
C(x_1, y_1) + C(x_2, y_2) \leq x_2 + y_2 - 1 + C(1 - x_2 + C(x_2, y_1), 1 - y_2 + C(x_1, y_2)).
\] (31)

This inequality is obtained from Eq. (22) by setting \( n = 3 \) and choosing \( x_3 = y_3 = 1 \) and therefore is fulfilled. Continuing in the same line as for the case of a chain of three rectangular boxes and with the same notational conventions, the following conditions are easily derived. If the nose \((x_1, y_2)\) of the \((X, Y)\)-staircase belongs to the \((X, Z)\)-staircase, then for even \( n = 2m \), the condition is

\[
\sum_i C(x_i, y_i) - \sum_i C(x_{2i-1}, y_{2i}) - \sum_i C(x_{2i}, y_{2i-1}) \\
\leq C(x_{2m} + \sum_i C(x_{2i}, y_{2i+1}) - \sum_i C(x_{2i}, y_{2i-1}), y_{2m} + \sum_i C(x_{2i+1}, y_{2i+1}) - \sum_i C(x_{2i-1}, y_{2i+1})),
\] (32)

while for \( n = 2m + 1 \) the condition is

\[
\sum_i C(x_i, y_i) - \sum_i C(x_{2i-1}, y_{2i}) - \sum_i C(x_{2i}, y_{2i-1}) \\
\leq x_{2m+1} + y_{2m+1} - 1 + C \left( 1 - x_{2m+1} + \sum_i C(x_{2i+1}, y_{2i}) - \sum_i C(x_{2i-1}, y_{2i}), \\
1 - y_{2m+1} + \sum_i C(x_{2i}, y_{2i+1}) - \sum_i C(x_{2i}, y_{2i-1}) \right).
\] (33)

If the nose \((x_1, y_2)\) of the \((X, Y)\)-staircase belongs to the \((Z, Y)\)-staircase, then for even \( n = 2m \), the condition is

\[
\sum_i C(x_i, y_i) - \sum_i C(x_{2i+1}, y_{2i}) - \sum_i C(x_{2i}, y_{2i+1}) \\
\leq x_{2m} + y_{2m} - 1 + C \left( 1 - x_{2m} + \sum_i C(x_{2i}, y_{2i-1}) - \sum_i C(x_{2i+1}, y_{2i+1}), \\
1 - y_{2m} + \sum_i C(x_{2i-1}, y_{2i}) - \sum_i C(x_{2i+1}, y_{2i}) \right).
\] (34)
while for \( n = 2m + 1 \) the condition is
\[
\sum_i C(x_i, y_i) - \sum_i C(x_{2i+1}, y_{2i}) - \sum_i C(x_{2i}, y_{2i+1})
\leq C \left( x_{2m+1} + \sum_i C(x_{2i-1}, y_{2i}) - \sum_i C(x_{2i+1}, y_{2i}), y_{2m+1} + \sum_i C(x_{2i}, y_{2i-1}) - \sum_i C(x_{2i}, y_{2i+1}) \right).
\] (35)

Note that all these conditions are fulfilled, as Eq. (32) is equivalent to Eq. (22) with \( n = 2m \), Eq. (33) is equivalent to Eq. (22) with \( n = 2m + 2 \) and \( x_{2m+2} = y_{2m+2} = 1 \), Eq. (34) is equivalent to Eq. (22) with \( n = 2m + 1 \) and \( x_{2m+1} = y_{2m+1} = 1 \), and Eq. (35) is equivalent to Eq. (22) with \( n = 2m + 1 \). Note also that if \( C \) is stable, then Eq. (32) is equivalent to Eqs. (34) and (33) is equivalent to Eq. (35).

To conclude, if \( C \) satisfies Eq. (22), then the maximal \( C \)-weighted area is obtained with extreme \((X, Z)\)- and \((Z, Y)\)-staircases, from which it follows that the function \( U^C \) defined in Eq. (23) is the upper bound function that determines the cycle-transitivity property of the probabilistic relation \( Q \) if the r.v. are uniformly distributed on disjoint sets of intervals. If the r.v. have arbitrary c.d.f., the staircases become curves that have the property that for any point on them all the points to the left of and below that point belong to the region of interest. It is clear that such a curve can always be approximated as closely as desired by a staircase (in the sense that the \( C \)-weighted area of the region enclosed by the curve and the staircase approximation can be made as small as required). As none of the combinations of staircases provides a greater upper bound than the one obtained with the extreme staircases, also in the limit none of the combinations of delimiting curves yields a greater upper bound. The situation of r.v. that are uniformly distributed on disjoint sets of intervals is therefore generic as far as the transitivity of \( Q \) is concerned.

One can easily verify that if \( C \) is a stable copula, then it holds that
\[
\beta + C(1 - \beta, \gamma) = \gamma + C(\beta, 1 - \gamma),
\]
from which Eq. (24) immediately follows. Note that without the concept of cycle-transitivity, it would be very difficult to describe this type of transitivity in a compact manner. This completes the proof of the theorem. \( \square \)

5. Artificial coupling with Frank copulas

It is natural to ask whether commutative copulas that fulfil condition Eq. (22) can be characterized in an alternative way. We consider again the case \( n = 2 \) which corresponds to Eq. (30):
\[
C(x_1, y_1) + C(x_2, y_2) - C(x_1, y_2) - C(x_2, y_1) \leq C(x_2 - C(x_2, y_1), y_2 - C(x_1, y_2)),
\]
for all \( 0 \leq x_1 \leq x_2 \leq 1 \) and \( 0 \leq y_1 \leq y_2 \leq 1 \). The following example shows that this condition is not necessarily satisfied for any stable commutative copula.

Example 1. Consider the commutative copula \( C \) defined by
\[
C(x, y) = \begin{cases} 
\frac{1}{3} + \max(x + y - 1, 0), & \text{if } (x, y) \in [1/3, 2/3]^2, \\
\min(x, y), & \text{elsewhere}.
\end{cases}
\] (36)

It is the ordinal sum \((1/3, 2/3, T_1)\) with \( T_1 \) linearly rescaled to the square \([1/3, 2/3]^2\). It is easily verified that \( C \) is stable (as it is a ‘symmetrical’ ordinal sum of Frank copulas [19]). Let \( x_1 = y_1 = 1/4 \) and \( x_2 = y_2 = 3/4 \). The left-hand side of Eq. (30) becomes \( C(1/4, 1/4) + C(3/4, 3/4) - C(1/4, 3/4) - C(3/4, 1/4) = 1/4 + 3/4 - 1/4 - 1/4 = 1/2 \), while the right-hand side evaluates to \( C(x_2 - C(x_2, y_1), y_2 - C(x_1, y_2)) = C(1/2, 1/2) = 1/3 \), showing that Eq. (30) does not hold for all \( 0 \leq x_1 \leq x_2 \leq 1 \) and \( 0 \leq y_1 \leq y_2 \leq 1 \). Note that for the same choice of points also Eq. (31) fails to hold.

From this example we also learn that ‘symmetrical’ ordinal sums of Frank copulas do not always fulfil the condition of Theorem 1. For the Frank copulas themselves, however, condition Eq. (30) is always satisfied.
Proposition 8. For any Frank copula and any \(0 \leq x_1 \leq x_2 \leq 1\) and \(0 \leq y_1 \leq y_2 \leq 1\), it holds that
\[
C(x_1, y_1) + C(x_2, y_2) - C(x_1, y_2) - C(x_2, y_1) \leq C(x_2 - C(x_2, y_1), y_2 - C(x_1, y_2)).
\]

Proof. The proof goes by straightforward computation. Let \(C = T^F_\lambda\) and assume that \(\lambda > 1\). Condition Eq. (30), rewritten in the form
\[
C(u, s) + C(v, t) - C(u, t) - C(v, s) \leq C(v - C(v, s), t - C(u, t)),
\]
with \(u \leq v\) and \(s \leq t\), is equivalent with
\[
[\lambda - 1 + (\lambda^x - 1)(\lambda^y - 1)]\frac{[\lambda - 1 + (\lambda^x - 1)(\lambda^t - 1)][\lambda - 1] - 1}{[\lambda - 1 + (\lambda^x - 1)(\lambda^t - 1)][\lambda - 1 + (\lambda^y - 1)(\lambda^s - 1)]}.
\]
Since
\[
\frac{\lambda^x - C(x, y)}{\lambda^x - 1} = \frac{\lambda^x(\lambda - 1)}{\lambda - 1 + (\lambda^x - 1)(\lambda^y - 1)} - 1 = \frac{(\lambda^x - 1)(\lambda - \lambda^y)}{\lambda - 1 + (\lambda^x - 1)(\lambda^y - 1)},
\]
the above inequality is also equivalent with
\[
(\lambda - 1)^2(\lambda^y - \lambda^x)(\lambda^z - \lambda^w) \leq (\lambda^y - 1)(\lambda - \lambda^w)(\lambda^z - 1)(\lambda - \lambda^z).
\]
This inequality holds for all \(u, v, s, t \in [01]\) since
\[
(\lambda - 1)(\lambda^y - \lambda^x) \leq (\lambda^y - 1)(\lambda - \lambda^x),
\]
is equivalent to \((\lambda - \lambda^y)(\lambda^x - 1) \geq 0\), and similarly,
\[
(\lambda - 1)(\lambda^z - \lambda^y) \leq (\lambda^z - 1)(\lambda - \lambda^y),
\]
is equivalent to \((\lambda - \lambda^z)(\lambda^y - 1) \geq 0\). For \(\lambda < 1\), the computations are similar. For \(\lambda = 1\), the proof is immediate. \(\square\)

Based on numerous numerical experiments, we dare to put forward the following conjecture.

Conjecture 1. For any Frank copula condition (22) is satisfied.

Assuming the conjecture is true, Theorem 1 applies in particular to the Frank copula family. For \(C = T^F_\lambda\), it then holds that the probabilistic relation \(Q\) generated by a collection of r.v. is cycle-transitive w.r.t. the upper bound function \(U^F_\lambda\) given by:
\[
U^F_\lambda(x, \beta, \gamma) = \beta + T^F_\lambda(1 - \beta, \gamma) = \beta + \gamma - T^F_\lambda(\beta, \gamma) = S^F_{1/\lambda}(\beta, \gamma).
\]
In the above transition, we have used the fact that \(T^F_\lambda(1 - x, y) = y - T^F_{1/\lambda}(x, y)\) \([14]\).

Since for \(\lambda < \lambda^r\) it holds that \(T^F_\lambda \geq T^F_{\lambda^r}\), it also follows that \(U^F_\lambda \geq U^F_{\lambda^r}\). Therefore, the lower the value of \(\lambda\) when the r.v. are coupled by \(T^F_\lambda\), the weaker the type of transitivity exhibited by the probabilistic relation generated by these r.v. In particular, the strongest type of transitivity is encountered when coupling by \(T_L\), the weakest when coupling by \(T_M\).

Let us discuss the three main copulas again.

(i) For \(C = T_L = T^F_\infty\), it follows from Eq. (37) that \(Q\) is cycle-transitive w.r.t. the upper bound function \(U^F_\infty\) given by
\[
U^F_\infty(x, \beta, \gamma) = \max(\beta, \gamma) = \gamma = U_{ps}(x, \beta, \gamma).
\]
In other words, \(Q\) is partially stochastic transitive (Proposition 7).

(ii) For \(C = T_P = T^F_1\), we retrieve the well-known case of independent r.v. (Proposition 5), with
\[
U^F_1(x, \beta, \gamma) = \beta + \gamma - \beta\gamma = U_D(x, \beta, \gamma).
\]

(iii) For \(C = T_M = T^F_0\), it follows from Eq. (37) that \(Q\) is cycle-transitive w.r.t. the upper bound function \(U^F_0\) given by
In other words, $Q$ is $T_L$-transitive (Proposition 6).

6. Conclusion

In this paper, we have studied the probabilistic relation generated by a collection of random variables, obtained by computing winning probabilities from artificially constructed bivariate c.d.f. Provided the copula involved in this construction satisfies some additional condition, this probabilistic relation turns out to exhibit an interesting type of transitivity that can be expressed in the cycle-transitivity framework. The generated probabilistic relation can therefore be seen as a graded alternative to stochastic dominance. In future work, we will try to characterize the (stable) commutative copulas fulfilling the additional condition encountered, and at least confirm that Frank copulas effectively fulfill that condition. Additionally, we will examine how the knowledge of this type of transitivity can help us come up with meaningful crisp partial orders on a collection of random variables, providing alternatives to the classical stochastic dominance notions.

References