Probabilistic constructions of discrete copulas
Olivier P. Faugeras

To cite this version:
Olivier P. Faugeras. Probabilistic constructions of discrete copulas. 2012. <hal-00751393>

HAL Id: hal-00751393
https://hal.archives-ouvertes.fr/hal-00751393
Submitted on 13 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Probabilistic constructions of discrete copulas

Olivier P. Faugeras

Toulouse School of Economics - Université Toulouse 1 Capitole - GREMAQ, Manufacture des Tabacs, Bureau MF319, 21 Allée de Brienne, 31000 Toulouse, France.

Abstract

For a multivariate vector $X$ with discrete components, we construct, by means of explicit randomised transformations of $X$, multivariate couplings of copula representers $U$ associated to $X$. As a result, we show that any copula can be constructed in this manner, giving a full probabilistic characterisation of the set of copula functions associated to $X$. The dependence properties of these copula representers are contrasted with those of the original $X$: the impact on independence and on the concordance order of this added randomisation structure is elucidated and quantified. Some explicit formulas are given in the bivariate and pattern recognition case. At last, an application of these constructions to the empirical copula shows they can easily be simulated when the distribution function is unknown.

Keywords: Copula, Multivariate coupling, Probabilistic construction, Distributional transform, Dependence, Concordance, Kendall’s $\tau$, Empirical copula

1. Introduction

The paper is organised as follows: in the present section 1, we motivate our study of copula functions by distinguishing between the analytical and probabilistic approaches. In section 2, we give a first probabilistic construction which is generalised in section 3. The dependence properties of the former construction is studied in section 4, while some particular cases and applications are presented in section 5.
1.1. Analytical versus probabilistic characterisation of measures

Let \( \mu \) be a probability measure on the Euclidean measurable space \( (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \).

It is classical material in probability theory to show that one can restrict the study of the measure \( \mu \) to some subsets of the sigma algebra, such as the set of lower orthants \( \{(-\infty, x]\} = \{(-\infty, x_1] \times \ldots \times (-\infty, x_d]\} \), hence defining the cumulative distribution function (cdf) \( F : \mathbb{R}^d \mapsto [0, 1] \) as

\[
F(x) = \mu\{(-\infty, x] \}
\]

This induces some properties on \( F \):

P.i) \( F \) is continuous from above;

P.ii) \( F \) is \( d \)-increasing, viz. \( \Delta_A F \geq 0 \) for bounded rectangles \( A \);

P.iii) \( F(+\infty, \ldots , +\infty) = 1 \) and \( \lim_{x \downarrow y} F(x) = 0 \) if at least one coordinate of \( y \) is \(-\infty\).

(see e.g. Billingsley [2] chapter 12, Resnick [12] chapter 2-3 or Shiryaev [16] p. 160). These properties thus restrict the set of functions which can qualify as cdf and can conversely be used as an analytical definition of a cdf, as in e.g. Shiryaev [16] definition 2 p. 160, or Nelsen [11] section 2.4. Indeed, to any function on \( \mathbb{R}^d \) satisfying the above properties P.i)-iii), there exists a unique probability measure on \( \mathbb{R}^d \) corresponding to that cumulative distribution function, see e.g. Billingsley [2] theorem 12.4 for the univariate case and theorem 12.5 for the multivariate case, or Shiryaev [2] theorem 2 p. 160.

In addition to such an analytical characterisation of measures, one can alternatively give a probabilistic characterisation. More precisely, starting from a cdf \( F \) (in the sense of a function satisfying properties P.i)-P.iii)), one can construct on a given probability space \( (\Omega, \mathcal{A}, P) \) a vector-valued random variable \( X \) whose cdf is precisely \( F \). An explicit construction in the univariate case, where the probability space is chosen to be \( ([0, 1], \mathcal{B}([0, 1]), \text{Leb}) \), is given in the well-known Skorohod representation or quantile transform theorem, see e.g. Williams [21] section 3.12, Resnick [12] section 2.4 and 2.5, Billingsley [2] theorem 14.1. It relies on explicitly setting \( X = Q(U) \), where \( Q \) is the generalised inverse of \( F \) or quantile function, viz.

\[
Q(t) = \inf\{x : F(x) \geq t\}, \quad 0 < t < 1;
\]

with \( Q(0) = \lim_{t \downarrow 0} Q(t) \), \( Q(1) = \lim_{t \uparrow 1} Q(t) \), and \( U \) is uniform on \([0, 1]\).
Although the same holds for a multivariate cdf as is shown in, e.g. Shiryaev [16] theorem 2.3.3 p. 160 or Billingsley [2] theorem 12.5, the proof is generally based on some more advanced result of measure theory such as Caratheodory’s extension theorem and is generally not as explicit as in the one dimensional case. It is thus desirable to be able to produce an explicit construction in the multivariate case, which generalises to $([0,1]^d, B([0,1]^d)$ the previous univariate quantile transform, as stated in Deheuvels [3] proposition 1.2, who claims that one can define a collection $U = (U_1, \ldots, U_d)$ of uniform on $(0,1)$ random variables, together with a multivariate $X = (X_1, \ldots, X_d)$, on the same probability space $(\Omega, \mathcal{A}, P)$, in such a way that

i) $X$ has c.d.f. $F$, $P(X \leq x) = F(x)$

ii) $X$ is explicitly defined as $X = Q(U) = (Q_1(U_1), \ldots, Q_d(U_d))$, where $Q_i$ is the $i$th marginal quantile function.

Such a probabilistic representation of measures via random variables is related to the more general concept of coupling, see e.g. Thorisson [19]: a representation or copy of a random variable $X$ is a random variable $\hat{X}$ with the same distribution as $X$. A coupling of a collection $\{X_1, \ldots, X_d\}$ of random variables (possibly defined on different probability spaces) is a joint construction of random variables $(\hat{X}_1, \ldots, \hat{X}_d)$ defined on the same probability space, and such that $\hat{X}_i \overset{d}{=} X_i$, $i = 1, \ldots, d$.

1.2. On the usefulness of both approaches

At first sight, the above discussion, on whether one should characterise multivariate Euclidean probability measures analytically or probabilistically, may appear philosophical and a matter of personal taste.

However, Nelsen [11] p. 243 argue it is not the case as some operations on distribution functions are not derivable from operations on the corresponding random variables. More precisely, if $X \sim F$, $Y \sim G$ and $\Psi(F, G)$ is a binary operation on the two cdf $F, G$, then, $\Psi$ is said to be derivable if there exists a Borel-measurable function $Z(\cdot, \cdot)$ such that $Z(X, Y)$ has cdf $\Psi(F, G)$. Nelsen cite the mixture operation,

$$\Psi(F, G) = pF + (1 - p)G, \text{ for } 0 < p < 1,$$

as an example of an operation which is not derivable. This lead Schweizer and Sklar [15] to conclude that “the distinction between working directly with distributions functions [...] and working with them indirectly, via random
variables, is intrinsic and not just a matter of taste” and Alsina et al. [1] that “the classical model for probability theory– which is based on random variables defined on a common probability space– has its limitations”.

On the other hand, Thorisson [19] p. 5 argue in favour of the probabilistic approach, and that “a coupling characterisation of a distributional property deepens our understanding of that property”. He substantiates his claim with an example on stochastic domination, where simple proofs can be obtained once a coupling has been constructed. Indeed, by being constructive, a coupling characterization makes emergence of properties more explicit.

Our aim in this article is to substantiate Thorisson’s view, by constructing several explicit representations of $X$ on $([0,1]^d, \mathcal{B}([0,1]^d))$. We will see in the next subsection 1.3 that such joint constructions are intrinsically connected with the concept of copula functions, and in subsection 1.4 that such epistemological distinctions are not superfluous, except when all coordinate of $X$ are continuous.

In order for a probabilistic construction of copula functions to be made possible, we first need to alleviate the preliminary objection raised by Nelsen. Indeed, it can be noted that the mixture operation in his example can be recovered from operations on r.v. by allowing randomized operations. Indeed, if $I \sim B(1,p)$ is a Bernoulli r.v. independent of $(X,Y)$, then $Z = IX + (1-I)Y$ has cdf $\Psi(F,G)$. Such a representation of a mixture is often called a latent variable representation of $\Psi$ and is easily generalised to a mixture of more than two components. Allowing an extra randomization is fundamental and will be seen to be the key element to give a probabilistic characterisation of copula functions, see section 2 below.

1.3. Copulas

The same epistemological distinction occurs for defining copula functions. A $d$-dimensional copula function $C$ can be defined,

- either analytically, as a function defined on $[0,1]^d$, grounded and $d$-increasing with univariate marginals, see Nelsen [11] chapter 2,

- or probabilistically as the restriction to $I^d$ of the multivariate cdf of a random vector $U$ with each marginal uniformly distributed on $I = [0,1]$, see below.

Their main interest lies in the celebrated Sklar’s theorem:
Theorem 1.1 (Sklar-analytic). For every Multivariate cdf $F$, with marginal cdf $G = (G_1, \ldots, G_d)$, there exists some copula function $C$ such that the fundamental identity holds

$$F(x) = C(G(x))$$ (2)

where $G(x) = (G_1(x_1), \ldots, G_d(x_d))$. Conversely, if $C$ is a copula and $G$ a set of marginal distribution functions, then the function $F$ defined by (2) is a joint distribution function with marginals $G$.

The latter theorem thus defines a (set of) copula associated with $F$ (or $X$). Its proof is illustrative of the analytical approach to copulas: the set of pairs $\{(G(x), F(x))\}$ which satisfies equation (2) defines a unique $d$–variate function $C'$ on $\text{Ran}G$, see [11] lemma 2.3.4. The latter function is a sub-copula, viz. a function satisfying the same analytical properties as of a copula, except it is defined only on the subset $S_1 \times \ldots \times S_d \subset I^d$, where the $S_i \subset I$ and contain 0 and 1. The latter function is then extended to $I^d$, first by continuity on $\{S_1 \times \ldots \times S_d\}$, then to $I^d$ by a bilinear interpolation preserving the analytical properties of a copula, see Nelsen [11] lemma 2.3.5.

Deheuvels [3] takes the probabilistic side in the epistemological divide, claiming that relation (2) is “often mistakenly used as a definition of a copula” associated to $X$ (or $F$). He prefers to define a copula associated with $X$ through a multivariate quantile coupling. We can reformulate his proposition 1.2 in a theorem like as follows:

Theorem 1.2 (Sklar-probabilistic). We may define (a representation of) $X$ on a probability space jointly with a vector $U$ such that

i) each marginal of $U$ is uniform $(0, 1)$;

ii) $X = Q(U)$.

The cdf of a vector $U$ satisfying the previous conditions is a copula function $C$ associated with $X$. Moreover, Sklar’s identity (2) is necessary and sufficient for $C$ to be a copula associated with $X$.

Note that Deheuvels does not distinguish between the original $X$ and its representation, as they are equal a.s., nor gives an explicit construction of $U$ when at least one of the marginals $G_j$ is discontinuous.
1.4. Precopulas

To substantiate the difference between the analytical and probabilistic views, it is pertinent to define the concept of precopulas of the first and second type introduced by Deheuvels in [3], as, respectively,

i) \( C^*(t) = F(Q(t)) \);

ii) \( C^{**}(t) = P(G(X) \leq t) \).

When the marginals \( G \) are all continuous, this analytical-probabilistic distinction is superfluous, \( C \) is unique and can be explicitly defined either way, as

\[ C(t) = C^*(t) = C^{**}(t). \]

To the contrary, when at least one of the marginals \( G_j \) is discontinuous, the two approaches do not allow to explicitly define the (whole set of) copulas associated with \( X \). Analytically, the problem is twofold:

i) for a discrete marginal, several choice of generalised inverse, thus of quantile functions, are possible, in particular the left-continuous version \( Q(t) \) of \( 1 \) and its right continuous one \( Q_+(t) = \lim_{\varepsilon \to 0} Q(t + \varepsilon) \), for \( t \in (0,1) \), see Deheuvels [3]. So once an arbitrary version of the quantile functions has been chosen, one can not recover the whole set of possible copulas;

ii) equation (2) only uniquely defines a function on the range of the marginals. So “inverting” such a relation as with the precopula of the first type approach does not lead to a function satisfying the analytical properties of a copula (in particular the uniformity of marginals).

Probabilistically, similar problems occur:

i) Several representations of the random variables may exist and remain indistinguishable, as noted in e.g. [21] p. 34. with the example of \( Q(U) \equiv Q_+(U) \);

ii) each transformed marginal \( G_j(X_j) \) remains discontinuous and is no longer uniform, so the cdf of \( G(X) \) no longer has uniform \((0,1)\) marginals, and \( C^{**} \) is thus not a copula.

This limitation is exemplified in Deheuvels’s [3] proposition 1.3 and 1.4. which show that the precopulas of the first and second type, although they remain uniquely defined for each specified \( F \) (or \( X \)), are closely connected with any copula \( C \) associated to \( X \), in the sense that \( C^* \) and \( C^{**} \) agree with
$C$ on the set of points of left-increase of $G$ (i.e. on the support of $X$ for a discrete $X$) and satisfy $C^{**}(u) \leq C(u) \leq C^*(u)$, $u \in [0,1]^d$.

Genest and Neslehova [8] illustrate the above difficulties by introducing different set of functions: let $A$ the set of sub-copulas which satisfies (2), $\mathcal{H}$ the set of cdf, and $\mathcal{C}$ the set of copulas. For the analytical approach, they define $B$ (respectively $C$) as the precopula of the first type obtained by using the left- (respectively right-) continuous quantile functions, viz. $B = F \circ Q^{-1}$ and $C = F \circ Q^{-1}$. For the probabilistic approach, they define $D$ (respectively $E$) the joint cdf of the vector obtained by using the right- (respectively left-) continuous version of probability integral transforms, viz. $D(u) = P(G(X) \leq u)$, (respectively, $E(u) = P(G_-(X) \leq u)$). Their proposition 1 show that for precopulas of the first type, $B \in A$, but $A \notin \mathcal{H}$ whereas $C \in \mathcal{H}$ but $C \notin A \cap \mathcal{C}$, and for precopulas of the second type, $D \in A \cap \mathcal{H}$ but $D \notin \mathcal{C}$ whereas $E = C$.

2. A first probabilistic construction for discrete vectors: $R$-copulas

Let $X = (X_1, \ldots, X_d)$ a discrete $d$-dimensional vector defined on a given probability space $(\Omega, \mathcal{A}, P)$, with cumulative distribution $F$ and marginal cdfs $G = (G_1, \ldots, G_d)$. Our goal is to give a probabilistic construction of the set of copulas associated with the discrete $X$, i.e. to define the set of discrete copulas as precopulas of the second type. The discussion of the preceding section can be illustrated in the following diagram,

$$
\begin{array}{c}
F \xrightarrow{q} X \\
c \downarrow \quad \quad r \downarrow \\
C \xleftarrow{p} U
\end{array}
$$

where $q$ stands for the symbolic mapping corresponding to the quantile coupling operation, that is to say which constructs a random vector from a given cdf, and $p$ its inverse, i.e. which defines a cdf from a random vector. The analytic approach can be described via the map $c$, whereas a probabilistic approach follows $p \circ r$, where $r$ is the mapping which transforms a random $X$ to a representation $U$ whose distribution is the copula $C$ associated to $X$. The latter approach amounts to give an explicit probabilistic construction of the set of possible $U$ in theorem 1.2, i.e. to explicitly define the map $r$.

In the continuous case, $r$ is simply the probability integral transform, viz. $r(X) = G(X)$. To follow a probabilistic approach in the discontinuous case,
we need to extend the probability integral transform to non-continuous cdf. The discussion of section 1.1 suggests that we have to define a randomized transformation, which we present below.

2.1. The \( \mathcal{R} \)-distributional transform

In the univariate case, Rüschendorff [13] introduces a random probability integral transform, which uniformly randomizes the jumps of the cdf, called the distributional transform.

Definition 2.1. Let \( X \sim F \) univariate. The modified c.d.f. is defined as

\[
F(x, \lambda) := P(X < x) + \lambda P(X = x) = F_-(x) + \lambda (F(x) - F_-(x))
\]

Let \( V \overset{d}{=} U(0, 1) \) with \( V \perp X \), then the distributional transform \( \mathcal{R} : X \mapsto U \) of \( X \) is defined by

\[
U := \mathcal{R}(X) = F(X, V) = F_-(X) + V(F(X) - F_-(X))
\]

This transform reduces to the classical probability integral transform when \( X \) is continuous. Its interest lies in the following theorem, proved in Rüschendorff [13], proposition 2.1.

Theorem 2.2. One has,

i) \( U \overset{d}{=} U(0, 1) \)

ii) \( X = F^{-1}(U) \) a.s.

In other words, the quantile transform is exactly the inverse of the distributional transform.

2.2. The \( \mathcal{R} \)-multivariate distributional transform quantile coupling

The distributional transform allows to define a coupling of \( X \), which we call the multivariate distributional transform quantile coupling, similarly to proposition 1.2 of Deheuvels [3]:

Theorem 2.3. We can extend the probability space to carry a vector of marginally uniform \((0, 1)\) random variables \( \mathbf{V} = (V_1, \ldots, V_d) \), independent of \( \mathbf{X} \), so that we can construct (a representation of) \( \mathbf{X} \), together with a random vector \( \mathbf{U} = (U_1, \ldots, U_d) \in (0, 1)^d \), with the following properties

i) \( \mathbf{U} = G(\mathbf{X}, \mathbf{V}) = (G_1(X_1, V_1), \ldots, G_d(X_d, V_d)) \);
ii) $X = Q(U)$ a.s.;

iii) $U$ has uniform $(0, 1)$ marginals.

Proof. $X$ is defined on $(Ω, A, P)$ with value in $(\mathbb{R}^d, B(\mathbb{R}^d))$. On $(0, 1)^d$, define, for $i = 1, \ldots, d$, $π_i$ as the $i$th coordinate projection mapping $π_i(v) = v_i$ for $v = (v_1, \ldots, v_d) \in (0, 1)^d$. Let $µ$ a probability measure on $([0, 1]^d, B([0, 1]^d))$ with uniform marginals, viz. $µ ◦ π_i^{-1}(A) = Leb(A)$, for all $A ∈ B([0, 1])$.

Define the extension $(\tilde{Ω}, \tilde{A}, \tilde{P})$ of the probability space $(Ω, A, P)$ by setting

i) $\tilde{Ω} = \omega × [0, 1]^d$;

ii) $B(\mathbb{R}^d) ⊗ B([0, 1]^d)$ the product sigma field;

iii) $\tilde{P} = P × µ$.

Define $ζ_1(\tilde{ω}) = ω$ and $ζ_2(\tilde{ω}) = v$ for $\tilde{ω} = (ω, v) ∈ \tilde{Ω}$. Set $\tilde{X}(\tilde{ω}) = X(ζ_1(\tilde{ω}))$ and $\tilde{V}(\tilde{ω}) = V(ζ_2(\tilde{ω}))$. Then $\tilde{P}(\tilde{X} ∈ A) = P(X ∈ A)$ for $A ∈ B(\mathbb{R}^d)$ and $\tilde{P}(\tilde{V} ∈ B) = µ(B)$ for $B ∈ B([0, 1]^d)$. Moreover,

$$\tilde{P}(\tilde{X} ∈ A, \tilde{V} ∈ B) = P(X ∈ A) × µ(B)$$

$$= P(X ∈ A) × P(V ∈ B)$$

In other words, we have constructed, on a common probability space, random variables, $\tilde{X}, \tilde{V}$ which have the same law as, respectively, $X, V$ and which are independent of each other.

From now on, we will make use of the convention of the common probability space, by identifying $\tilde{X}, \tilde{V}$ and $X, V$. Set $U = G(X, V)$ and $\tilde{X} = Q(U)$. By theorem 2.2, each marginal $U_i$ of $U$ is uniform and $X_i = \tilde{X}_i$ a.s. Moreover, $U$ is in the random cube $[G_-(X), G(X)]$ w.p. 1. For any $u$ in the cube $[G_-(x), G(x)]$, $G_-(u) = x$ jointly, thus $G_-(U) = X$ a.s.

Let’s call the vector $V$ a randomizer. As it is a vector with uniform marginals, its distribution $µ$ is described by a copula function $R$. Denote by $R^e$ the corresponding cdf, viz. the extension of $R$ to $\mathbb{R}^d$. (See Nelsen [11] p. 24 for a discussion between copula functions and random variables, via the extension of the former to $\mathbb{R}^d$). The choice of $V$ allows to explicitly define the copula associated with (that representation of) $X$ in a probabilistic manner, viz. as a precopula of the second type:

Corollary 2.4. Let $C$ the cdf corresponding to $U$ in the above construction, $C(u) = P(U ≤ u)$. Then, $C$ is a copula function associated to $X$. 

9
Proof.

\[ F(x) = P(X \leq x) = P(Q(U) \leq x) = P(U \leq G(x)) = C(G(x)). \]

In other words, \( U \) is a representation of \( C \), and the set of copulas associated with \( X \) in this construction is parametrized by the choice of the randomizer \( V \). We have thus defined our map \( p \circ r \) in diagram (3). More precisely, we have constructed the map \( (X, V) \mapsto C \), which defines \( C \) in a unique way. We call the set of copulas obtained in this manner, the \( R \)-copulas of \( X \).

**Remark 1.** The above multivariate distributional transform quantile coupling has not to be confused with what Rüschendorff calls the multivariate distributional transform, which, as generalisation of Rosenblatt’s transform, turn a \( d \)-dimensional vector into a \( U[(0,1)^d] \) one, and as such, “forget all dependence” in the original vector.

**Remark 2.** Rüschendorf [13] proposed a simplified version of the above construction with a comonotone vector \( V = (V, \ldots, V) \).

2.3. A general decomposition formula of \( R \)-copulas

The above construction yields an analytical formula for the set of admissible \( R \)-copulas as follows.

**Theorem 2.5.** Let \( R^e \) be the cdf of \( V \) in the above construction. Then,

\[ C(u) = \sum_x R^e \left( \frac{u - G_-(x)}{G(x) - G_-(x)} \right) P(X = x) \] (4)

where the sum is over all values of \( x \).

**Proof.**

\[
C(u) = P(U \leq u) = P(F(X,V) \leq u) \\
= E[P(F(X,V) \leq u|X)] = \int P(F(x,V) \leq u) dF(x) \\
= \sum_x P \left( V \leq \frac{u - G_-(x)}{G(x) - G_-(x)} \right) P(X = x)
\]

where the second line follows from independence and the third by the fact that \( X \) is discrete. \( \square \)
Remark 3. Note that, as noted above, given that $X = x$, $U = F(x, V)$ is in the cube $(G_-(x), G(x))$ w.p. 1. Therefore, for each marginal $i$,

$$P\left(V_i \leq \frac{u_i - G_{i-}(x_i)}{G_i(x_i) - G_{i-}(x_i)}\right) = \left\{\begin{array}{ll} 0, & \text{if } u_i \leq G_{i-}(x_i) \\ 1, & \text{if } u_i \geq G_i(x_i). \end{array}\right.$$  

In other words, the previous formula can be specialised for special cases, which will be done in section 5.1.

3. Two supplementary probabilistic constructions: $S$- and $S'$-copulas

3.1. Shorack’s transform

Rüschendorff’s transform in theorem 2.2 use the same randomization for each jump of the random variable. It is possible to generalise this transformation by allowing different randomizers for each jump, as suggested in Shorack [17].

Let $X \sim F$ a discrete univariate random variable. There is only a countable number of jumps \( \{x_1, x_2, \ldots\} \) and the cdf writes $F(x) = \sum_i f(x_i)1_{[x_i, \infty)}$, with $f(x_i) = F(x_i) - F(x_i-)$ the jump size at $x_i$. Set $x_0 = -\infty$. If $X$ has a finite number, say $k$, of values, set $x_{k+1} = \infty$, otherwise set $x_\infty = \infty$.

**Definition 3.1.** Let $(U_i)$ a sequence of uniform $[0, 1]$ random variables, independent of $X$. Shorack’s distributional transform is the mapping $S : X \mapsto U$, defined as

$$U := S(X) := F(X) - \sum_{i \geq 1} f(x_i)U_i1_{X=x_i}$$

Similarly to Rüschendorff’s transform, this randomized distributional transform also turns a r.v. $X$ into a uniform one, as proved in Shorack’s proposition 7.3.2, which we reproduces here for reference:

**Theorem 3.2.** With the notations of definition 3.2, one has,

\begin{itemize}
  \item[i)] $U \overset{d}{=} U_{(0,1)}$
  \item[ii)] $X = F^{-1}(U)$ a.s.
\end{itemize}
Proof.

\[ P(S(X) \leq u) = E[P(S(X) \leq u|X)] = \int P\left(F(x) - \sum_{i \geq 1} f(x_i)u_{1_{x=x_i}} \leq u\right) dF(x) = \sum_{i \geq 1} P(F(x_i) - f(x_i)u \leq u)P(X = x_i), \]

where we used \( X \perp (U_i) \) in the first line. Fix \( u \in [0,1] \). There exists a corresponding integer \( k \in \{1,2,\ldots,\infty\} \), such that \( u \in [F_-(x_{k-1}), F(x_k)] \). For each \( i \), \( F(x_i) - f(x_i)U_i \) is uniform on \((F_-(x_i), F(x_i)) = (F(x_{i-1}), F(x_i)) \). Therefore,

\[
P\left(U_i \geq \frac{F(x_i) - u}{f(x_i)}\right) = \begin{cases} 
0, & \text{if } i \geq k + 1 \\
1, & \text{if } i \leq k - 1 \\
1 - \frac{F(x_i) - u}{f(x_i)}, & \text{if } i = k 
\end{cases}
\]

thus

\[
P(S(X) \leq u) = F(x_{k-1}) + f(x_k)\left(1 - \frac{F(x_k) - u}{f(x_k)}\right) = u,
\]

with the obvious extension for \( u = 1 \), which proves i). Conditionally on \( X = x_i \), \( F_-(x_i) \leq S(x_i) \leq F(x_i) \) with probability one and for any \( u \in (F_-(x_i), F(x_i)) \), \( F^{-1}(u) = x_i \), thus \( P(F^{-1}(U) = X|X = x_i) = 1 \). In turn,

\[
P(F^{-1}(U) = X) = \sum_{i \geq 1} P(F^{-1}(U) = X|X = x_i)P(X = x_i) = 1,
\]

which is ii).

\[ \square \]

Remark 4. Shorack [17] defines his transform with the extra assumption that the set of \((U_i)\) uniform r.v. be independent of each other, which is not required for theorem 3.2 to hold true.

3.2. Relation with R"uschendorf’s transform and the modified Shorack’s transform

In the one dimensional case, Shorack’s transform writes, conditionally on \( X = x_i \),

\[
S(X = x_i) = F(x_i) - U_i f(x_i) = F_-(x_i) + V_i (F(x_i) - F_-(x_i)),
\]

12
with $V_i = 1 - U_i$. In other words, Shorack’s transform is a conditional Rüschendorff’s transform with a reversed randomizer $V_i = 1 - U_i$. Moreover, whenever the randomizing variables are the same, viz. $U_i = U$ for all $i$, Shorack’s transform reduces to Rüschendorf’s one with $V = 1 - U$.

In order to reconcile with Rüschendorff’s transform, we define the following modified Shorack’s (univariate) transform $S'$ as a conditional Rüschendorf’s transform as follows:

**Definition 3.3.** For $X$ univariate, define, conditionally on $X = x_i$, the univariate r.v. 

$$S'(X = x_i) = F_-(x_i) + U_if(x_i) = F(x_i, U_i),$$

where $U_i \sim U_{[0,1]}$. The modified Shorack univariate transform is set as

$$S'(X) = \sum_{x_i} 1_{X=x_i} S'(X = x_i).$$

Then, theorem 3.2 remains obviously valid with this modified transform $U := S'(X)$ instead of $U := S(X)$.

3.3. **$S$-copulas for bivariate discrete random variables**

As with Rüschendorff’s transform, Shorack’s transform or its modification allows to obtain a probabilistic proof of Sklar’s theorem, via a multivariate coupling.

For reasons to become clear later, we restrict ourselves to the bidimensional case for Shorack’s transform. Let $(X, Y)$ be a discrete bivariate vector with cdf $H$ and marginals $F$ and $G$: $(X, Y)$ has a countable number $k = (k_X, k_Y) \in \overline{\mathbb{N}} \times \overline{\mathbb{N}}$ of values with strictly positive mass. Denote by \{ $x_i, i = 1, \ldots, k_X$ \} (resp. \{ $y_j, j = 1, \ldots, k_Y$ \}), these jump points of $X$ (resp. $Y$), with jump size $f(x_i) = F(x_i) - F(x_i -)$ (resp. $g(y_j) = G(y_j) - G(y_j -)$). Define a **randomisation structure** $S$ as a set of bivariate uniform random variables $(U_i, V_j)$, with copula $C_{ij}$ for $i = 1, \ldots, k_X, j = 1, \ldots, k_Y$. Construct the pair of Shorack’s tranform,

$$S(X) = F(X) - \sum_i f(x_i)U_i1_{X=x_i}$$

$$S(Y) = G(Y) - \sum_j g(y_j)V_j1_{Y=y_j}$$

Then, if $\tilde{C}$ is the cdf of $(S(X), S(Y))$, then $\tilde{C}$ is a copula associated to $X$: 13
Theorem 3.4. We can extend the probability space to carry the randomisation structure \( \mathcal{S} \), with \((U_i, V_j) \sim C_{ij}\) and \( \mathcal{S} \) independent of \((X, Y)\), so that we can construct (a representation of) \((X, Y)\), together with a pair \((\tilde{U}, \tilde{V})\), with the following properties

i) \(\tilde{U} = \mathcal{S}(X), \tilde{V} = \mathcal{S}(Y)\) are uniform \([0, 1]\);

ii) \(X = F^{-1}(\tilde{U}), Y = G^{-1}(\tilde{V})\);

iii) The cdf \(\tilde{C}\) of \((\tilde{U}, \tilde{V})\) satisfies Sklar’s theorem.

Proof. Similarly to theorem 2.3, define \(\pi_{ij}\) the coordinate projection mapping \(\pi_{ij}(u, v) = (u_i, v_j)\) for \((u, v) \in [0, 1]^{k_Y} \times [0, 1]^{k_X}\). Let \(\mu_{ij}\) be the probability measure on \(([0, 1]^2, \mathcal{B}([0, 1]^2))\) corresponding to the copula functions \(C_{ij}\). Let \(\mu\) be a probability measure on \(([0, 1]^{k_Y} \times [0, 1]^{k_X}, \mathcal{B}([0, 1]^{k_Y} \times [0, 1]^{k_X}))\) satisfying \(\mu \circ \pi_{ij}^{-1} = \mu_{ij}, \mu \circ \pi_i^{-1} = \mu \circ \pi_j^{-1} = \text{Leb}\). The existence of \(\mu\) is guaranteed by the Daniell-Kolmogorov extension theorem, as the compatibility conditions \(\mu_{ij} \circ \pi_i^{-1} = \mu \circ \pi_i^{-1} = \text{Leb}\) are automatically satisfied in this bivariate case. One can thus define jointly \((X, Y)\) and a countable set \(((U_i, V_j), i \in \{1, \ldots, k_X\}, j \in \{1, \ldots, k_Y\})\) of bivariate vectors on a common probability space, with \((U_i, V_j) \sim \mu_{ij}\), such that \((X, Y)\) is independent of \(((U_i, V_j), i \in \{1, \ldots, k_X\}, j \in \{1, \ldots, k_Y\})\). The rest of the proof is similar to that of theorem 2.3 with theorem 3.2 instead of 2.2 and is thus omitted. \(\square\)

Similarly to section 2, we call the set of copulas which can be obtained via this construction, the set of \(\mathcal{S}\)-copulas associated to \((X, Y)\).

Remark 5. Note that by choosing a randomization structure reduced to a singleton, i.e. if \(C_{ij} = R\) for all \(i, j\), one recovers the previous \(R\)-copula construction. Therefore, the \(\mathcal{S}\)-copula construction contains the \(R\)-copula one. We chose to separate this two constructions, as the \(R\)-copula construction has the advantage of simplicity and may prove sufficient for some applications, see section 4.

3.4. Generalisation to higher dimensions by the modified Shorack’s transform: conditional \(\mathcal{S}'\) copulas

When the dimension of \(X\) is higher than 2, defining Shorack’s transform as above with a multidimensional array of randomizers independent of \(X\) may lead to problems of existence of such an array of random variables.
For example, in the 3-dimensional case, if \( X = (X, Y, Z) \) has cdf \( F \) and marginals \( G_1, G_2, G_3 \), and jump points \( \{x_i, y_j, z_k\} \), setting

\[
\tilde{U}_X = G_1(X) - \sum_i (G_1(x_i) - G_1(x_i^-)) U_i 1_{X=x_i}
\]

\[
\tilde{U}_Y = G_2(Y) - \sum_j (G_2(y_j) - G_2(y_j^-)) V_j 1_{Y=y_j}
\]

\[
\tilde{U}_Z = G_3(Z) - \sum_k (G_3(z_k) - G_3(z_k^-)) W_k 1_{Z=z_k}
\]

with \((U_i, V_j, W_k) \sim C_{ijk}\) some trivariate copula functions, imposes some compatibility conditions on the randomization structure \( S = \{C_{ijk}\} \), in particular that \( \mu_{ijk} \circ \pi_i = \mu_{ijk'} \circ \pi_i \), for \( k \neq k' \) with obvious notations. The set of randomizing copulas \( C_{ijk} \) thus can not be chosen freely. For more details on existence of multivariate distributions given some finite dimensional projections, see e.g. Joe [10].

In order to avoid such incompatibilities issues, it is better to use the modified Shorack’s transform: let \( X = (X(1), \ldots, X(d)) \) a \( d \)-variate discrete vector. \( X \) has a countable number \( k = (k_1, \ldots, k_d) \) of (possibly infinite) values. For each multidimensional index \( i = (i(1), \ldots, i(d)) \), corresponding to each value \( x_i \) of \( X \), pick a random vector \( U_i = (U_i(1), \ldots, U_i(d)) \) with copula \( C_i \). The set of vectors \( U_i \) for \( i \) corresponding to each value \( x_i \) of \( X \) will be called a (modified) randomisation structure \( S' \) associated with \( X \). Given \( X = x_i \), define the \( d \)-dimensional vector of modified Shorack transforms

\[
S'(X) = \sum_{x_i} 1_{X=x_i} S'(X = x_i)
\]

where \( S'(X = x_i) \) is the \( d \)-variate vector of modified Shorack transforms,

\[
S'(X = x_i) = (S'(X(1) = x(1)_i), \ldots, S'(X(d) = x(d)_i))
\]

with \( x_i = (x(1)_i, \ldots, x(d)_i) \).

**Theorem 3.5.** We can extend the probability space to carry the randomisation structure \( S' = \{U_i, i \in k\} \), with \( U_1 \sim C_1 \), so that we can construct (a representation of) \( X \), together with a \( d \)-variate vector \( \tilde{U} \), with the following properties

i) \( \tilde{U} = S'(X) \) has uniform \([0,1]\) marginals;
ii) $X = G^{-1}(\tilde{U})$;

iii) The cdf $C$ of $\tilde{U}$ satisfies Sklar’s theorem.

Proof. $X$ has a countable number $k = (k_1, \ldots, k_d)$ of (possibly infinite) values. Let $U = (U_i, i \in k)$ the vector stacking the set of $d$-variate vector $U_i$. Let $\mu_i$ the measure corresponding to the $d$-variate copula $C_i$. Let $K(\cdot)$ the probability kernel from $(\mathbb{R}^d, B(\mathbb{R}^d))$ to $([0,1]^k, B([0,1]^k)$ such that

$$K(x, \cdot) = \sum_{i \in k} 1_{x=x_i} \mu_i(\cdot).$$

Define the extension of $(\Omega, A, P)$ as $(\bar{\Omega}, \bar{A}, \bar{P})$ with

i) $\bar{\Omega} = \Omega \times [0,1]^k$;

ii) $\bar{A} = A \otimes B([0,1]^k)$;

iii) $\bar{P}(A \times B) = \int_A K(X(\omega), B) dP(\omega)$ for $A \times B \in \bar{A}$.

For $\bar{\omega} = (\omega, v)$, set $\zeta_1(\bar{\omega}) = \omega$, $\zeta_2(\bar{\omega}) = v$, $\bar{X}(\bar{\omega}) = X(\zeta_1(\bar{\omega}))$, $\tilde{U}(\bar{\omega}) = U(\zeta_2(\bar{\omega}))$. Then, $\bar{P}(\bar{X} \in A) = P(X \in A)$, and the conditional distribution of $\bar{U}$ given $\bar{X}$ is $K(\bar{X}, \cdot)$. This conditional extension of the probability space allows to identify $\bar{U}, \bar{X}$ and $X, U$ and to consider them jointly defined on a common probability space. Setting $\tilde{U} = S'(X)$ yields the claimed result, in a similar way to the proof of theorem 2.3.

The obvious modification of theorem 2.5 yields explicit the parametrisation of the set of $S$ or $S'$-copulas by the randomisation structure:

**Theorem 3.6.** Let $C^e_i$ the cdf of $U_i$ with copula function $C_i$, for $i \in k$ in the above construction,

$$C(u) = \sum_i C^e_i \left( \frac{u - G_-(x_i)}{G(x_i) - G_-(x_i)} \right) P(X = x_i)$$

Proof. Similar to that of theorem 2.5.

3.5. The probabilistic structure of discrete copulas

From a multivariate discrete vector $X \sim F$, we have constructed families of copulas of increasing complexity (respectively $R$-, $S$-, $S'$-copulas) associated to $X$, via an explicit construction of the representers $U$ of $C$. We show below a converse for the last construction, i.e. that any multivariate discrete copula is a $S'$-copula.
Theorem 3.7. Let $C$ a copula function, $G$ a vector of discrete marginal cdf and $F$ the joint cdf defined by Sklar’s theorem,

$$F(x) = C \circ G(x), \quad \forall x \in \mathbb{R}^d.$$ 

Then $C$ is a $S'$-copula.

Proof. Equivalently, the theorem asserts that if $U$ is a $d-$dimensional, marginally uniformly distributed, random vector defined canonically on $([0, 1]^d, \mathcal{B}([0, 1]^d))$, with law $dC$, so that $X$ is the random vector, defined on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, by $X = G^{-1}(U)$, then there exists a countable collection $(V_x)$ of independent, marginally uniform, random vectors of $([0, 1]^d, \mathcal{B}([0, 1]^d))$, independent of $U$, so that $U$ is equal in distribution to a modified Shorack transform $S'(X)$, with $(V_x)$ as randomization structure.

Set $A_x$ the hypercube $A_x := [G^{-1}(x), G(x)]$. Since $G$ is discrete, the set of $A_x$ define a countable partition of the unit hypercube, so that $P(X = x) = P(U \in A_x) > 0$. Since the latter probability is strictly positive, one can define, for each $x$, the conditional probability measure $\mu_{/A_x}(\cdot) := P(U \in \cdot | U \in A_x) = \frac{P(U \in \cdot \cap A_x)}{P(U \in A_x)}$.

Since the latter is a probability measure, there exist a random vector $C_x$ defined on $([0, 1]^d, \mathcal{B}([0, 1]^d))$ with law $\mu_{/A_x}$. Since $X$ has a countable number, say $k = k_1 \times \ldots \times k_d$, of values with positive probability mass, where each $k_i$ is possibly infinite, it is thus possible to define on the same probability space $U$ and a countable collection of independent random vectors $C_x$, indexed by $k$ many $x$, all independent of $U$, where $C_x$ has law $\mu_{/A_x}$; i.e. where $C_x$ has law the conditional law of $U$ given that $U \in A_x$. (Take $P \times \otimes_{x=1}^{k} \mu_{/A_x}$ as the probability measure on the corresponding product space and product sigma-algebra).

$U$ is equal in distribution to the weighted sum of independent r.v.

$$\tilde{U} := \sum_x C_x P(U \in A_x).$$
Indeed,

\[ P(\tilde{U} \in B) = \int_B \tilde{u}dP = \int_B \sum_x C_x P(U \in A_x)dP \]

\[ = \sum_x \left( \int_B C_x dP \right) P(U \in A_x) \]

\[ = \sum_x P(U \in B|U \in A_x) P(U \in A_x) \]

\[ = \sum_x P(U \in B \cap A_x) = P(U \in B). \]

From \( C_x \) define \( V_x \) by

\[ C_x = G(x, V_x). \]

where \( G(x, \lambda) \) is the vector of extended marginal cdf, and on which we based in section 3.4 what we called the modified Shorack’s multivariate transform or the conditional multivariate Ruschendorf transform. In other words, \( V_x = G_x^{-1}(C_x) \), where, for \( x = (x(1), \ldots, x(d)) \),

\[ G_x^{-1}(u) = \left( \begin{array}{c} \frac{u_1 - G_1(x(1)\downarrow)}{g_1(x(1))} \\ \vdots \\ \frac{u_d - G_d(x(d)\downarrow)}{g_d(x(d))} \end{array} \right) \]

with \( g_i(x(i)) = P(X(i) = x(i)) \) the marginal probability mass function: \( V_x \) is the image by \( G_x^{-1} \) of the law \( \mu_{/A_x} \). One can now define a representation of \( U \) as a modified Shorack transform \( S'(X) \), with \( (V_x) \) as randomization structure: define \( S'(X) \) such that, conditionally on \( X = x \), \( S'(x) = G(x, V_x) \), viz.

\[ S'(X) = \sum_x S'(x)1_{X=x}. \]

Then,

\[ P(S'(X) \in B) = \int_B \sum_x 1_{X=x} G(x, V_x)dP \]

\[ = \sum_x \int_B 1_{X=x} C_x dP \]

\[ = \sum_x P(U \in B|U \in A_x) P(U \in A_x) \]

\[ = P(U \in B). \]
4. Dependence properties of $R$-copulas

We illustrate below how the previous probabilistic constructions influence “dependence” properties of random variables. For simplicity, we reduce the study to $R$–copulas.

4.1. Independence and product copula

In Genest et al. [8] section 4.1, it is argued that whereas the copula function characterizes independence or comonotonicity for continuous random variables, it is no longer the case for discrete data. For the latter, it is shown that

$$C = \Pi \Rightarrow X \perp Y$$

where $\Pi(u, v) = uv$ is the product copula, while the converse is not true, see their example 5. The probabilistic approach with Rüschendorf’s transform allows to remedy this defect, with the following proposition,

**Proposition 4.1.** Choose the randomizers $V_1 \perp V_2$ in the construction of theorem 2.3. Then, $X \perp Y \Rightarrow C = \Pi$.

*Proof.* Let the copula representers be $U_1 = F(X, V_1)$, $U_2 = G(Y, V_2)$. As $V_1 \perp V_2$, $X \perp Y$, $(X, Y) \perp (V_1, V_2)$, the law of $((X, V_1), (Y, V_2))$ is $P((X, V_1), (Y, V_2)) = P_X \otimes P_Y \otimes P_{V_1} \otimes P_{V_2} = P(X, V_1) \otimes P(Y, V_2)$. That is to say, $(X, V_1) \perp (Y, V_2)$ thus $U_1 \perp U_2$ therefore $C = \Pi$. ☐

4.2. Impact on the concordance order

A classical way to characterize dependence properties of bivariate $(X, Y)$ random variables is through the concept of concordance and discordance. Informally, a pair of random variables are concordant if “large” values of one tend to be associated with “large” values of the other and “small” values of one with “small” values of the other, see Nelsen [11] chapter 5, or [4]. Precisely, if $(X_1, Y_1)$ and $(X_2, Y_2)$ are independent copies of $(X, Y)$, then the $(X_i, Y_i)$ are said to be concordant if $(X_2 - X_1)(Y_2 - Y_1) > 0$ holds true whereas they are said to be discordant when the reverse inequality is valid. Related to these ideas is the concordance partial order $\prec_c$ on bivariate vectors: let
(X_1, Y_1) \sim H_1 \text{ and } (X_2, Y_2) \sim H_2 \text{ two vectors with identical marginals } \overset{d}{=} X_2, Y_1 \overset{d}{=} Y_2, \text{ then }

(X_1, Y_1) \prec_c (X_2, Y_2) \iff H_1(s, t) \leq H_2(s, t) \quad (5)

holds for all \( s, t \in \mathbb{R} \). In (5), \( H_1 \) and \( H_2 \) can equivalently be replaced with the joint survival functions \( \overline{H}_1(s, t) := P(X_1 > s, Y_1 > t) \), \( \overline{H}_2(s, t) := P(X_2 > s, Y_2 > t) \). This relation thus encapsulates the idea of comparing vectors by their probability of (upper or lower) concordance. The corresponding concordance order on distribution functions will also be denoted by \( H_1 \prec_c H_2 \).

In our context of probabilistic constructions of copula representers, the question arises how the concordance order is transformed by the randomization structure. This question can be understood in the two following ways.

4.2.1. Influence of the randomizers on the concordance order

Let \( (X, Y) \sim H \) a discrete cdf with marginal cdfs \( F, G \). Let \( (U_1, V_1) \sim R_1 \) and \( (U_2, V_2) \sim R_2 \) two pairs of marginally uniform random variables with copula \( R_1, R_2 \). Construct the pairs of \( R \)-copula representers, \( (U_1^X, U_1^Y) := (F(X, U_1), G(Y, V_1)) \) and \( (U_2^X, U_2^Y) := (F(X, U_2), G(Y, V_2)) \) with randomisation structure \( (U_1, V_1) \) and \( (U_2, V_2) \), respectively. The influence of the randomisers on the copula representers is shown in the following proposition.

**Proposition 4.2.** If \( R_1 \prec_c R_2 \), then \( (U_1^X, U_1^Y) \prec_c (U_2^X, U_2^Y) \).

**Proof.** Since \( R_1(u, v) \leq R_2(u, v) \) for all \((u, v) \in [0, 1]^2\), (4) entails that that, for all \(0 \leq s, t \leq 1\),

\[
P(U_1^X \leq s, U_1^Y \leq t) = \sum_{x,y} R_1^x \left( \frac{s - F_-(x)}{f(x)}, \frac{t - G_-(y)}{g(y)} \right) P(X = x, Y = y) \\
\leq \sum_{x,y} R_2^x \left( \frac{s - F_-(x)}{f(x)}, \frac{t - G_-(y)}{g(y)} \right) P(X = x, Y = y) \\
= P(U_2^X \leq s, U_2^Y \leq t),
\]

where \( f(x), g(y) \) stand for the probability mass functions of \( X, Y \) at \( x, y \). \(\square\)
4.2.2. Preservation of the concordance order at the copula level

Let \((X_1, Y_1) \sim H_1\) and \((X_2, Y_2) \sim H_2\) two discrete vectors with identical marginals \(X_1 \overset{d}{=} X_2 \sim F, Y_1 \overset{d}{=} Y_2 \sim G\). Let \((U, V)\) a randomization structure of marginally uniforms r.v. with copula function \(R\). Construct the pairs of \(R\)-copula representers, \((U_1^X, U_1^Y) := (F(X_1, U), G(Y_1, V))\) and \((U_2^X, U_2^Y) := (F(X_2, U), G(Y_2, V))\) with the same randomisation structure.

**Proposition 4.3.** \((X_1, Y_1) \prec_c (X_2, Y_2)\) implies \((U_1^X, U_1^Y) \prec_c (U_2^X, U_2^Y)\).

**Proof.** Fix \((s, t) \in (0, 1)^2\). From (4), the cdf of the copula representers writes

\[
P(U_1^X \leq s, U_1^Y \leq t) = \int \phi_{s,t}(x, y) dH_1(x, y)
\]

with \(\phi_{s,t}\) is the bounded positive function defined by

\[
\phi_{s,t}(x, y) := R^e\left(\frac{s - F_-(x)}{f(x)}, \frac{t - G_-(y)}{g(y)}\right)
\]

for \((x, y)\) in the support of \(F, G\). We extend it by setting \(\phi_{s,t}(\infty, y) = \phi_{s,t}(x, \infty) = 0\). \((s, t) \mapsto \phi_{s,t}(x, y)\) is the cdf of a pair of random variables marginally uniformly distributed on \((F_-(x), F(x))\) and \((G_-(y), G(y))\). In other words, it writes

\[
\phi_{s,t}(x, y) = \begin{cases}
0 & \text{if } s < F_-(x) \text{ or } t < G_-(y) \\
1 & \text{if } s > F(x) \text{ and } t > G(y) \\
\frac{s - F_-(x)}{F(x) - F_-(x)} & \text{if } F_-(x) \leq s \leq F(x) \text{ and } t > G(y) \\
\frac{t - G_-(y)}{g(y) - G_-(y)} & \text{if } s > F(x) \text{ and } G_-(y) \leq t \leq G(y) \\
R\left(\frac{s - F_-(x)}{F(x) - F_-(x)}, \frac{t - G_-(y)}{g(y) - G_-(y)}\right) & \text{if } F_-(x) \leq s \leq F(x) \text{ and } G_-(y) \leq t \leq G(y)
\end{cases}
\]

Since \((s, t)\) are fixed, introducing \(x^* = F^{-1}(s), y^* = G^{-1}(t)\), so that \(F_-(x^*) < s \leq F(x^*)\) and \(G_-(y^*) < t \leq G(y^*)\), allows to rewrite \(\phi_{s,t}\) as

\[
\phi_{s,t}(x, y) = \begin{cases}
0 & \text{if } x > x^* \text{ or } y > y^* \\
1 & \text{if } x < x^* \text{ and } y < y^* \\
\frac{s - F_-(x^*)}{f(x^*)} & \text{if } x = x^* \text{ and } y < y^* \\
\frac{t - G_-(y^*)}{g(y^*)} & \text{if } x < x^* \text{ and } y = y^* \\
k_{s,t} & \text{if } x = x^* \text{ and } y = y^*
\end{cases}
\]
with \( k_{s,t} := R\left(\frac{s-F(x^*)}{f(x^*)}, \frac{t-G(y^*)}{g(y^*)}\right) \). Fréchet-Hoeffding bounds on \( R \) entails that \((x, y) \mapsto \phi_{s,t}(x, y)\) is \( \Delta \)-Monotone, in the sense that

\[
\Delta \phi_{s,t}(x_1, y_1, x_2, y_2) := \phi_{s,t}(x_1, y_1) + \phi_{s,t}(x_2, y_2) - \phi_{s,t}(x_2, y_1) - \phi_{s,t}(x_1, y_2)
\]

is \( \geq 0 \) for any two points \((x_1, y_1), (x_2, y_2)\) in the support of \( F, G \). It thus define a (discrete) measure \( \mu_{s,t}(\cdot) \) on the set of points \((x, y)\) in the support of \( F, G \), by setting

\[
\mu_{s,t}(\cdot) = \Delta \phi_{s,t}(\cdot)_{(\infty, \infty)}^{(\infty, \infty)}
\]

where \([x, \infty)\) stands for the discrete interval \( \{x, \ldots, \infty\} \) of points in the support of \( F \) larger than or equal to \( x \), and similarly for \([y, \infty)\]. Since \( \phi_{s,t}(x, y) = \Delta \phi_{s,t}(\cdot)_{(x, y)} \), we can identify \( \mu_{s,t} \) with \( \phi_{s,t} \) and the integration by part formula of Rüschendorf [14] in his theorem 3 is valid, as \( H_1(x, y) = \Delta H_1^{(x,y)} \). Indeed,

\[
\int \Delta \phi_{s,t}(x, y) \, dH_1(x, y) = \int \int 1_{u \in [x, \infty), v \in [y, \infty]} \, d\phi_{s,t}(u, v) \, dH_1(x, y) = \int \int dH_1(x, y) \, d\phi_{s,t}(u, v) = \int \Delta H_1^{(x, y)} \, d\phi_{s,t}(u, v)
\]

by Fubini. Thus,

\[
P(U_1^X \leq s, U_1^Y \leq t) = \int \phi_{s,t}(x, y) \, dH_1(x, y) = \int H_1(x, y) \, d\phi_{s,t}(x, y)
\]

and the assumption \( H_1 \prec_c H_2 \) yields, by a reverse integration by parts, that

\[
P(U_1^X \leq s, U_1^Y \leq t) \leq P(U_2^X \leq s, U_2^Y \leq t).
\]

\(\square\)

4.3. Kendall’s \( \tau \)

A measure of the strength of association based on concordance and discordance ideas is Kendall’s \( \tau \). For \((X_1, Y_1)\) and \((X_2, Y_2)\) two independent copies of \((X, Y)\) with cdf \( H \), it is defined by

\[
\tau(X, Y) := \tau(H) := P(\text{concordance}) - P(\text{discordance}) = P((X_2 - X_1)(Y_2 - Y_1) > 0) - P((X_2 - X_1)(Y_2 - Y_1) < 0)
\]
With continuous random variables, $P(\text{tie}) = P(X_2 - X_1)(Y_2 - Y_1) = 0 = 0$, thus

$$
\tau(X, Y) = 2P(\text{concordance}) - 1 = 4 \int_{[0,1]^2} C(u,v)dC(u,v) - 1 := \tau(C),
$$

where $C$ is the (unique) copula of $H$, see [11, 4]. As explained in [8] and in [4], $\tau(X,Y)$ only depends on the copula function and as such remains invariant w.r.t. any strictly increasing transformations of the variables. Hence, it is scale invariant in the utmost manner and thus is an attractive measure of dependence in this continuous case. A probabilistic way to view this is that since $\tau(X,Y) = \tau(F(X), G(Y))$, there is agreement of the coefficient at the level of observations $(X,Y)$ and at the copula representers’ one.

However, these nice properties break down in the discrete case. In particular, due to the possible presence of ties,

$$
\tau(X, Y) = 2P(\text{concordance}) - 1 + P(\text{tie}) = 4P(X_2 < X_1, Y_2 < Y_1) - 1 + P(X_1 = X_2 \text{ or } Y_1 = Y_2)
$$

which restricts the range of possible values for $\tau(X,Y)$, as explained in [4]. Moreover, since “each possible choice of copula $C \in C_H$ leads to different values for $\tau(C)$”, “the probabilistic and analytical definitions of $\tau$ […] do not coincide”, in the sense that one may have that $\tau(X,Y) \neq \tau(C)$, see [8] section 4.3. These authors review some modified versions of the coefficient and [4] propose a “continuing technique” which turns count data $(X,Y)$ into continuous ones $(X^*, Y^*)$ so that $\tau(X,Y) = \tau(X^*, Y^*)$.

The next proposition relates Kendall’s $\tau$ at the observational and copula level in the proposed probabilistic construction. To that purpose, let $F, G$ be the marginal cdfs of $(X,Y)$ and $f, g$ their probability mass functions. Let $(U_1, V_1), (U_2, V_2)$ be i.i.d. randomization structure with copula $R$ and independent of $(X_1, Y_1)$ and $(X_2, Y_2)$. Construct the pair of copula representers,

$$
(U_1^X, U_1^Y) = (F_-(X_1) + U_1 f(X_1), G_-(Y_1) + V_1 g(Y_1))
$$
$$
(U_2^X, U_2^Y) = (F_-(X_2) + U_2 f(X_2), G_-(Y_2) + V_2 g(Y_2))
$$

Proposition 4.4.

$$
\tau(X,Y) = \tau(U_1^X, U_1^Y) - \tau(R)P(X_1 = X_2, Y_1 = Y_2)
$$
Proof.

\[ P(U_1^X < U_2^X, U_1^Y < U_2^Y) \]

\[ = \int P \left( \begin{array}{c} F_-(x_1) + U_1 f(x_1) < F_-(x_2) + U_2 f(x_2), \\ G_-(y_1) + V_1 g(y_1) < G_-(y_2) + V_2 g(y_2) \end{array} \right) \, dH(x_1, y_1) dH(x_2, y_2) \]

Since \( F_-(x_1) + U_1 f(x_1) \) is uniformly distributed on \((F_-(x_1), F(x_1))\) and similarly for the other variables, the probability under the integral sign is equal to

\[
\begin{cases}
1 & \text{if } x_1 < x_2 \text{ and } y_1 < y_2 \\
\frac{1}{2} & \text{if } x_1 = x_2 \text{ and } y_1 < y_2 \\
0 & \text{if } x_1 = x_2 \text{ and } y_1 > y_2 \\
\frac{1}{2} & \text{if } y_1 = y_2 \text{ and } x_1 < x_2 \\
0 & \text{if } y_1 = y_2 \text{ and } x_1 > x_2 \\
P(U_1 < U_2, V_1 < V_2) & \text{if } x_1 = x_2 \text{ and } y_1 = y_2
\end{cases}
\]

Note that,

\[ P(U_1 < U_2, V_1 < V_2) = \int R(u, v) dR(u, v) = \frac{\tau(R) + 1}{4}. \]

Therefore,

\[ P(U_1^X < U_2^X, U_1^Y < U_2^Y) \]

\[ = P(X_1 < X_2, Y_1 < Y_2) + \frac{\tau(R) + 1}{4} P(X_1 = X_2, Y = 1 = Y_2) \]

\[ + \frac{1}{2} \left( P(X_1 = X_2, Y_1 < Y_2) + P(Y_1 = Y_2, X_1 < X_2) \right) \]

By symmetry,

\[ P((U_2^X - U_1^X)^2(Y_2^Y - U_1^Y)^2 > 0) \]

\[ = P((X_2 - X_1)(Y_2 - Y_1) > 0) \]

\[ + \frac{1}{2} \left( P(X_1 = X_2, Y_1 \neq Y_2) + P(Y_1 = Y_2, X_1 \neq X_2) \right) \]

\[ + \frac{\tau(R) + 1}{2} P(X_1 = X_2, Y_1 = Y_2), \]

so that,

\[ \tau(X, Y) = 2P(\text{concordance of } (X, Y)) - 1 + P(\text{tie of } (X, Y)) \]

\[ = 2P(\text{concordance of } (U_1^X, U_1^Y)) - 1 - \tau(R)P(X_1 = X_2, Y_1 = Y_2) \]

\[ = \tau(U_1^X, U_1^Y) - \tau(R)P(X_1 = X_2, Y_1 = Y_2), \]

24
as \((U_X, U_Y)\) have continuous marginals.

The following obvious corollary gives a generalization of [4]'s continuing technique to preserve concordance coefficients.

**Corollary 4.5.** \(R = \Pi\) implies \(\tau(H) = \tau(C)\).

5. Some examples and applications

We show below some examples and applications of the previous results. For simplicity, we restrict ourselves to the \(R\)-copula construction. Similar results for the other constructs can be obtained in a similar manner.

5.1. Bivariate discrete \(R\)-copulas

5.1.1. Analytical formula

In the bivariate case, we can exhibit a somehow more tractable formula for the set of \(R\)-copulas compatible with its parent discrete distribution. For an univariate cdf \(G\), define its left continuous version as

\[
G_-(x) = P(X < x) = \lim_{\varepsilon \downarrow 0} G(x - \varepsilon).
\]

Moreover, remark 3 for a bidimensional copula function \(C\) means that its extension \(C^e\) writes

\[
C^e(u, v) = \begin{cases} 
  C(u, v) & (u, v) \in (0, 1)^2 \\
  u & u \in (0, 1), v > 1 \\
  v & v \in (0, 1), u > 1 \\
  1 & u > 1, v > 1 \\
  0 & u \leq 0 \text{ or } v \leq 0 
\end{cases}
\]

**Proposition 5.1.** For a bivariate discrete distribution, the set of \(R\)-copulas constructed in the first probabilistic approach writes

\[
C(u) = F(G_{1\rightarrow} \circ Q_1(u_1), G_{2\rightarrow} \circ Q_2(u_2)) + (u_1 - G_{1\rightarrow} \circ Q_1(u_1))P(X_2 < Q_2(u_2)|X_1 = Q_1(u_1)) + (u_2 - G_{2\rightarrow} \circ Q_2(u_2))P(X_1 < Q_1(u_1)|X_2 = Q_2(u_2)) + R \left( \frac{u_1 - G_{1\rightarrow}(Q_1(u_1))}{G_1(Q_1(u_1)) - G_{1\rightarrow}(Q_1(u_1))}, \frac{u_2 - G_{2\rightarrow}(Q_2(u_2))}{G_2(Q_2(u_2)) - G_{2\rightarrow}(Q_2(u_2))} \right) \times P(X_1 = Q_1(u_1), X_2 = Q_2(u_2)) \tag{6}
\]

where \(R\) is the copula function corresponding to the cdf of the randomizers.

**Proof.** Omitted. Available on request. \(\square\)
5.1.2. Example: bivariate Bernoulli

Applied to the simplest bivariate distribution, a bivariate Bernoulli distribution defined by

\[ P(X = 0, Y = 0) = a \quad P(X = 0, Y = 1) = b \]
\[ P(X = 1, Y = 0) = c \quad P(X = 1, Y = 1) = d, \]

with \( d = 1 - a - b - c \), and

\[ 0 \leq a, b, c \leq 1, \quad 0 \leq a + b, a + c \leq 1, \quad 0 \leq a + b + c \leq 1, \]

the previous proposition writes

**Proposition 5.2.** The set of \( R \)-copula writes

\[
C(u_1, u_2) = \begin{cases} 
  aR \left( \frac{u_1}{a+b}, \frac{u_2}{a+c} \right), & 0 < u_1 < a + b, 0 < u_2 < a + c \\
  a \frac{u_2}{a+c} + cR \left( \frac{u_1-(a+b)}{c+d}, \frac{u_2}{a+c} \right), & a + b < u_1 < 1, 0 < u_2 < a + c \\
  a \frac{u_1}{a+b} + bR \left( \frac{u_1}{a+b}, \frac{u_2-(a+c)}{b+d} \right), & 0 < u_1 < a + b, a + c < u_2 < 1 \\
  a \frac{u_1}{a+b} + b \frac{u_2-(a+c)}{b+d} + cR \left( \frac{u_1-(a+b)}{c+d}, \frac{u_2-(a+c)}{b+d} \right), & a + b < u_1 < 1, a + c < u_2 < 1 \\
  +dR \left( \frac{u_1-(a+b)}{c+d}, \frac{u_2-(a+c)}{b+d} \right), & a + b < u_1 < 1, a + c < u_2 < 1 
\end{cases}
\]

where \( R \) is the copula of the randomizers.

**Proof.** Theorem 2.5 specialises as follows:

\[
C(u_1, u_2) = aR \left( \frac{u_1}{a+b}, \frac{u_2}{a+c} \right) + bR \left( \frac{u_1}{a+b}, \frac{u_2-(a+c)}{b+d} \right) + cR \left( \frac{u_1-(a+b)}{c+d}, \frac{u_2}{c+d} \right) + dR \left( \frac{u_1-(a+b)}{c+d}, \frac{u_2-(a+c)}{b+d} \right),
\]

Taking into account remark 3 leads to the claimed result. \( \Box \)
5.1.3. Pattern recognition with a continuous regressor

Let \( \mathbf{X} = (X, Y) \) a bivariate vector with \( X \) continuous and \( Y \) discrete with values in \( \{0, 1\} \). Such a setting is often encountered in fields like pattern recognition, supervised learning or classification, where one tries to predict a binary outcome measurement \( Y \) based on a set of continuous features \( X \), see e.g. Hastie et al [9]. Denote \( F_{X,Y} \) the joint cdf of \( (X, Y) \), and \( F \) and \( G \) their respective marginals. The distribution of \( (X, Y) \) is completely specified once one has determined \( p_0 = P(Y = 0) \), and the two functions \( x \to F_{X|Y=0}(x), \ x \to F_{X|Y=1}(x) \). Assume \( 0 < p_0 < 1 \) to avoid degeneracy of \( Y \). By precising theorem 2.5, we describe below the structure of \( R \)-copulas associated with \( F_{X,Y} \).

**Proposition 5.3.** Let \( C \) a \( R \)-copula associated to \( (X, Y) \) in the above setting. Then \( C \) is unique and given by

\[
C(u_1, u_2) = F_{X|Y=0}(F_X^{-1}(u_1))(u_2 \mathbb{1}_{u_2 < p_0} + p_0 \mathbb{1}_{u_2 \geq p_0})
+ F_{X|Y=1}(F_X^{-1}(u_1))(u_2 - p_0) \mathbb{1}_{u_2 > p_0}
\]

for \( 0 < u_1, u_2 < 1 \) with \( F(x) = p_0 F_{X|Y=0}(x) + (1 - p_0) F_{X|Y=1}(x) \).

**Proof.** Set \((V_1, V_2) \perp (X, Y)\) a randomizer. The distributional transform reduces to \( U_1 = F(X), U_2 = G(Y, V_2) \). Then, for \( 0 < u_1, u_2 < 1 \),

\[
C(u_1, u_2) = P(U_1 \leq u_1, U_2 \leq u_2)
= \sum_{y=0,1} P(F(X) \leq u_1, V_2 \leq \frac{u_2 - G_-(y)}{P(Y = y)} | Y = y) P(Y = y)
= \sum_{y=0,1} P(V_2 \leq \frac{u_2 - G_-(y)}{P(Y = y)}) P(F(X) \leq u_1 | Y = y) P(Y = y)
\]

by independence. As \( F \) is continuous, \( P(F(X) \leq u_1 | Y = y) = F_{X|Y=y}(F_X^{-1}(u_1)) \).
Since \( V_2 \sim U_{(0,1)} \),

\[
P(V_2 \leq u_2/p_0) = \frac{u_2}{p_0} \mathbb{1}_{u_2 < p_0} + \mathbb{1}_{u_2 \geq p_0}
\]
and

\[
P(V_2 \leq \frac{u_2 - p_0}{1 - p_0}) = \frac{u_2 - p_0}{1 - p_0} \mathbb{1}_{p_0 < u_2}.
\]

Rearranging yields the claimed result. \( \square \)

Note that the extension is straightforward for several continuous regressors \( X_1, \ldots, X_d, \) or for \( Y \) taking more than two values.
5.2. The empirical copula

5.2.1. Introduction

A prominent example of a discrete multivariate distribution is the empirical distribution function associated with an i.i.d. sample $X_1, \ldots, X_n$ of observations from a parent distribution $F$, with marginals $G$ and associated copula $C$. Define the joint empirical distribution function as,

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{X_i \leq x},$$

and the vector of marginal ecdf $G_n := (G_{n,1}, \ldots, G_{n,d})$ with

$$G_{n,j}(x_j) = F_n(\infty, \ldots, \infty, x_j, \infty, \ldots, \infty), \quad j = 1, \ldots, d.$$  

As $F_n$ is a genuine cdf, there exists, by theorem 1.1, a (set of) copulas $C_n$ associated to $F_n$, viz. such that

$$F_n(x) = C_n(G_n(x)). \quad (7)$$

$C_n$ is an empirical copula function associated to the sample $X_1, \ldots, X_n$. As the marginals $G_n$ are discontinuous, such empirical copula is not unique.

The discussion in the introduction appears relevant to stress the often overlooked fact that so-called “empirical copulas” encountered in the literature are not genuine copulas. Indeed, Rüschendorff [13], equation (25), defines the empirical copula as a precopula of the second type, viz. the cdf of the multivariate rank vector $G_n(X_i)$,

$$C_n(u) := C_n^*(u) = P(G_n(X_i)), \quad u \in I^d.$$  

whereas Van der Vaart and Wellner [20] section 3.9.4.4. p.389 or Fermanian [7] defines the empirical copula as a precopula of the first type, viz. as

$$C_n(u) := C_n^{**}(u) = F_n \circ G_n^{-1}(u), \quad u \in I^d.$$  

In both cases, the objects are not copulas per se, as they do not have uniform marginals. When the marginals $G$ are continuous, $C_n^*$ and $C_n^{**}$ agree with $C_n$ on the grid of points $\mathcal{L}_n = \{k/n\}$, with $k = (k_1, \ldots, k_d) \in \{0, \ldots, n\}^d$ as shown in Deheuvels [3] proposition 1.5.

Yet, this discrepancy can have practical harmful consequences. For example, when one wants to perform a Cramér-von Mises or Kolmogorov-Smirnov
type Goodness of fit test, a careless substitution of $C_n$ by $C_n^*$ or $C_n^{**}$ in
statistics like
\[ T_n := \int_{[0,1]^d} (C_n(u) - C_\theta(u))^2 dC_\theta(u) \]
where \{C_\theta, \theta \in \Theta\} is a set of candidate copulas, may lead to unwanted errors
due to the divergence w.r.t. $C_n$ away of the grid points $L_n$, cf. Durrleman et
al. [5].

5.2.2. The $R$-empirical copula

The proposed constructions ($R$, $S$, $S'$ copulas) allow to remedy this defect
by specifying a particular candidate. This can be useful when one needs
that $C_n$ be a genuine cdf, e.g. for computing quantities like $\int \phi dC_n$ for a
measurable integrable function $\phi$, see e.g. Tchen [18] corollary 2.2. For
simplicity, we restrict ourselves below to the construction of $R$-empirical
copula.

Let $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ the empirical measure associated with the sample
$X_1, \ldots, X_n$. We now work conditionnally on the sample. Let $X_n^*$ distributed
according to $\mu_n$ and $V$ a vector of uniform marginals distributed according
to a copula function $R$, independent of $X_n^*$. One can define the $R$-empirical
copula associated to $F_n$ by
\[ C_n^R(u) = P(G_n(X_n^*, V)) \leq u) \]

$X_n^*$ is the bootstrap version of $X$ obtained by resampling the original
data. The above probability can be computed by Monte-Carlo: construct a
bootstrap sample ($X_{1,n}^*, \ldots, X_{n,n}^*$) of $B$ i.i.d. r.v. distributed according to
$\mu_n$. Generate (independently of the bootstrap sample) an independent se-
cquence of randomizers ($V_1, \ldots, V_B$) distributed according to $R$. A bootstrap
representer of $C_n^R$ is obtained, by setting
\[ U_i^* = G_n(X_{i,n}^*, V_i), \quad i = 1, \ldots, B \]

By corollary 2.2, $U_i^*$ are i.i.d. distributed according to $C_n^R$, conditionally on
$F_n$. $C_n^R$ can be estimated by the bootstrap empirical cdf, viz.
\[ C_{n,B}^{*,R}(u) := \frac{1}{B} \sum_{i=1}^B 1_{U_i^* \leq u}. \]

Indeed, one has the following convergence theorem:
Theorem 5.4. \( \sqrt{B} (C_{n,B}^R - C_n^R) \xrightarrow{d} \mathcal{G}_{C_n^R} \), as \( B \to \infty \), conditionally on \( F_n \), where \( \mathcal{G}_{C_n^R} \) is a \( C_n^R \)-Brownian Bridge, viz. a centered Gaussian process with covariance function

\[
E[\mathcal{G}_{C_n^R}(u)\mathcal{G}_{C_n^R}(v)] = C_n^R(u \wedge v) - C_n^R(u)C_n^R(v)
\]

Proof. Conditionally on \( F_n \), \( U_i^* \) are i.i.d. with cdf \( C_n \). A simple conditional application of Donsker's theorem yields weak convergence of the bootstrap ecdf to the corresponding \( C_n \) Brownian bridge. \( \square \)

5.3. Simulating the \( R \)-empirical copula: a bivariate Bernoulli example

We illustrate below how one can practically simulate the \( R \)-empirical copula of a bivariate Bernoulli distribution.

5.3.1. Simulating the empirical distributional transform

For a univariate Bernoulli, \( X \sim B(1, p) \), with \( q = 1 - p \), the distributional transform can be written as

\[
F(X, V) = Vq\mathbb{1}_{X=0} + (q + VP)\mathbb{1}_{X=1} = Vq(1 - X) + (q + VP)X
\]

with \( V \sim U_{[0,1]} \). If \( p \) is unknown but that one observes an i.i.d. sample \( X_1, \ldots, X_n \), a bootstrap synthetic sample \( X_1^*, \ldots, X_B^* \) of \( B = 10000 \) realizations distributed according to \( F_n \) can be simulated. Conditionally on \( F_n \), the bootstrap r.v. \( X^* \) is distributed according to a \( B(1, 1 - N_0/n) \) distribution, with \( N_0 := \#\{X_i = 0, i = 1, \ldots, n\} = F_n(0) \). The empirical distributional transform, i.e. the distributional transform of \( X^* \), turns \( X^* \) into \( U^* = F_n(X^*, V) \). The latter is uniformly distributed on \( [0,1] \), conditionally on \( F_n \), as is illustrated by the plot of its ecdf \( \frac{1}{B} \sum_{i=1}^{B} 1_{U_i^* \leq u} \) in figure 1.

5.3.2. Bootstrapping the \( R \)-empirical copula

For the bivariate Bernoulli distribution of section 5.1.2, one can construct from a sample of \( (X, Y) \) values bootstrap replications \( (X^*, Y^*) \) together with randomizers \( (V_1, V_2) \), as explained in section 5.2.2. We chose \( V_1 \perp V_2 \), viz. \( R = \Pi \). A scatterplot of the copula representers \( (U_1, U_2) = (F_n(X^*, V_1), G_n(Y^*, V_2)) \) is shown in figure 2. The interest of the ecdf transform becomes clear: whereas a scatterplot of the original data gives only the four points \( (0,0), (0,1), (1,0), (1,1) \) and give little visual information on the dependence structure, the mass concentrated in the original sample has been spread at the copula representers level and cures the too-many-ties problem of the ranks, see Genest et al. [8] section 6.
Figure 1: Empirical cdf of $U^*$

Figure 2: Scatter plot of the copula representers for the bivariate Bernoulli with $a = 0.2$, $b = 0.1$, $c = 0.3$
6. Conclusion

We presented several explicit probabilistic constructions of copulas associated with a discrete vector $X$. These constructions were obtained by multivariate extensions of randomised probability integral transforms, which allowed to define a multivariate coupling, viz. a representation of a vector $U$ whose cdf is a copula function satisfying Sklar’s theorem for $X$. The obtained families of $R,S,S’$ -copulas, indexed by a randomisation structure of increasing complexity, may serve different purposes. The most sophisticated construction gives a probabilistic characterisation of the set of copula functions associated to $X$: every copula associated to $X$ is an $S’$-copula. The less sophisticated $R$ -copula construct is e.g. sufficient to construct “a” copula which preserve the “dependence” properties of the original $X$ at the copula level, obtaining a generalisation of [4]’s continuing technique. Moreover, the impact of the added randomisation in the latter construct on independence and concordance order of the original $X$ was elucidated and quantified via Kendall’s $\tau$. Explicit analytical formulas of the obtained copulas can then be derived in particular cases, like the bivariate one or for pattern recognition. As an important application, when the true cdf of $X$ is unknown, one obtain simple ways to turn the empirical copula function into a genuine copula function, which can in turn be easily simulated.

Maybe the most important interest of such constructions is of practical and statistical nature: by showing that one has to enlarge the probability space to carry an extraneous randomisation, there are important implications from an inferential point of view. We recommend the interested reader to refer to the companion paper [6], where we explore the epistemological issues involved in estimating parametric copula models for discrete-valued data.

References


2000.
