Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 98 (2007) 544 – 567
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

On rank correlation measures for non-continuous random variables

Johanna Nešlehová

RiskLab, Department of Mathematics, ETH Zurich, CH-8092 Zurich, Switzerland

Received 25 May 2005
Available online 25 January 2006

Abstract

For continuous random variables, many dependence concepts and measures of association can be expressed in terms of the corresponding copula only and are thus independent of the marginal distributions. This interrelationship generally fails as soon as there are discontinuities in the marginal distribution functions. In this paper, we consider an alternative transformation of an arbitrary random variable to a uniformly distributed one. Using this technique, the class of all possible copulas in the general case is investigated. In particular, we show that one of its members—the standard extension copula introduced by Schweizer and Sklar—captures the dependence structures in an analogous way the unique copula does in the continuous case. Furthermore, we consider measures of concordance between arbitrary random variables and obtain generalizations of Kendall's tau and Spearman's rho that correspond to the sample version of these quantities for empirical distributions.

AMS 1991 subject classification: 62H20; 62G20

Keywords: Copula; Empirical copula; Kendall's tau; Measures of association; Non-continuous; Spearman's rho

1. Introduction

Monotonic dependence between random variables is of key importance in many practical applications, see for instance Lehmann [12], Jogdeo [9] or Embrechts et al. [4] and McNeil et al. [15] for some recent contexts in insurance and finance. Since the early works of Hoeffding [6], Kruskal [10] and Lehmann [11], numerous measures of monotonic dependence between random variables and/or samples have been proposed and studied extensively. In most cases, however, the random variables involved are assumed to have continuous distribution functions. A partic-

E-mail address: johanna@math.ethz.ch.

0047-259X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2005.11.007

ularly elegant contribution to the understanding of monotonic dependence between continuous random variables provides the copula approach, see for instance Schweizer and Wolff [22], and Nelsen [17,18]. If the random variables under study have continuous distribution functions, the corresponding copula is unique and remains the same if the random variables are (almost surely) subject to strictly increasing transformations, such as the change of scale or location. As monotonic dependence also has this invariance property, Scarsini [20] shows that it can be determined from the corresponding copula alone. Consequently, concordance measures like Kendall's tau and Spearman's rho can be expressed solely in terms of the corresponding copula.

The main contribution of this paper is a generalization of rank correlation measures for non-continuous random variables. Marshall [14] obtains a number of counterexamples through which he shows that an arbitrary measure of association depending solely on the copula is generally trivial, i.e. a constant. Involving marginal distributions becomes inevitable, but the question how this should be accomplished is far from being well understood. Few attempts have been made, however, see for example Hoeffding [7], Tajar et al. [23], Tajar et al. [24], Mesfioui and Tajar [16] and Denuit and Lambert [3] who address the purely discrete case. Though the latter paper by Denuit and Lambert [3] investigates similar issues, the present paper offers alternative proofs of common results and discusses several more general problems.

In this paper, a technique allowing to adapt the copula-based approach to the general non-continuous case is presented. It relies on an alternative transformation of an arbitrary random variable to a uniformly distributed one. It becomes clear that the key role of the unique copula in the continuous case is taken over by the so-called standard extension copula introduced by Schweizer and Sklar [21]. This result allows for generalizations of rank correlation measures, in particular of Kendall's tau and Spearman's rho.

The paper is organized as follows: basic notation and theoretical background with special focus on copulas and concordance is established in Section 2. Typical fallacies that occur when allowing for discontinuities are highlighted. In Section 3, the technique relying on an alternative transformation of the marginals is discussed and the main results concerning the dependence structures between arbitrary marginals obtained. These are used in Section 4 to generalize Kendall's tau and Spearman's rho and examine their properties. Finally in Section 5, empirical copulas are considered and it is shown that for empirical distributions, the generalizations of Kendall's tau and Spearman's rho coincide with their sample versions known from statistics.

2. Notations and preliminaries

The subject of our study will be a real-valued random vector $\boldsymbol{X} := (X_1, X_2)$ with joint distribution function $F_{\boldsymbol{X}}$ and marginals F_{X_1} and F_{X_2}. If not otherwise stated, F_{X_1} and F_{X_2} are arbitrary with ranges ran F_{X_1} and ran F_{X_2}. When confusion may arise, by continuous random variables (or marginals) we always mean random variables that have continuous distribution functions; absolute continuity of the distribution with respect to Lebesgue measure is, however, generally not assumed. The marginals are linked to the joint distribution function through the so-called *copula* function, which is a bivariate distribution function on $[0, 1]^2$ with uniform marginals. Copula functions are a key ingredient in the study of monotonic dependence. We will discuss this issue briefly below; for further details and proofs, see for instance Nelsen [18] or Joe [8].

The main result is the well-known Sklar's Theorem (cf. [21]) that guarantees that there exists at least one copula $\mathcal{C}_{\boldsymbol{X}}$ such that

$$F_{\boldsymbol{X}}(x, y) = \mathcal{C}_{\boldsymbol{X}}(F_{X_1}(x), F_{X_2}(y)) \quad \text{for all } x, y \in \mathbb{R}. \tag{1}$$

$\mathcal{C}_{\boldsymbol{X}}$ is uniquely determined on ran $F_{X_1} \times$ ran F_{X_2} and, due to its uniform continuity, even on the closure of ran $F_{X_1} \times$ ran F_{X_2}. Consequently, $\mathcal{C}_{\boldsymbol{X}}$ is unique if and only if F_{X_1} and F_{X_2} are continuous. In the general case, however, there exist several copulas satisfying (1) and we refer to them as to *possible copulas*. The perhaps best known technique for obtaining possible copulas is the one used by Schweizer and Sklar [21] extending the (unique) values of the copula from the closure of ran $F_{X_1} \times$ ran F_{X_2} to $[0, 1]^2$ by linear interpolation. The so-called *standard extension copula* results, which is formally defined as follows:

$$\mathcal{C}^S_{\boldsymbol{X}}(u_1, u_2) = (1-\lambda_1)(1-\lambda_2)\mathcal{C}_{\boldsymbol{X}}(a_1, a_2) + (1-\lambda_1)\lambda_2\mathcal{C}_{\boldsymbol{X}}(a_1, b_2) + \lambda_1(1-\lambda_2)\mathcal{C}_{\boldsymbol{X}}(b_1, a_2) + \lambda_1\lambda_2\mathcal{C}_{\boldsymbol{X}}(b_1, b_2) \tag{2}$$

with

$$\lambda_i = \begin{cases} \frac{u_i - a_i}{b_i - a_i} & \text{if } a_i < b_i \\ 1 & \text{if } a_i = b_i \end{cases} \quad \text{for } i = 1, 2,$$

and a_i and b_i being the least and the greatest element in the closure of ran F_{X_i} such that $a_i \leqslant u_i \leqslant b_i$, $i = 1, 2$.

A precise definition of concordance and concordance measures has been formulated by [20]. As we will consider solely random vectors with common marginals, it suffices to note that $\boldsymbol{X}$ is *more concordant* than $\boldsymbol{X}^*$ if and only if $F_{\boldsymbol{X}}(x_1, x_2) \geqslant F_{\boldsymbol{X}^*}(x_1, x_2)$ for all $(x_1, x_2) \in \mathbb{R}^2$.

Assume now that X_1 and X_2 have continuous distribution functions. In that case, $\boldsymbol{X}$ is more concordant than $\boldsymbol{X}^*$ if and only if the corresponding copulas satisfy $\mathcal{C}_{\boldsymbol{X}}(u, v) \geqslant \mathcal{C}_{\boldsymbol{X}^*}(u, v)$ for all $(u, v) \in [0, 1]^2$. Moreover, we have the following:

Definition 1 (*Scarsini [20]*). Let $\mathcal{L}(\Omega)$ denote a set of all real-valued continuous random variables on some probability space $(\Omega, \mathcal{A}, \mathrm{P})$. Consider $\varrho : \mathcal{L}(\Omega) \times \mathcal{L}(\Omega) \to \mathbb{R}$ that satisfies the following set of axioms:

A1. (symmetry) $\varrho(X_1, X_2) = \varrho(X_2, X_1)$,
A2. (normalization) $-1 \leqslant \varrho(X_1, X_2) \leqslant 1$,
A3. (independence) $\varrho(X_1, X_2) = 0$ if X_1 and X_2 are independent,
A4. (bounds) $\varrho(X_1, X_2) = 1$ if $X_2 = f(X_1)$ a.s. for f strictly increasing on the range of X_1 and $\varrho(X_1, X_2) = -1$ if $X_2 = f(X_1)$ a.s. for f strictly decreasing on the range of X_1,
A5. (change of sign) If T is strictly monotone on the range of X_1, then

$$\varrho(T(X_1), X_2) = \begin{cases} \varrho(X_1, X_2) & \text{if } T \text{ increasing}, \\ -\varrho(X_1, X_2) & \text{if } T \text{ decreasing}, \end{cases}$$

A6. (continuity) If (X_1^n, X_2^n) are pairs of (continuous) random variables converging in law to (X_1, X_2) (X_1 and X_2 being continuous), then

$$\lim_{n\to\infty} \varrho(X_1^n, X_2^n) = \varrho(X_1, X_2),$$

A7. (coherence) If (X_1^*, X_2^*) is more concordant than (X_1, X_2), then

$$\varrho(X_1^*, X_2^*) \geqslant \varrho(X_1, X_2),$$

then ϱ is called a *measure of concordance*.

Note that if ϱ depends only on the copula, i.e. if $\varrho(X_1, X_2) = \varrho(\mathcal{C}_{\boldsymbol{X}})$, the above axioms can be expressed solely in terms of copulas as well.

One way to obtain concordance measures satisfying Scarsini's definition is to use the so-called *concordance function*. Let $\boldsymbol{X} := (X_1, X_2)$ and $\boldsymbol{Y} := (Y_1, Y_2)$ be independent random vectors with common (arbitrary) marginals. The concordance function Q of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is given by

$$Q(\boldsymbol{X}, \boldsymbol{Y}) := \mathsf{P}[(X_1 - Y_1)(X_2 - Y_2) > 0] - \mathsf{P}[(X_1 - Y_1)(X_2 - Y_2) < 0]. \tag{3}$$

Note that Q is simply the difference between the probabilities of concordance and discordance of (X_1, X_2) and (Y_1, Y_2).

Measures of concordance between X_1 and X_2 now result from a suitable choice of the dependence structure of $\boldsymbol{Y}$. If $\boldsymbol{Y}$ is an independent copy of $\boldsymbol{X}$ then $Q(\boldsymbol{X}, \boldsymbol{Y})$ is Kendall's tau for X_1 and X_2. If, on the other hand, Y_1 and Y_2 are independent, $3Q(\boldsymbol{X}, \boldsymbol{Y})$ yields Spearman's rho for X_1 and X_2; for proofs and further details, see Nelsen [18, Section 5.1].

If the (common) marginals of $\boldsymbol{X}$ and $\boldsymbol{Y}$ have continuous distribution functions, then the concordance function solely depends on the corresponding copulas, see Nelsen [18]:

$$Q(\boldsymbol{X}, \boldsymbol{Y}) = 4\int \mathcal{C}_{\boldsymbol{X}}(u, v)\, d\mathcal{C}_{\boldsymbol{Y}}(u, v) - 1 = 4\int \mathcal{C}_{\boldsymbol{Y}}(u, v)\, d\mathcal{C}_{\boldsymbol{X}}(u, v) - 1. \tag{4}$$

For Kendall's tau and Spearman's rho it hence follows that, with the above specified choices of $\boldsymbol{Y}$,

$$\tau(X_1, X_2) := \tau(\mathcal{C}_{\boldsymbol{X}}) = Q(\boldsymbol{X}, \boldsymbol{Y}) = 4\int \mathcal{C}_{\boldsymbol{X}}(u, v)\, d\mathcal{C}_{\boldsymbol{X}}(u, v) - 1, \tag{5}$$

$$\rho(X_1, X_2) := \rho(\mathcal{C}_{\boldsymbol{X}}) = 3Q(\boldsymbol{X}, \boldsymbol{Y}) = 12\int \mathcal{C}_{\boldsymbol{X}}(u, v)\, du\, dv - 3. \tag{6}$$

One can verify that both τ and ρ fulfill Scarsini's definition; see Scarsini [20].

In case the marginals of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not continuous, neither the above axiomatic Definition 1 nor the use of the concordance function is clear. A first approach would be to simply substitute the unique copulas on the right hand side of (4) by some possible copulas. This technique is, however, questionable for two reasons. First, the result is neither the same for all possible copulas nor equal to Q as given by (3) in general. Secondly, even if (4) is fulfilled, Q does not yield a measure of concordance in the sense of Scarsini's definition. The major difficulty is normalization as this cannot be done without involving the marginals. These issues are illustrated in the following example.

Example 2. Let $\boldsymbol{X} := (X_1, X_2)$ and $\boldsymbol{Y} := (Y_1, Y_2)$ be independent random vectors with common Bernoulli marginals with probabilities of success equal to p and q, respectively. Note in particular that in this case, $F_{X_1}(0) = 1 - p$ and $F_{X_2}(0) = 1 - q$. In the interior of the unit square, the value of $\mathcal{C}_{\boldsymbol{X}}$ and $\mathcal{C}_{\boldsymbol{Y}}$ is uniquely determined solely in the point $(1 - p, 1 - q)$, where it must be equal to the corresponding distribution function evaluated at the point $(0, 0)$.

In order to illustrate the fallacies mentioned above, we assume for the sake of simplicity that X_1 and X_2 are independent. The choices of the dependence structure of $\boldsymbol{Y}$ that correspond to Kendall's tau and Spearman's rho then both reduce to the case of $\boldsymbol{Y}$ being an independent copy of $\boldsymbol{X}$.

The first fallacy is that the right hand side of (4) does not have the same value for all possible copulas. To see this, we consider two different possible copulas for $\boldsymbol{X}$ and $\boldsymbol{Y}$. One such copula

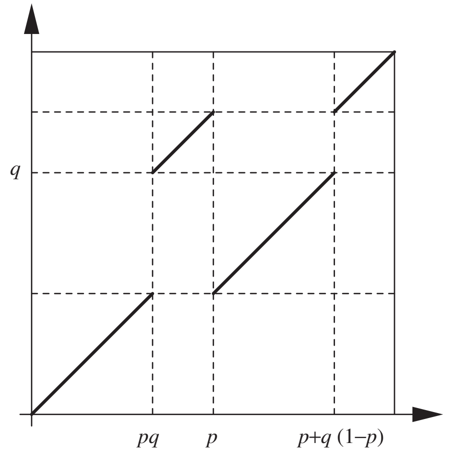


Fig. 1. Support of $\mathcal{C}_U$.

certainly is the independence copula Π given by $\Pi(u, v) = uv$ for any (u, v) in $[0, 1]^2$. On substituting $\mathcal{C}_X$ and $\mathcal{C}_Y$ on the right hand side of (4) by Π we obtain that

$$4 \int uv \, du \, dv - 1 = 0.$$

Secondly, consider for example the singular copula $\mathcal{C}_U$ with support consisting of the line segments described in Fig. 1. One can verify (see [19]) that $\mathcal{C}_U$ is a possible copula of $\boldsymbol{X}$ (and hence also of $\boldsymbol{Y}$) and further that

$$4 \int \mathcal{C}_U(u, v) \, d\mathcal{C}_U(u, v) - 1 = 1 - 4pq(1-p)(1-q),$$
$$4 \int \mathcal{C}_U(u, v) \, du \, dv - 1 = \tfrac{1}{3} - 2pq(1-p)(1-q)(p+q-2pq).$$

Neither of the expressions on the right side is equal to zero in general.

These results immediately imply that (4) does not hold. In the case of Bernoulli marginals, one can easily verify that the concordance function as given by (3) has a particularly simple form:

$$Q(\boldsymbol{X}, \boldsymbol{Y}) = \mathcal{C}_X(1-p, 1-q) + \mathcal{C}_Y(1-p, 1-q) - 2(1-p)(1-q). \tag{7}$$

For X_1 and X_2 independent this means that $Q(\boldsymbol{X}, \boldsymbol{Y}) = 0$. Hence, Eq. (4) holds for the independence copula, but no longer for the copula $\mathcal{C}_U$ defined above.

The third fallacy is that $Q(\boldsymbol{X}, \boldsymbol{Y})$ is not a measure of concordance in the sense of Scarsini's definition, because it violates the normalization axiom A2. To highlight this, note that by (7) and the Fréchet–Hoeffding inequality (see [18]) it follows that

$$2\max(p+q-1, 0) - 2pq \leqslant Q(\boldsymbol{X}, \boldsymbol{Y}) \leqslant 2\min(p, q) - 2pq \quad \text{for } \mathcal{C}_{\boldsymbol{Y}} = \mathcal{C}_{\boldsymbol{X}},$$
$$\max(p+q-1, 0) - pq \leqslant Q(\boldsymbol{X}, \boldsymbol{Y}) \leqslant \min(p, q) - pq \quad \text{for } \mathcal{C}_{\boldsymbol{Y}} = \Pi.$$

Neither the lower bounds are equal to -1 nor the upper ones to 1 in general. Furthermore, the bounds depend on the marginal parameters and have unequal absolute values.

The issue therefore becomes to first determine the interrelationship between the concordance function and possible copulas and second to construct measures of concordance (especially Kendall's tau and Spearman's rho) based upon Q. As will be shown in the next section, there exists a comparatively easy solution to the first question. The second task, however, is far more involved as there seem to exist several ways of accomplishing it. It also requires a slight modification of the above axiomatic definition of concordance measures that seems more suitable for the general non-continuous case. This will be the subject of Section 4.

3. The main technique

The major difficulties that arise when allowing for non-continuous distribution functions are caused by the fact that the transformed variable $F_X(X)$ is no longer uniform. A natural approach that overcomes this has also recently been considered by Denuit and Lambert [3] and Mesfioui and Tajar [16] for purely discrete random variables on a subset of $\mathbb{N}$. These authors construct a continuous "extension" of X by adding an independent and continuous random variable taking values in (0, 1). In the present paper, we follow a somewhat different route. One may also look at alternative transformations that would lead to the uniform distribution. In Section 3.1 we consider one such transformation used in simple hypothesis testing; see Ferguson [5]. This technique indeed enables to obtain several important results concerning the concordance function.

3.1. The univariate case

Suppose U is a uniform random variable independent of X, both U and X defined on some common probability space $(\Omega, \mathcal{A}, \mathsf{P})$, and consider the transformation $\psi : [-\infty, \infty] \times [0, 1] \to [0, 1]$ given by

$$\psi(x, u) = \mathsf{P}[X < x] + u\mathsf{P}[X = x] = F_X(x-) + u\Delta F_X(x), \tag{8}$$

with $\Delta F_X(x) = F_X(x) - F_X(x-)$. We further equip the space $[-\infty, \infty] \times [0, 1]$ with the following (lexicographical) order:

$$(x, u) \preceq (x^*, u^*) \quad \Leftrightarrow \quad (x < x^*) \vee (x = x^* \wedge u \leqslant u^*), \tag{9}$$

where "$\vee$" and "$\wedge$" denote the logical operations "or" and "and", respectively. With respect to this order, ψ becomes nondecreasing. Note that ψ is surjective, but not necessarily injective. A key result is now the following lemma:

Lemma 3. *Under the above assumptions, the random variable $\psi(X, U)$ is uniformly distributed. Moreover,*

$$\mathsf{P}[\psi(X, U) \leqslant \psi(x, u)] = \mathsf{P}[(X, U) \preceq (x, u)]. \tag{10}$$

Proof. Let $w \in [0, 1]$ and set

$$x(w) := F_X^{(-1)}(w+) \quad \text{and} \quad u(w) = \begin{cases} 1 & \text{if } \mathsf{P}[X = x(w)] = 0, \\ \dfrac{w - \mathsf{P}[X < x(w)]}{\mathsf{P}[X = x(w)]} & \text{otherwise}, \end{cases} \tag{11}$$

where $F_X^{(-1)}(\cdot+)$ denotes the right hand side limit of the generalized inverse of F_X given by $F_X^{-1}(u) := \inf\{x \in \mathbb{R} | F_X(x) \geqslant u\}$. One can verify that $\psi(x(w), u(w)) = w$ and that $\{\omega \in \Omega | \psi(X(\omega), U(\omega)) \leqslant w\} = \{\omega \in \Omega | (X(\omega), U(\omega)) \preceq (x(w), u(w))\}$. Consequently,

$$\begin{aligned} \mathsf{P}[\psi(X,U) \leqslant w] &= \mathsf{P}[(X,U) \preceq (x(w),u(w))] \\ &= \mathsf{P}[X < x(w)] + \underbrace{\mathsf{P}[U \leqslant u(w)]}_{=u(w)} \mathsf{P}[X = x(w)] \\ &= \psi(x(w), u(w)) = w. \end{aligned}$$

As $\mathsf{P}[(X,U) \preceq (x,u)] = \psi(x,u)$, (10) is straightforward. $\square$

In the subsequent discussions, the following identity will come in useful. From the uniformity of $\psi(X,U)$ we have that

$$\begin{aligned} \tfrac{1}{2} &= \mathrm{E}\psi(X,U) = \mathrm{E}(\mathrm{E}\psi(X,U)|X) \\ &= \mathrm{E}(F_X(X-) + \tfrac{1}{2}(F_X(X) - F_X(X-))), \end{aligned} \tag{12}$$

which in particular implies that

$$\mathrm{E}\left(\frac{(F_X(X) + F_X(X-))}{2}\right) = \frac{1}{2}. \tag{13}$$

Note 1. For a random variable X taking values solely in $\mathbb{N}$, an alternative transformation can be derived using the continuous extension method proposed by [3]. Because the continuous extension $X^* := X + (U-1)$ is a continuous random variable, $F_{X^*}(X + (U-1))$ is uniformly distributed. On using the expression for F_{X^*} derived in [3] it follows that $F_{X^*}(X + (U-1)) = F_X(X-1) + F_U(U)\Delta F_X(X)$. For U uniform this is precisely the above transformation $\psi(X,U)$.

3.2. Multivariate generalizations

In order to avoid notationally complex proofs, discussions below are restricted to the bivariate case only. Note, however, that all techniques that make sense in higher dimensions can indeed be generalized. As before, let $\boldsymbol{X} = (X_1, X_2)$ denote a random vector with arbitrary marginals and $\boldsymbol{U} = (U_1, U_2)$ a random vector with uniform marginals independent of $\boldsymbol{X}$. To this point, no restrictions on the dependence structure of $\boldsymbol{U}$ are necessary, in particular we do not need to assume that U_1 and U_2 are independent. Recall that according to Lemma 3, the componentwise transformed random vector $\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U}) := \big(\psi(X_1,U_1), \psi(X_2,U_2)\big)$ has uniform marginals. The following proposition now emphasizes that the dependence structures of $\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U})$ and $\boldsymbol{X}$ are closely related.

Proposition 4. *For any dependence structure of* $\boldsymbol{U}$, *the unique copula* $C_{\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U})}$ *of* $\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U})$ *is a possible copula of* $\boldsymbol{X}$. *Furthermore, if* U_1 *and* U_2 *are independent then* $C_{\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U})}$ *is the standard extension copula of Schweizer and Sklar and*

$$\begin{aligned} &\mathsf{P}[\psi(X_1,U_1) \leqslant \psi(x_1,u_1), \psi(X_2,U_2) \leqslant \psi(x_2,u_2)] \\ &\qquad = \mathsf{P}[(X_1,U_1) \preceq (x_1,u_1), (X_2,U_2) \preceq (x_2,u_2)]. \end{aligned} \tag{14}$$

Proof. Along the same lines as in (11) of the proof of Lemma 3, for $(w_1, w_2) \in [0,1]^2$ one can construct the points $(x_1(w_1), u_1(w_1))$ and $(x_2(w_2), u_2(w_2))$ and argue that

$$\begin{aligned}
&\mathsf{P}[\psi(X_1,U_1)\leqslant w_1, \psi(X_2,U_2)\leqslant w_2] \\
&\quad = \mathsf{P}[X_1 < x_1(w_1), X_2 < x_2(w_2)] \\
&\qquad + u_1(w_1)\mathsf{P}[X_1 = x_1(w_1), X_2 < x_2(w_2)] \\
&\qquad + u_2(w_2)\mathsf{P}[X_1 < x_1(w_1), X_2 = x_2(w_2)] \\
&\qquad + \mathcal{C}_{\boldsymbol{U}}(u_1(w_1), u_2(w_2))\mathsf{P}[X_1 = x_1(w_1), X_2 = x_2(w_2)]. \qquad (15)
\end{aligned}$$

Now suppose that $(w_1, w_2) \in \operatorname{ran} F_{X_1} \times \operatorname{ran} F_{X_2}$. In this case, we either have $x_i(w_i) = F_{X_i}^{(-1)}(w_i)$ or $x_i(w_i) > F_{X_i}^{(-1)}(w_i)$ for $i = 1, 2$. The first situation, however, implies $F_{X_i}(x_i(w_i)) = w_i$ and hence $u_i(w_i) = 1$. In the latter case, $\mathsf{P}[F_{X_i}^{(-1)}(w_i) < X_i < x_i(w_i)] = 0$ which yields $\mathsf{P}[X_i < x_i(w_i)] = F_{X_i}(F_{X_i}^{(-1)}(w_i)) = w_i$. Consequently, depending upon whether $\mathsf{P}[X_i = x_i(w_i)] > 0$ or not, $u_i(w_i)$ equals either 0 or 1. In either case, (15) leads to

$$\begin{aligned}
\mathcal{C}_{\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U})}(w_1, w_2) &= \mathsf{P}[\psi(X_1,U_1)\leqslant w_1, \psi(X_2,U_2)\leqslant w_2] \\
&= \mathsf{P}[X_1 \leqslant F_{X_1}^{(-1)}(w_1), X_2 \leqslant F_{X_2}^{(-1)}(w_2)] = \mathcal{C}_{\boldsymbol{X}}(w_1, w_2).
\end{aligned}$$

To prove the second part of the lemma, first note that in case of independent U_i's, (15) can be rewritten as

$$\begin{aligned}
&\mathsf{P}[\psi(X_1,U_1)\leqslant w_1, \psi(X_2,U_2)\leqslant w_2] \\
&\quad = (1-u_1(w_1))(1-u_2(w_2))\mathsf{P}[X_1 < x_1(w_1), X_2 < x_2(w_2)] \\
&\qquad + u_1(w_1)(1-u_2(w_2))\mathsf{P}[X_1 \leqslant x_1(w_1), X_2 < x_2(w_2)] \\
&\qquad + (1-u_1(w_1))u_2(w_2)\mathsf{P}[X_1 < x_1(w_1), X_2 \leqslant x_2(w_2)] \\
&\qquad + u_1(w_1)u_2(w_2)\mathsf{P}[X_1 \leqslant x_1(w_1), X_2 \leqslant x_2(w_2)]. \qquad (16)
\end{aligned}$$

Without loss of generality, assume that $w_i \notin \operatorname{ran} F_{X_i}$ for $i = 1, 2$. One can verify that the least and the greatest element in $\overline{\operatorname{ran} F_{X_i}}$ satisfying $a_i \leqslant w_i \leqslant b_i$ is given by $F_{X_i}(x_i(w_i)-)$ and $F_{X_i}(x_i(w_i))$, respectively. Furthermore, note that in this case λ_i from (2) is hence equal to $u_i(w_i)$. Consequently, from (16) and (2) we have that

$$\begin{aligned}
\mathcal{C}_{\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U})}(w_1, w_2) &= (1-\lambda_1)(1-\lambda_2)\mathcal{C}_{\boldsymbol{X}}\big(F_{X_1}(x_1(w_1)-), F_{X_2}(x_2(w_2)-)\big) \\
&\quad + \lambda_1(1-\lambda_2)\mathcal{C}_{\boldsymbol{X}}\big(F_{X_1}(x_1(w_1)), F_{X_2}(x_2(w_2)-)\big) \\
&\quad + (1-\lambda_1)\lambda_2\mathcal{C}_{\boldsymbol{X}}\big(F_{X_1}(x_1(w_1)-), F_{X_2}(x_2(w_2))\big) \\
&\quad + \lambda_1\lambda_2\mathcal{C}_{\boldsymbol{X}}\big(F_{X_1}(x_1(w_1)), F_{X_2}(x_2(w_2))\big) = \mathcal{C}_{\boldsymbol{X}}^S(w_1, w_2). \qquad (17)
\end{aligned}$$

Finally,

$$\begin{aligned}
&\mathsf{P}[(X_1,U_1) \preceq (x_1,u_1), (X_2,U_2) \preceq (x_2,u_2)] \\
&\quad = (1-u_1)(1-u_2)\mathsf{P}[X_1 < x_1, X_2 < x_2] \\
&\qquad + u_1(1-u_2)\mathsf{P}[X_1 \leqslant x_1, X_2 < x_2] \\
&\qquad + u_2(1-u_1)\mathsf{P}[X_1 < x_1, X_2 \leqslant x_2] + u_1u_2\mathsf{P}[X_1 \leqslant x_1, X_2 \leqslant x_2] \\
&\quad = \mathcal{C}_{\boldsymbol{X}}^S\Big(F_{X_1}(x_1-) + u_1\big(\Delta F_{X_1}(x_1)\big), F_{X_2}(x_2-) + u_2\big(\Delta F_{X_2}(u_2)\big)\Big) \\
&\quad = \mathcal{C}_{\boldsymbol{X}}^S(\psi(x_1,u_1), \psi(x_2,u_2)) = \mathcal{C}_{\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U})}(\psi(x_1,u_1), \psi(x_2,u_2)), \qquad (18)
\end{aligned}$$

which completes the proof. □

Furthermore, one can argue that the concordance function of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is linked to the concordance function of the vectors transformed by $\boldsymbol{\Psi}$, provided, however, that the uniform random vectors used in the transformations have independent marginals.

Theorem 5. *Suppose that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent bivariate random vectors with common marginals and further that $\boldsymbol{U}$ and $\boldsymbol{V}$ are iid bivariate random vectors with independent uniform marginals assumed independent of $\boldsymbol{X}$ and $\boldsymbol{Y}$. The concordance function of $\boldsymbol{X}$ and $\boldsymbol{Y}$ satisfies*

$$Q(\boldsymbol{X},\boldsymbol{Y}) = Q(\boldsymbol{\Psi}(\boldsymbol{X},\boldsymbol{U}),\boldsymbol{\Psi}(\boldsymbol{Y},\boldsymbol{V})). \tag{19}$$

It moreover follows that

$$Q(\boldsymbol{X},\boldsymbol{Y}) = 4\int \mathcal{C}^S_{\boldsymbol{X}}(u,v)\,d\mathcal{C}^S_{\boldsymbol{Y}}(u,v) - 1 = 4\int \mathcal{C}^S_{\boldsymbol{Y}}(u,v)\,d\mathcal{C}^S_{\boldsymbol{X}}(u,v) - 1 \tag{20}$$

and

$$Q(\boldsymbol{X},\boldsymbol{Y}) = \mathrm{E}[F_{\boldsymbol{X}}(Y_1,Y_2) + F_{\boldsymbol{X}}(Y_1-,Y_2) + F_{\boldsymbol{X}}(Y_1,Y_2-) + F_{\boldsymbol{X}}(Y_1-,Y_2-) - 1]. \tag{21}$$

Proof. First recall that $\{(X_1,U_1)\succ(Y_1,V_1)\} = \{X_1>Y_1\}\cup\{X_1=Y_1\}\cap\{U_1>V_1\}$. From this and (14) one argues that

$$\begin{aligned}
&\mathsf{P}\big(\psi(X_1,U_1)>\psi(Y_1,V_1),\psi(X_2,U_2)>\psi(Y_2,V_2)\big)\\
&\quad= \mathsf{P}\big(X_1>Y_1,X_2>Y_2\big) + \mathsf{P}\big(X_1=Y_1,X_2>Y_2\big)\mathsf{P}(U_1>V_1)\\
&\qquad+\mathsf{P}\big(X_1>Y_1,X_2=Y_2\big)\mathsf{P}(U_2>V_2)\\
&\qquad+\mathsf{P}\big(X_1=Y_1,X_2=Y_2\big)\mathsf{P}(U_1>V_1)\mathsf{P}(U_2>V_2)\\
&\quad= \mathsf{P}\big(X_1>Y_1,X_2>Y_2\big) + \tfrac{1}{2}\mathsf{P}\big(X_1=Y_1,X_2>Y_2\big) + \tfrac{1}{2}\mathsf{P}\big(X_1>Y_1,X_2=Y_2\big)\\
&\qquad+\tfrac{1}{4}\mathsf{P}\big(X_1=Y_1,X_2=Y_2\big),
\end{aligned}$$

which yields

$$\begin{aligned}
&\mathsf{P}\big((\psi(X_1,U_1)-\psi(Y_1,V_1))(\psi(X_2,U_2)-\psi(Y_2,V_2))>0\big)\\
&\quad= \mathsf{P}\big((X_1-Y_1)(X_2-Y_2)>0\big) + \tfrac{1}{2}\mathsf{P}\big(X_1=Y_1,X_2\neq Y_2\big)\\
&\qquad+\tfrac{1}{2}\mathsf{P}\big(X_1\neq Y_1,X_2=Y_2\big) + \tfrac{1}{2}\mathsf{P}\big(X_1=Y_1,X_2=Y_2\big)\\
&\quad= \mathsf{P}\big((X_1-Y_1)(X_2-Y_2)>0\big) + \tfrac{1}{2}\mathsf{P}\big(X_1=Y_1\vee X_2=Y_2\big).
\end{aligned}$$

A similar computation yields for the probability of discordance of the transformed vectors $\mathsf{P}\big((X_1-Y_1)(X_2-Y_2)<0\big)+\frac{1}{2}\mathsf{P}\big(X_1=Y_1\vee X_2=Y_2\big)$; hence (19) holds. Expression (20) follows directly by Proposition 4 and (4). Along the same lines as in (18) one finally obtains from (14) that

$$\begin{aligned}
\int \mathcal{C}^S_{\boldsymbol{X}}\,d\mathcal{C}^S_{\boldsymbol{Y}} &= \mathsf{P}\big(\psi(X_1,U_1)\leqslant\psi(Y_1,V_1),\psi(X_2,U_2)\leqslant\psi(Y_2,V_2)\big)\\
&= \mathsf{P}\big((X_1,U_1)\preceq(Y_1,V_1),(X_2,U_2)\preceq(Y_2,V_2)\big)\\
&= \mathrm{E}\Big(\mathrm{E}\big(\mathrm{E}(1_{\{(X_1,U_1)\preceq(Y_1,V_1),(X_2,U_2)\preceq(Y_2,V_2)\}}|\boldsymbol{Y})\big|\boldsymbol{V}\big)\Big)\\
&= \mathrm{E}\Big(\mathrm{E}\big((1-V_1)(1-V_2)F_{\boldsymbol{X}}(Y_1-,Y_2-) + V_1(1-V_2)F_{\boldsymbol{X}}(Y_1,Y_2-)\\
&\qquad+(1-V_1)V_2F_{\boldsymbol{X}}(Y_1-,Y_2) + V_1V_2F_{\boldsymbol{X}}(Y_1,Y_2)\big|\boldsymbol{Y}\big)\Big)\\
&= \tfrac{1}{4}\mathrm{E}\big(F_{\boldsymbol{X}}(Y_1-,Y_2-) + F_{\boldsymbol{X}}(Y_1,Y_2-) + F_{\boldsymbol{X}}(Y_1-,Y_2) + F_{\boldsymbol{X}}(Y_1,Y_2)\big),
\end{aligned} \tag{22}$$

which yields identity (21). □

Note 2. Identity (21) is rather technical; it allows in particular for a comparatively quick calculation of $Q(\boldsymbol{X}, \boldsymbol{Y})$ for purely discrete marginals. It is interesting to note, however, that in the continuous case, $\int \mathcal{C}_{\boldsymbol{X}}\, d\mathcal{C}_{\boldsymbol{Y}}$ equals $\mathrm{E}(F_{\boldsymbol{X}}(Y_1, Y_2))$, see Nelsen [18]. The right hand side of (21) hence compresses into $\mathrm{E}[4F_{\boldsymbol{X}}(Y_1, Y_2) - 1]$. As soon as the marginal distribution functions have discontinuities, $\int \mathcal{C}^S_{\boldsymbol{X}}\, d\mathcal{C}^S_{\boldsymbol{Y}}$ is rather replaced by the mean of the one sided limits, $\frac{1}{4}\mathrm{E}\big(F_{\boldsymbol{X}}(Y_1-, Y_2-) + F_{\boldsymbol{X}}(Y_1, Y_2-) + F_{\boldsymbol{X}}(Y_1-, Y_2) + F_{\boldsymbol{X}}(Y_1, Y_2)\big)$, as shown in (22).

4. Measures of concordance

4.1. Scarsini's definition revisited

Motivated by Theorem 5, we now study the standard extension copula more carefully. First, the interpretation given in Proposition 4 provides an elegant tool through which several important properties of the standard extension copula can be obtained. Some of these generalize the properties of the unique copula corresponding to distributions with continuous marginals:

Corollary 6. *For a random vector $\boldsymbol{X}$ with arbitrary marginals the following results hold*:
(1) *A random vector $\boldsymbol{X}^*$ having the same marginals as $\boldsymbol{X}$ is more concordant than $\boldsymbol{X}$ if and only if $\mathcal{C}^S_{\boldsymbol{X}}(u, v) \leqslant \mathcal{C}^S_{\boldsymbol{X}^*}(u, v)$ for any $(u, v) \in [0, 1]^2$.*
(2) *If T is strictly increasing and continuous on* ran X_1, *the standard extension copulas of $\boldsymbol{X}$ and $(T(X_1), X_2)$ are the same.*
(3) *If T is strictly decreasing and continuous on* ran X_1, *the standard extension copula of $(T(X_1), X_2)$ is given by $v - \mathcal{C}^S_{\boldsymbol{X}}(1-u, v)$ for any $(u, v) \in [0, 1]^2$.*

Proof. Throughout the proof, assume that the vectors $\boldsymbol{U}$ and $\boldsymbol{U}^*$, which will be used for the transformation $\boldsymbol{\Psi}$ from Section 3.2, are iid with uniform and independent marginals and are moreover independent of $\boldsymbol{X}$ and $\boldsymbol{X}^*$, respectively. Statement (1) follows from the fact that $\boldsymbol{X}^*$ is more concordant than $\boldsymbol{X}$ if and only if the transformed vector $\boldsymbol{\Psi}(\boldsymbol{X}^*, \boldsymbol{U}^*)$ is more concordant than $\boldsymbol{\Psi}(\boldsymbol{X}, \boldsymbol{U})$. The "if" part follows from (18) by setting $u_1 = u_2 = 1$; the "only if" part can be verified by combining (18) and (14). For (2) and (3), first note that the distribution function of $T(X_1)$ is given by $F_{T(X_1)}(x) = F_{X_1}(T^{-1}(x))$ if T is increasing and by $F_{T(X_1)}(x) = 1 - F_{X_1}(T^{-1}(x)-)$ if T is decreasing. One easily shows that this leads to $\psi(T(X_1), U_1) = \psi(X_1, U_1)$ in the former case and to $\psi(T(X_1), U_1) = 1 - \psi(X_1, 1 - U_1)$ in the latter. From this it immediately follows that the copulas corresponding to $(\psi(X_1, U_1), \psi(X_2, U_2))$ and $(\psi(T(X_1), U_1), \psi(X_2, U_2))$ are equal if T is increasing. Otherwise, observe first that the copula of $(\psi(X_1, U_1), \psi(X_2, U_2))$ coincides with the copula of $(\psi(X_1, 1 - U_1), \psi(X_2, U_2))$ as $1 - U_1$ is uniformly distributed and independent of U_2. The well known result by Schweizer and Wolff [22], concerning the changes the unique copula in the continuous case undergoes under strictly monotone transformations of the marginals, finally yields that the copula of $(1 - \psi(X_1, 1 - U_1), \psi(X_2, U_2))$ is indeed given by $v - \mathcal{C}^S_{\boldsymbol{X}}(1 - u, v)$. $\square$

The standard extension copula, however, does not share all properties with the unique copula corresponding to a distribution function with continuous marginals. One such issue is weak convergence. The standard extension copulas corresponding to a weak convergent sequence of random vectors $\boldsymbol{X}_n$ do not necessarily converge pointwise on the entire unit square if the limiting vector has non-continuous marginals. In general, the pointwise convergence holds only on the product of the ranges of the limiting marginal distribution functions; see Lindner and Szimayer

[13]. The convergence on the entire unit square can be achieved only in special cases, one such being that the supports of the marginal distribution functions do not change with n; for details see Nešlehová [19].

Another issue is that though the standard extension copula coincides with the independence copula if the marginals are independent, it is always different from the Fréchet–Hoeffding bounds, even if the marginals are countermonotonic and comonotonic, respectively (i.e. when the upper, respectively lower, Fréchet bound is a possible copula of $\boldsymbol{X}$). This is due to the fact that as soon as the closure of the product of the ranges of the marginal distribution functions does not fill out the entire unit square, the standard extension copula cannot be singular. Although the standard extension copula is bounded from below and above by standard extension copulas corresponding to the perfect monotonic case, these bounds are not simply related, unless, according to Corollary 6, the monotonic functions are continuous.

Corollary 6 and the fact that the concordance function depends solely on the standard extension copula now motivate a generalization of Scarsini's definition of concordance and concordance measures to the case of non-continuous random variables. According to Corollary 6, provided $\boldsymbol{X}^*$ and $\boldsymbol{X}$ have common marginals, $\boldsymbol{X}^*$ is more concordant than $\boldsymbol{X}$ if and only if the corresponding standard extension copulas satisfy $\mathcal{C}^S_{\boldsymbol{X}}(u, v) \leqslant \mathcal{C}^S_{\boldsymbol{X}^*}(u, v)$. By the above discussions, however, it seems meaningful to change two of Scarsini's axioms as follows:

A4*. (bounds) $\varrho(X_1, X_2) = 1$ if $X_2 = f(X_1)$ a.s. for a strictly increasing and continuous function f on the range of X_1 and $\varrho(X_1, X_2) = -1$ if $X_2 = f(X_1)$ a.s. for a strictly decreasing and continuous function f on the range of X_1.

A5*. (change of sign) If T is strictly monotone and continuous on the range of X_1, then

$$\varrho(T(X_1), X_2) = \begin{cases} \varrho(X_1, X_2) & \text{if } T \text{ increasing,} \\ -\varrho(X_1, X_2) & \text{if } T \text{ decreasing.} \end{cases}$$

The sixth axiom, however, still remains somewhat questionable as it is not satisfied by the most interesting concordance measures, unless restrictions on the distribution functions of the weak convergent sequence are made; see Nešlehová [19].

4.2. Kendall's tau

In the case of continuous marginals Kendall's tau is defined as the concordance function between the random vector $\boldsymbol{X}$ and an independent copy $\boldsymbol{Y}$, say. If the marginals are not necessarily continuous, the concordance function equals Kendall's tau of the corresponding standard extension copula, $\tau(\mathcal{C}^S_{\boldsymbol{X}})$, i.e.

$$Q(\boldsymbol{X}, \boldsymbol{Y}) = \tau(\mathcal{C}^S_{\boldsymbol{X}}) = 4 \int \mathcal{C}^S_{\boldsymbol{X}}(u, v)\, d\mathcal{C}^S_{\boldsymbol{X}}(u, v) - 1.$$

The question remains, however, as to whether $\tau(\mathcal{C}^S_{\boldsymbol{X}})$ fulfills the (Scarsini) axioms as amended above. The key issue again becomes the investigation of the relationship between the dependence structure of $\boldsymbol{X}$ and the dependence structure of the transformed vector $\boldsymbol{\Psi}(\boldsymbol{X}, \boldsymbol{U})$ represented by the standard extension copula. If the marginals of $\boldsymbol{X}$ are independent, the same is true for the marginals of $\boldsymbol{\Psi}(\boldsymbol{X}, \boldsymbol{U})$ and as a consequence, $\tau(\mathcal{C}^S_{\boldsymbol{X}})$ equals zero. On the other hand, however, the above concordance function generally cannot reach the bounds ± 1 if the marginals are comonotonic and countermonotonic, respectively. If we denote by $\mathcal{M}^S_{\boldsymbol{X}}$ and $\mathcal{W}^S_{\boldsymbol{X}}$ the standard extension copulas

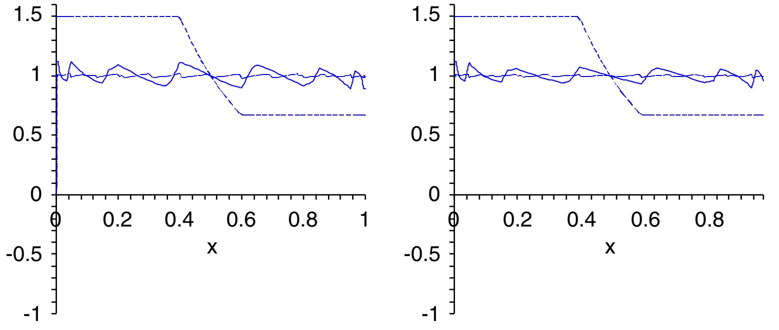


Fig. 2. $|\tau(\mathcal{M}_{\boldsymbol{X}}^S)/\tau(\mathcal{W}_{\boldsymbol{X}}^S)|$ (left) and $|\rho(\mathcal{M}_{\boldsymbol{X}}^S)/\rho(\mathcal{W}_{\boldsymbol{X}}^S)|$ (right) for binomial distributions $F_{X_1} = \mathcal{B}(n, 0.4)$ and $F_{X_2} = \mathcal{B}(n, x)$ with $n = 1$ (dashed line), 4 (solid line) and 10 (dotted line) as functions of the parameter x of F_{X_2} taking values in $(0, 1)$.

corresponding to comonotonic and countermonotonic marginals, respectively, we have that

$$\tau(\mathcal{W}_{\boldsymbol{X}}^S) \leqslant \tau(\mathcal{C}_{\boldsymbol{X}}^S) \leqslant \tau(\mathcal{M}_{\boldsymbol{X}}^S),$$

but $|\tau(\mathcal{W}_{\boldsymbol{X}}^S)| \neq |\tau(\mathcal{M}_{\boldsymbol{X}}^S)|$ in general. The relationship between these bounds is rather complex as illustrated by Fig. 2. It is, however, possible to bound the concordance function by less sharp bounds that are far more easy to handle.

Corollary 7. *Under the hypotheses of Theorem* 5 *it follows that*

$$|Q(\boldsymbol{X}, \boldsymbol{Y})| \leqslant \sqrt{\big(1 - \mathrm{E}(\Delta F_{X_1}(X_1))\big)\big(1 - \mathrm{E}(\Delta F_{X_2}(X_2))\big)}. \tag{23}$$

Proof. The difference between the probabilities of concordance and discordance can also be rewritten as

$$\begin{aligned} Q = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} & 1_{\{(-\infty, y_1) \times (-\infty, y_2)\}}(x_1, x_2) + 1_{\{(y_1, \infty) \times (y_2, \infty)\}}(x_1, x_2) \\ & - 1_{\{(-\infty, y_1) \times (y_2, \infty)\}}(x_1, x_2) - 1_{\{(y_1, \infty) \times (-\infty, y_2)\}}(x_1, x_2)\, dF_{\boldsymbol{X}}(x_1, x_2)\, dF_{\boldsymbol{Y}}(y_1, y_2). \end{aligned} \tag{24}$$

Consider now the function $g : \mathbb{R}^2 \to \{-1, 0, 1\}$ given by $g(a, b) = \operatorname{sign}(b - a)$ if $a \neq b$ and by zero otherwise. With this notation, Q equals

$$Q = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} g(x_1, y_1) g(x_2, y_2)\, dF_{\boldsymbol{X}}(x_1, x_2)\, dF_{\boldsymbol{Y}}(y_1, y_2).$$

By Hölder's inequality, it first follows that

$$\begin{aligned} |Q| \leqslant & \sqrt{\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} g^2(x_1, y_1)\, dF_{\boldsymbol{X}}(x_1, x_2)\, dF_{\boldsymbol{Y}}(y_1, y_2)} \\ & \times \sqrt{\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} g^2(x_2, y_2)\, dF_{\boldsymbol{X}}(x_1, x_2)\, dF_{\boldsymbol{Y}}(y_1, y_2)}. \end{aligned}$$

Furthermore, the right hand side can be simplified as follows:

$$\begin{aligned}
&\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} g^2(x_1, y_1)\, dF_{\boldsymbol{X}}(x_1, x_2)\, dF_{\boldsymbol{Y}}(y_1, y_2)\\
&\quad = \int_{\mathbb{R}^2} g^2(x_1, y_1)\, dF_{X_1}(x_1)\, dF_{X_1}(y_1)\\
&\quad = \int_{\mathbb{R}^2} 1_{(-\infty, y_1)}(x_1)\, dF_{X_1}(x_1)\, dF_{X_1}(y_1) + \int_{\mathbb{R}^2} 1_{(y_1,\infty)}(x_1)\, dF_{X_1}(x_1)\, dF_{X_1}(y_1)\\
&\quad = \int_{\mathbb{R}} F_{X_1}(y_1-)\, dF_{X_1}(y_1) + \int_{\mathbb{R}^2} 1_{(-\infty, x_1)}(y_1)\, dF_{X_1}(y_1)\, dF_{X_1}(x_1)\\
&\quad = 2\mathrm{E}(F_{X_1}(X_1-)) \overset{(12)}{=} 1 - \mathrm{E}(\Delta F_{X_1}(X_1)).
\end{aligned}$$

By symmetry it follows that

$$\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} g^2(x_2, y_2)\, dF_{\boldsymbol{X}}(x_1, x_2)\, dF_{\boldsymbol{Y}}(y_1, y_2) = 1 - \mathrm{E}(\Delta F_{X_2}(X_2)),$$

which completes the proof. □

The following proposition states that the right hand side in (23) corresponds to the sharper kind of monotonicity required in A4*.

Proposition 8. *Assume that the marginals of* $\boldsymbol{X}$ *satisfy* $X_2 = T(X_1)$ *a.s. for some strictly monotone and continuous transformation* T *on* ran X_1. *Then*

$$\tau(\mathcal{C}^S_{\boldsymbol{X}}) = \begin{cases} 1 - \mathrm{E}(\Delta F_{X_1}(X_1)) & \text{for } T \text{ increasing},\\ -1 + \mathrm{E}(\Delta F_{X_1}(X_1)) & \text{for } T \text{ decreasing}. \end{cases} \tag{25}$$

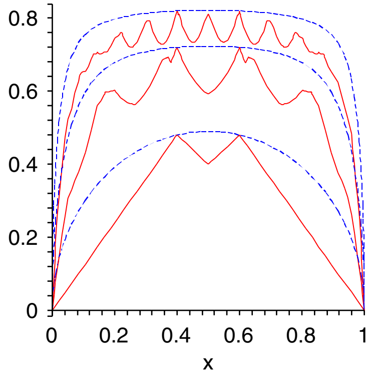
Proof. Assume first that T is increasing. In that case $F_{X_2}(x) = F_{X_1}(T^{-1}(x))$ and (21) yields

$$\begin{aligned}
\tau(\mathcal{C}^S_{\boldsymbol{X}}) &= \mathrm{E}[\min(F_{X_1}(X_1), F_{X_1}(X_1)) + \min(F_{X_1}(X_1-), F_{X_1}(X_1))\\
&\quad + \min(F_{X_1}(X_1), F_{X_1}(X_1-)) + \min(F_{X_1}(X_1-), F_{X_1}(X_1-)) - 1]\\
&= \mathrm{E}[F_{X_1}(X_1) + 3F_{X_1}(X_1-) - 1]\\
&= 4\mathrm{E}F_{X_1}(X_1-) + \mathrm{E}[(\Delta F_{X_1}(X_1)) - 1] \overset{(12)}{=} 1 - \mathrm{E}(\Delta F_{X_1}(X_1)).
\end{aligned}$$

For T decreasing we have that $F_{X_2}(x) = 1 - F_{X_1}(T^{-1}(x)-)$, which leads to

$$\begin{aligned}
\tau(\mathcal{C}^S_{\boldsymbol{X}}) &= \mathrm{E}[\max(F_{X_1}(X_1) - F_{X_1}(X_1-), 0) + \max(F_{X_1}(X_1-) - F_{X_1}(X_1-), 0)\\
&\quad + \max(F_{X_1}(X_1) - F_{X_1}(X_1), 0) + \max(F_{X_1}(X_1-) - F_{X_1}(X_1), 0) - 1]\\
&= \mathrm{E}(\Delta F_{X_1}(X_1)) - 1. \qquad \square
\end{aligned}$$

The less sharp bounds $\sqrt{\left(1 - E(\Delta F_{X_1}(X_1))\right)\left(1 - E(\Delta F_{X_2}(X_2))\right)}$ are compared with the sharper ones given by $\max(|\tau(\mathcal{M}^S_X)|, |\tau(\mathcal{W}^S_X)|)$ in Fig. 3. With Proposition 8, it is now natural to generalize Kendall's tau for non-continuous random variables in the following way.



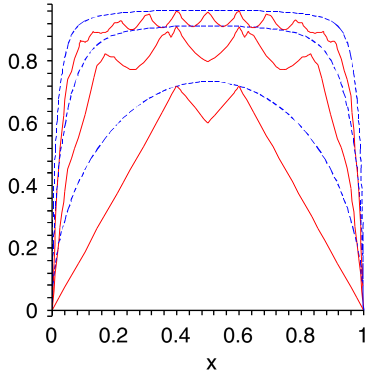


Fig. 3. $\max(|\tau(\mathcal{M}^s_X)|, |\tau(\mathcal{W}^s_X)|)$ (solid line) and the less sharp bounds (dashed line) for Kendall's tau (left) and $\max(|\rho(\mathcal{M}^s_X)|, |\rho(\mathcal{W}^s_X)|)$ (solid line) and the less sharp bounds (dashed line) for Spearman's rho (right) for binomial distributions $F_{X_1} = \mathcal{B}(n, 0.4)$ and $F_{X_2} = \mathcal{B}(n, x)$ with $n = 1$ (bottom curves), 4 (middle curves) and 10 (top curves), again as functions of the parameter x of F_{X_2} taking values in (0, 1).

Definition 9. Let $\boldsymbol{X} = (X_1, X_2)$ be a bivariate random vector with arbitrary marginals, then the *non-continuous version of Kendall's tau* is given by

$$\tau(X_1, X_2) = \frac{4\int \mathcal{C}^S_{\boldsymbol{X}}\, d\mathcal{C}^S_{\boldsymbol{X}} - 1}{\sqrt{(1 - \mathrm{E}(\Delta F_{X_1}(X_1))(1 - \mathrm{E}(\Delta F_{X_2}(X_2))}}. \tag{26}$$

This quantity has many properties similar to those of Kendall's tau for distributions with continuous marginals. In fact, one can verify by Corollary 6 and the properties of Kendall's tau for continuous distributions that τ satisfies A1–A3, A4* and A5* as well as A7. Note, however, that for a weak convergent sequence $\{\boldsymbol{X}_n\}$, the convergence of the corresponding sequence of Kendall's taus generally fails; see Nešlehová [19] for details.

The normalization used in Definition 9 is certainly not the only possible; see Denuit and Lambert [3] who obtain another version of Kendall's tau by an alternative normalization of $\tau(\mathcal{C}^S_{\boldsymbol{X}})$. The above generalization of Kendall's tau is, however, supported by the fact that it coincides with the known population version of this measure for empirical distributions as will be discussed later in Section 5.

4.3. Spearman's rho

In order to obtain a generalization of Spearman's rho one can proceed similarly as with Kendall's tau.

If the marginals of $\boldsymbol{Y}$ are independent copies of the marginals of $\boldsymbol{X}$, we obtain

$$\rho(\mathcal{C}^S_{\boldsymbol{X}}) = 3Q(\boldsymbol{X}, \boldsymbol{Y}) = 12\int \mathcal{C}^S_{\boldsymbol{X}}(u, v)\, du\, dv - 3,$$

which is equal to Spearman's rho of the transformed vector $\boldsymbol{\Psi}(\boldsymbol{X}, \boldsymbol{U})$. A difficulty is again caused by the fact that the marginals of $\boldsymbol{\Psi}(\boldsymbol{X}, \boldsymbol{U})$ cannot be perfectly monotonic dependent. In analogy to $\tau(\mathcal{C}^S_{\boldsymbol{X}})$, we have that $\rho(\mathcal{W}^S_{\boldsymbol{X}}) \leqslant \rho(\mathcal{C}^S_{\boldsymbol{X}}) \leqslant \rho(\mathcal{M}^S_{\boldsymbol{X}})$ but $|\rho(\mathcal{W}^S_{\boldsymbol{X}})| \neq |\rho(\mathcal{M}^S_{\boldsymbol{X}})|$ in general. The interrelationship between the bounds is complicated and not easily evaluated analytically; see Fig. 2

for an illustration. In order to find a suitable normalization that would be comparatively easy to handle, one can, however, generalize the approach chosen by Hoeffding [7]. First note that $\rho(\mathcal{C}^S_{\boldsymbol{X}})$ is also equal to the (linear) correlation coefficient of $\psi(X_1, U_1)$ and $\psi(X_2, U_2)$. Furthermore,

$$\begin{aligned}
\mathrm{E}(\psi(X_1, U_1)\psi(X_2, U_2)) &= \mathrm{E}(\mathrm{E}(\psi(X_1, U_1)\psi(X_2, U_2)|\boldsymbol{X}))\\
&= \mathrm{E}\big(F_{X_1}(X_1-)F_{X_2}(X_2-)\\
&\quad +\tfrac{1}{2}F_{X_2}(X_2-)(F_{X_1}(X_1) - F_{X_1}(X_1-))\\
&\quad +\tfrac{1}{2}F_{X_1}(X_1-)(F_{X_2}(X_2) - F_{X_2}(X_2-))\\
&\quad +\tfrac{1}{2}(F_{X_1}(X) - F_{X_1}(X_1-))(F_{X_2}(X_2) - F_{X_2}(X_2-))\big)\\
&= \mathrm{E}\left(\frac{(F_{X_1}(X_1) + F_{X_1}(X_1-))(F_{X_2}(X_2) + F_{X_2}(X_2-))}{4}\right).
\end{aligned}$$

On combining this result with (13) one finds that $\rho(\mathcal{C}^S_{\boldsymbol{X}})/12$ also equals the covariance between $(F_{X_i}(X_i) + F_{X_i}(X_i-))/2$, $i = 1, 2$. Hence, it seems suitable to divide $\rho(\mathcal{C}^S_{\boldsymbol{X}})/12$ by the square root of variances of these random variables. These are evaluated in the following lemma.

Lemma 10. *With the above notation*,

$$\operatorname{var}\left(\frac{(F_{X_i}(X_i) + F_{X_i}(X_i-))}{2}\right) = \frac{1}{12}\left[1 - \mathrm{E}\big(\Delta F_{X_i}(X_i)\big)^2\right], \quad i = 1, 2. \tag{27}$$

Proof. First note that, for $i = 1, 2$,

$$\begin{aligned}
\tfrac{1}{3} &= \mathrm{E}(\psi^2(X_i, U_i)) = \mathrm{E}(\mathrm{E}(\psi^2(X_i, U_i)|X_i))\\
&= \mathrm{E}\Big(F_{X_i}(X_i-)^2 + F_{X_i}(X_i-)F_{X_i}(X_i) - F_{X_i}(X_i-)^2\\
&\quad +\tfrac{1}{3}\big(F_{X_i}(X_i)^2 - 2F_{X_i}(X_i)F_{X_i}(X_i-) + F_{X_i}(X_i-)^2\big)\Big),
\end{aligned}$$

which after some minor algebraic simplifications leads to

$$\mathrm{E}\left(\frac{(F_{X_i}(X_i) + F_{X_i}(X_i-))^2}{4}\right) = \frac{1}{3} - \mathrm{E}\left(\frac{(F_{X_i}(X_i) - F_{X_i}(X_i-))^2}{12}\right). \tag{28}$$

The result now follows immediately on combining (28) and (13). □

Hence, $|\rho(\mathcal{C}^S_{\boldsymbol{X}})| \leqslant \sqrt{\big(1 - \mathrm{E}(\Delta F_{X_1}(X_1))^2\big)\big(1 - \mathrm{E}(\Delta F_{X_2}(X_2))^2\big)}$. The new bounds $\pm\sqrt{\big(1 - \mathrm{E}(\Delta F_{X_1}(X_1))^2\big)\big(1 - \mathrm{E}(\Delta F_{X_2}(X_2))^2\big)}$ are less sharp, meaning that they are not necessarily attained when the marginals are countermonotonic and comonotonic, respectively; see Fig. 3. They are, however, reached when the marginals are a.s. strictly monotone and continuous functions of one another, which leads to the following definition.

Definition 11. Let $\boldsymbol{X} = (X_1, X_2)$ be a bivariate random vector with arbitrary marginals, then the *non-continuous version of Spearman's rho* is given by

$$\rho(X_1, X_2) = \frac{12\int \mathcal{C}^S_{\boldsymbol{X}}(u, v)\, du\, dv - 3}{\sqrt{\big(1 - \mathrm{E}(\Delta F_{X_1}(X_1))^2\big)\big(1 - \mathrm{E}(\Delta F_{X_2}(X_2))^2\big)}}. \tag{29}$$

As in the case of Kendall's tau it immediately follows that ρ satisfies A1–A3 and A7. Axioms A4* and A5* can be obtained easily by noting that $F_{T(X_i)}(T(X_i)) + F_{T(X_i)}(T(X_i)-)$ equals $F_{X_i}(X_i) + F_{X_i}(X_i-)$ if T is strictly increasing and continuous on ran X_i and $-(F_{X_i}(X_i) + F_{X_i}(X_i-))$ if T is strictly decreasing and continuous on ran X_i. The only difficulty is again the weak convergence. There the situation is analogous to Kendall's tau; see Nešlehová [19] for details.

4.4. Examples

One of the major differences between the continuous and non-continuous case is that the above defined generalizations of Kendall's tau and Spearman's rho reach the bounds 1 and -1 for sharper kind of monotonic dependence, as described in Axiom A4*. In order to illustrate the performance of τ and ρ, we hence consider several examples of comonotonic and countermonotonic marginals.

Example 12. Let X_1 and X_2 follow binomial distributions $\mathcal{B}(n_1, p)$ and $\mathcal{B}(n_2, q)$, respectively. As can be easily verified, in the special case when $n_1 = n_2 = 1$, τ and ρ are equal. In this situation, τ and ρ even coincide with the linear correlation coefficient.

Fig. 4 gives Kendall's tau and Spearman's rho for $p = 0.4$ and q ranging from 0 to 1 (values on the *x*-axis). Both measures seem to behave similarly, although the values of ρ are slightly larger in the comonotonic case; this illustrates the result obtained by Capéraaà and Genest [1] and Mesfioui and Tajar [16] that Kendall's tau is less than or equal to Spearman's rho for positively dependent random variables. As described in Axiom A5*, there is a symmetry between the graphics in the upper and lower panels in Fig. 4: for $n_1 = n_2$ fixed, the τ and ρ evaluated in q in the upper picture is equal to $-\tau$ and $-\rho$ evaluated in $1-q$ in the lower picture. Note that for X_1 and X_2 comonotonic, $\tau = \rho = 1$ if and only if $n_1 = n_2$ and $p = q$. In analogy, if X_1 and X_2 are countermonotonic, $\tau = \rho = -1$ if and only if $n_1 = n_2$ and $p = 1-q$. The plots moreover show that τ and ρ converge to 0 if q converges to either 0 or 1. This can be explained by the fact that, for q equal to 0 or 1, X_2 is a.s. a constant and hence independent of X_1. Also note that Kendall's tau and Spearman's rho get closer to ± 1 with increasing *n*. The linear correlation coefficient performs similarly to τ and ρ; this is, however, no longer the case if one alters the support of either X_1 or X_2.

Example 13. The previous example can be generalized by considering marginal distribution functions of the form

$$F_{X_1} = \lambda F_1 + (1-\lambda)F_2 \quad \text{and} \quad F_{X_2} = \mu G_1 + (1-\mu)G_2, \tag{30}$$

where F_1 and G_1 are distribution functions of purely discrete rvs and F_2 and G_2 of continuous rvs, respectively. Because of the symmetry between the countermonotonic and comonotonic case, we focus on the comonotonic case only, i.e. we calculate Kendall's tau and Spearman's rho of $\mathcal{M}(F_{X_1}, F_{X_2})$, where $\mathcal{M}$ is the Fréchet–Hoeffding upper bound. Note that the normalizing functions of Kendall's tau and Spearman's rho, i.e. the functions in the denominator of (26) and (29) respectively, do not depend on F_2 or G_2.

Fig. 5 illustrates a situation similar to the upper panel of Fig. 4. Here we set $\lambda = \mu = 0.5$ and consider F_1 and G_1 the distribution functions of binomial rvs, $\mathcal{B}(n_1, p)$ and $\mathcal{B}(n_2, q)$, respectively. Furthermore, F_2 and G_2 are distribution functions of standard normal $\mathcal{N}(0, 1)$ rvs. As in Fig. 4, Kendall's tau and Spearman's rho are displayed as functions of the parameter q of G_1, ranging from 0 to 1. Note that as in the previous Example 12, the values of both Kendall's tau and Spearman's rho tend to 1 with increasing n_1 and n_2 and are exactly equal to 1 for $p = q$. However, although

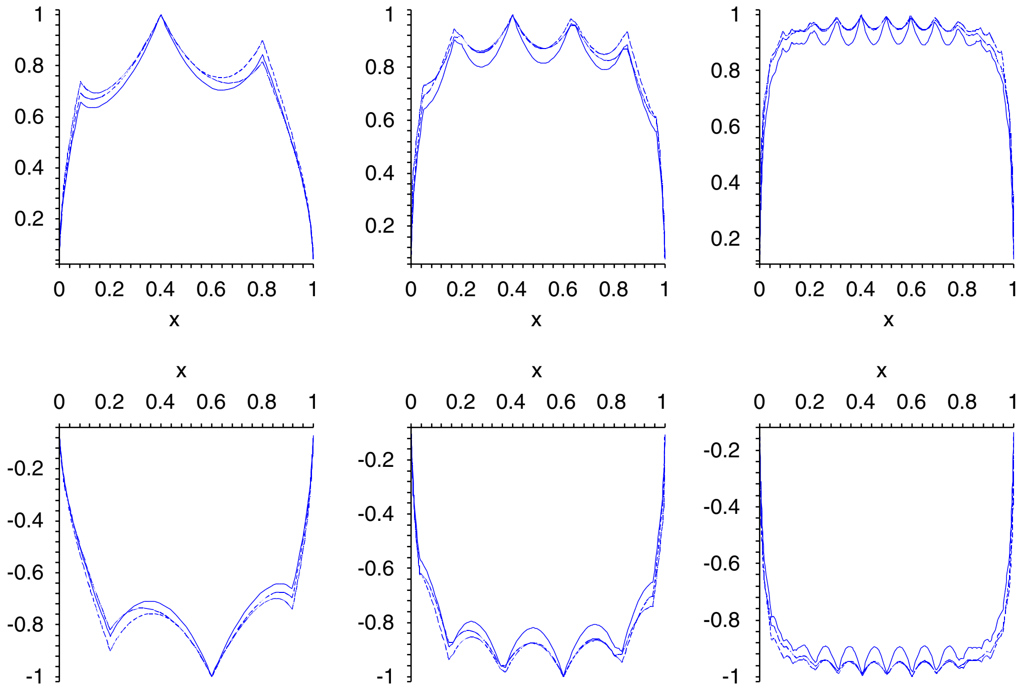


Fig. 4. Kendall's tau (solid lines), Spearman's rho (dashed lines) and linear correlation coefficient (dotted lines) for comonotonic (upper panel) and countermonotonic (lower panel) binomial random variables.

Fig. 5 shows similarities with Fig. 4, the pictures are not entirely the same. The difference is caused by the fact that the continuous distribution functions F_2 and G_2 do enter through the calculation of the numerators in (26) and (29) and hence do influence the values of Kendall's tau and Spearman's rho.

This issue is illustrated in Fig. 6. Here we set $\lambda = \mu = 0.5$ and $F_1 = G_1$. In the plots (a) and (b), $F_1 = G_1 = \mathcal{B}(4, 0.4)$ and $F_2 = \mathcal{N}(0, 1)$. Kendall's tau and Spearman's rho are then calculated for various choices of G_2: we consider $G_2 = \mathcal{N}(x, 1)$ for $x \in [0, 2.5]$ in Fig. 6(a) and $G_2 = \mathcal{N}(0, x)$ for $x \in [1, 3]$ in Fig. 6(b), respectively. Note that for the values of the parameters which yield $F_2 = G_2$, $\tau = \rho = 1$.

In Fig. 6(c), Kendall's tau is displayed as function of the parameter q of $G_1 = \mathcal{B}(4, q)$. In comparison with Fig. 5(b), different distribution functions F_2 and G_2 are considered:

- the solid line in Fig. 6(c) displays Kendall's tau for F_2 and G_2 standard normal as in Fig. 5(b);
- the dashed line in Fig. 6(c) displays Kendall's tau for F_2 Student's *t* with 5 degrees of freedom and G_2 standard normal;
- the dotted line in Fig. 6(c) displays Kendall's tau for F_2 Student's *t* with 2 degrees of freedom and G_2 standard normal;
- and finally, the dot-dashed line in Fig. 6(c) displays Kendall's tau for F_2 and G_2 both Student's *t* with 2 degrees of freedom.

Spearman's rho behaves similarly, although the difference appears to be much less significant and barely noticeable visually.

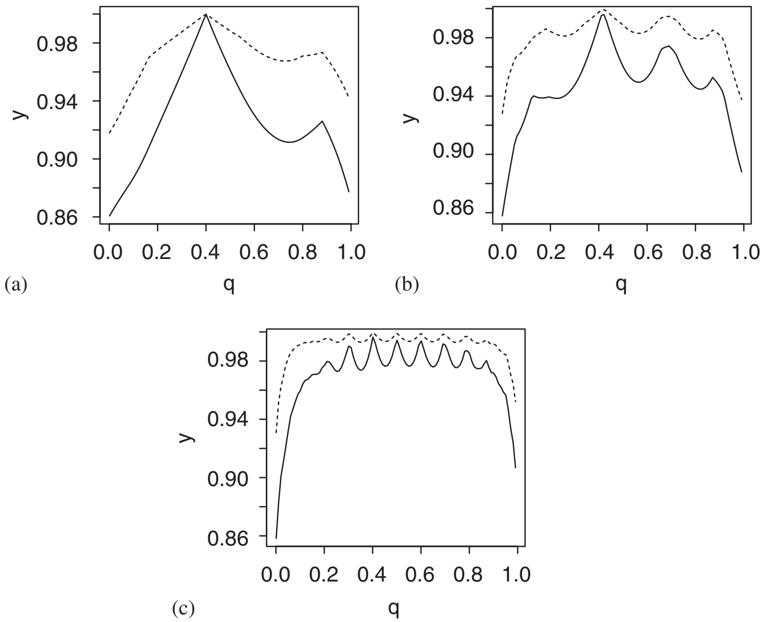


Fig. 5. Kendall's tau (solid lines), Spearman's rho (dashed lines) for F_1 and G_1 binomial, F_2 and G_2 standard normal and $\lambda = \mu = 0.5$.

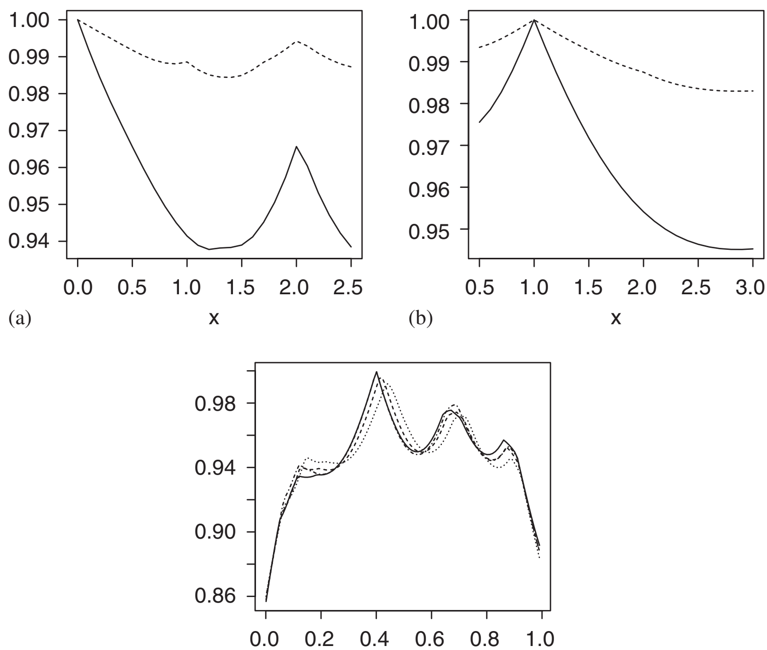


Fig. 6. Comparison of Kendall's tau (solid lines) and Spearman's rho (dashed lines) for various choices of F_2 and G_2 (plots (a) and (b)) and Kendall's tau as function of the parameter of G_1 for various choices of F_2 and G_2 (plot (c)).

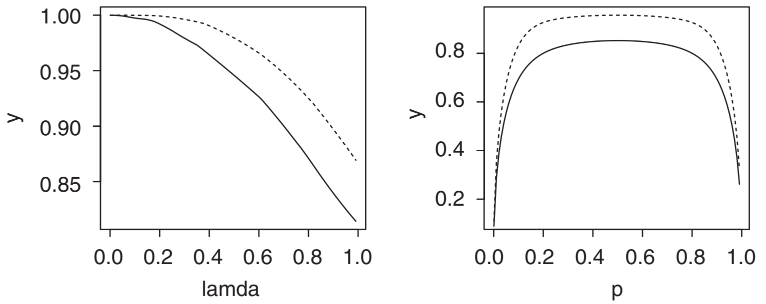


Fig. 7. Kendall's tau (solid lines) and Spearman's rho (dashed lines) as functions of λ for $\mu = \lambda$, $F_1 = \mathcal{B}(4, 0.4)$, $G_1 = \mathcal{B}(4, 0.8)$, and $F_2 = G_2 = \mathcal{N}(0, 1)$ (left); Kendall's tau (solid lines) and Spearman's rho (dashed lines) as functions of q for $\lambda = 1$, $\mu = 0$, $F_1 = \mathcal{B}(4, p)$, $G_1 = \mathcal{B}(4, 0.8)$, and $F_2 = G_2 = \mathcal{N}(0, 1)$ (right).

Finally, we examine the influence of the parameters λ and μ. For $\lambda = \mu = 1$, the situation compresses into the purely discrete case discussed in the previous example. For $\lambda = \mu = 0$, both marginals are continuous and hence $\tau = \rho = 1$.

Fig. 7 shows τ and ρ as function of $\lambda \in [0, 1]$, for $\mu = \lambda$, F_1 and G_1 binomial and F_2 and G_2 standard normal (left picture). If $\lambda = 1$ and $\mu = 0$, the joint distribution has one margin purely discrete and the other continuous. After some straightforward calculations, it can be verified that in this case, $\tau = \sqrt{1 - \mathrm{E}(\Delta F_{X_1}(X_1))}$ and $\rho = \sqrt{1 - \mathrm{E}(\Delta F_{X_1}(X_1))^2}$. The values of τ and ρ are displayed in Fig. 7 (right).

5. Empirical distributions

In this section we focus on a special family of discrete distributions—the empirical distributions corresponding to bivariate random samples. We show that the versions of Kendall's tau and Spearman's rho obtained above are equal to the sample versions of these quantities known from statistics. In order to state the results, some additional notation is required. Assume we are given a sample $\{x_k, y_k\}_{k=1}^n$ of size n from an arbitrary bivariate distribution function H with marginals F and G, say. As H is not necessarily continuous, ties in the observations are possible. This means that the empirical distribution functions, henceforth denoted by $\widehat{H}_n$, $\widehat{F}_n$ and $\widehat{G}_n$, can have jumps of size greater than $1/n$. Suppose that there are r distinct values of the x_k's, $\xi_1 < \cdots < \xi_r$, and s distinct values of the y_k's, $\eta_1 < \cdots < \eta_s$. Furthermore, set $u_i := \#\{x_k | x_k = \xi_i\}$, $v_j := \#\{y_k | y_k = \eta_j\}$ and $w_{ij} := \#\{(x_k, y_k) | x_k = \xi_i \wedge y_k = \eta_j\}$ as well as $p_i := u_i/n$, $q_j := v_j/n$ and $h_{ij} := w_{ij}/n$ for the corresponding frequencies. Finally, we consider order statistics and ranks, each understood componentwise. The order statistics will be denoted by $x_{(i)}$ and $y_{(i)}$, respectively, $1 \leqslant i \leqslant n$. As there are possibly ties in the observations, the ranks that will enter into the subsequent calculations will be twofold. First the ordinary ranks given by $R(x_k) := \sum_{i=1}^n 1_{(x_i \leqslant x_k)}$ and $R(y_k) := \sum_{i=1}^n 1_{(y_i \leqslant y_k)}$. Secondly, suppose that i is such that $x_k = \xi_i$. Then the average rank of x_k is the following:

$$\overline{R}(x_k) = \begin{cases} \frac{1+\cdots+u_1}{u_1} = \frac{u_1+1}{2} & \text{if } i = 1, \\ \sum_{j=1}^{i-1} u_j + \frac{u_i+1}{2} & \text{otherwise.} \end{cases}$$

Analogously, $\overline{R}(y_k)$ denotes the average rank of y_k.

Observe that any copula corresponding to $\widehat{H}_n$ is uniquely determined only in points $(R(x_{(i)})/n, R(y_{(j)})/n)$ where its value equals the number of pairs (x, y) in the sample that simultaneously satisfy $x \leqslant x_{(i)}$ and $y \leqslant y_{(j)}$ divided by n. The linear interpolation of these values leading to the standard extension copula of $\widehat{H}_n$ has been considered by [2] in order to construct distribution-free tests of independence and is referred to as "empirical copula". He considers merely the case of continuous H and hence of samples without ties. In the general case, the standard extension copula of $\widehat{H}_n$ is, however, a natural generalization of Deheuvels' definition; we will therefore also refer to it as the *empirical copula* of the sample. Note, however, that several alternative definitions of the empirical copula exist; see for instance [25, Section 3.9.4.4]. These are asymptotically equal to Deheuvels' definition, but generally not continuous and hence not proper copulas.

We now turn back to Kendall's tau and Spearman's rho as defined in Definitions 9 and 11.

Theorem 14. *Kendall's τ corresponding to the empirical distribution function $\widehat{H}_n$ of a sample $\{x_k, y_k\}_{k=1}^n$ from an arbitrary bivariate distribution H equals the sample version of Kendall's tau,*

$$\widehat{\tau} = \frac{\#[\textit{concordant pairs}] - \#[\textit{discordant pairs}]}{\sqrt{\binom{n}{2} - u}\sqrt{\binom{n}{2} - v}}, \tag{31}$$

where $u = \sum_{k=1}^{r} \binom{u_k}{2}$ and $v = \sum_{l=1}^{s} \binom{v_l}{2}$.

Proof. On substituting the empirical distribution functions in (24) one finds, after some simplification, that the difference between the probabilities of concordance and discordance equals

$$Q = \sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{i-1}\sum_{l=1}^{j-1}\big(h_{kl}h_{ij} + h_{ij}h_{kl} - h_{kj}h_{il} - h_{il}h_{kj}\big),$$

which leads to

$$\tau = 2\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{i-1}\sum_{l=1}^{j-1}\big(h_{kl}h_{ij} - h_{kj}h_{il}\big) \Big/ \sqrt{1 - \sum_{i=1}^{r} p_i^2}\sqrt{1 - \sum_{j=1}^{s} q_j^2}\,. \tag{32}$$

Since $h_{ij}h_{kl}$ equals $w_{ij}w_{kl}/n^2$ if (ξ_i, η_j) and (ξ_k, η_l) are in the sample (concordant pair if $k < i$ and $l < j$) and $h_{il}h_{kj}$ equals $w_{il}w_{kj}/n^2$ if (ξ_i, η_l) and (ξ_k, η_j) are in the sample (discordant pair if $k < i$ and $l < j$), the numerator in (32) equals

$$\frac{2}{n^2}\big(\#[\text{concordant pairs}] - \#[\text{discordant pairs}]\big).$$

In addition, we examine the quantities $n^2/2(1 - \sum_{i=1}^{r} p_i^2)$. If no ties are present, then $r = n$, ξ_i equals the ith order statistic of $\{x_k\}_{k=1}^n$ for $i = 1, \ldots, n$ and $p_i = 1/n$. Therefore,

$$\frac{n^2}{2}\left(1 - \sum_{i-1}^{r} p_i^2\right) = \frac{n^2}{2}\left(1 - \sum_{i=1}^{n}\frac{1}{n^2}\right) = \binom{n}{2}.$$

Otherwise, $p_i = u_i/n$ and we get

$$\frac{n^2}{2}\left(1-\sum_{i=1}^{r}\frac{u_i^2}{n^2}\right)=\frac{n^2}{2}\left(\sum_{i=1}^{n}\frac{n-1}{n^2}-\sum_{i=1}^{r}\frac{u_i^2-u_i}{n^2}\right)=\binom{n}{2}-\sum_{i=1}^{r}\binom{u_i}{2}=\binom{n}{2}-u.$$

As $\frac{n^2}{2}(1-\sum_{j=1}^{s} q_j^2)$ can be re-written similarly, the theorem follows. □

Theorem 15. *Under the hypotheses of Theorem* 14, *Spearman's ρ corresponding to the empirical distribution function $\widehat{H}_n$ equals the sample version of Spearman's rho*,

$$\widehat{\rho}=\frac{\sum_{k=1}^{n}(\overline{R}(x_k)-\overline{R}_x)(\overline{R}(y_k)-\overline{R}_y)}{\sqrt{\sum_{k=1}^{n}(\overline{R}(x_k)-\overline{R}_x)^2\sum_{k=1}^{n}(\overline{R}(y_k)-\overline{R}_y)^2}}, \tag{33}$$

where $\overline{R}_x$ and $\overline{R}_y$ are given by $\overline{R}_x=\frac{1}{n}\sum_{i=1}^{n}\overline{R}(x_i)$ and $\overline{R}_y=\frac{1}{n}\sum_{j=1}^{n}\overline{R}(y_j)$.

Proof. First note that, by definition of the average ranks,

$$\overline{R}_x=\frac{1}{n}\sum_{k=1}^{r}u_k\frac{\left(\sum_{i=1}^{k-1}u_i+1\right)+\cdots+\left(\sum_{i=1}^{k-1}u_i+u_k\right)}{u_k}=\frac{1}{n}\sum_{i=1}^{n}i=\frac{n+1}{2}.$$

Similarly $\overline{R}_y=(n+1)/2$ and hence $\overline{R}_y=\overline{R}_x$. According to the discussion in Section 4.3, $\rho(X,Y)=\mathrm{corr}(\tilde{X},\tilde{Y})$, where $\tilde{X}$ and $\tilde{Y}$ are random variables with expectations $\frac{1}{2}$ given by

$$\mathsf{P}\left[\tilde{X}=\frac{\widehat{F}(\xi_i)+\widehat{F}(\xi_{i-1})}{2}\right]=p_i,$$
$$\mathsf{P}\left[\tilde{Y}=\frac{\widehat{G}(\eta_j)+\widehat{G}(\eta_{j-1})}{2}\right]=q_j,$$
$$\mathsf{P}\left[\tilde{X}=\frac{\widehat{F}(\xi_i)+\widehat{F}(\xi_{i-1})}{2},\tilde{Y}=\frac{\widehat{G}(\eta_j)+\widehat{G}(\eta_{j-1})}{2}\right]=h_{ij},$$

for $i=1,\dots,r$ and $j=1,\dots,s$. One easily verifies that $\frac{\widehat{F}(\xi_i)+\widehat{F}(\xi_{i-1})}{2}-\frac{1}{2}=\frac{1}{n}(\overline{R}(x_k)-\overline{R}_x)$ for any x_k with $x_k=\xi_i$ and similarly that $\frac{\widehat{G}(\eta_j)+\widehat{G}(\eta_{j-1})}{2}-\frac{1}{2}=\frac{1}{n}(\overline{R}(y_l)-\overline{R}_y)$ for any y_l such that $y_l=\eta_j$. Because there are exactly u_i such x_k's, v_j such y_l's and finally w_{ij} observations in $\{x_k,y_k\}_{k=1}^{n}$ equal to (ξ_i,η_j),

$$\begin{aligned}\mathrm{cov}(\tilde{X},\tilde{Y})&=\sum_{i=1}^{r}\sum_{j=1}^{s}h_{ij}\left(\frac{\widehat{F}(\xi_i)+\widehat{F}(\xi_{i-1})}{2}-\frac{1}{2}\right)\left(\frac{\widehat{G}(\eta_j)+\widehat{G}(\eta_{j-1})}{2}-\frac{1}{2}\right)\\&=\frac{1}{n^3}\sum_{k=1}^{n}(\overline{R}(x_k)-\overline{R}_x)(\overline{R}(y_k)-\overline{R}_y)\end{aligned}$$

and

$$\mathrm{var}(\tilde{X})=\sum_{i=1}^{r}p_i\left(\frac{\widehat{F}(\xi_i)+\widehat{F}(\xi_{i-1})}{2}-\frac{1}{2}\right)^2=\frac{1}{n^3}\sum_{k=1}^{n}(\overline{R}(x_k)-\overline{R}_x)^2.$$

Hence,

$$\rho(X,Y)=\frac{\operatorname{cov}(\tilde X,\tilde Y)}{\sqrt{\operatorname{var}(\tilde X)\operatorname{var}(\tilde Y)}}=\frac{\sum_{k=1}^{n}(\overline{R}(x_k)-\overline{R}_x)(\overline{R}(y_k)-\overline{R}_y)}{\sqrt{\sum_{k=1}^{n}(\overline{R}(x_k)-\overline{R}_x)^2\sum_{k=1}^{n}(\overline{R}(y_k)-\overline{R}_y)^2}},$$

which is (33). $\square$

These results support the choice of the normalization used in Definitions 9 and 11 of τ and ρ. Also, Theorems 14 and 15 generalize the statement that the sample versions of Kendall's tau and Spearman's rho can be expressed in terms of the empirical copula. In the general case, however, the values of the marginal empirical distribution functions enter into the calculations as well.

6. Discussion

In this paper, we obtained generalizations for Kendall's tau and Spearman's rho for arbitrary random variables. These depend on the corresponding standard extension copula but also on the marginal distribution functions. The fact that marginal distribution functions take influence upon the dependence structure is characteristic for non-continuous distributions. In the case of concordance measures, this "nuisance" causes difficulties that are basically twofold. On one hand, the measures typically do not reach the bounds ± 1 for countermonotonic and comonotonic marginals. If this requirement is, however, loosened in the sense that ± 1 is attained when the marginals are strictly monotonic and continuous transformations of one another, measures that fulfill the desirable property indeed exist. This turns out to be the case for both Kendall's tau and Spearman's rho. The second difficulty is that the measures of concordance corresponding to a weakly convergent sequence of random vectors may not converge; see Nešlehová [19]. Further research, however, on this issue would certainly be welcome. Otherwise, the obtained generalizations of Kendall's tau and Spearman's rho share all the properties those quantities have in the case of continuous distribution functions. Moreover, the constructions rely on an alternative transformation of an arbitrary random variable to a uniformly distributed one. This technique may prove useful in further investigations of dependence structures in the general case.

The results derived in this paper contribute mainly to a *quantification* of monotonic dependence. The *modeling* side, however, still remains challenging. To illustrate this, Fig. 8 shows

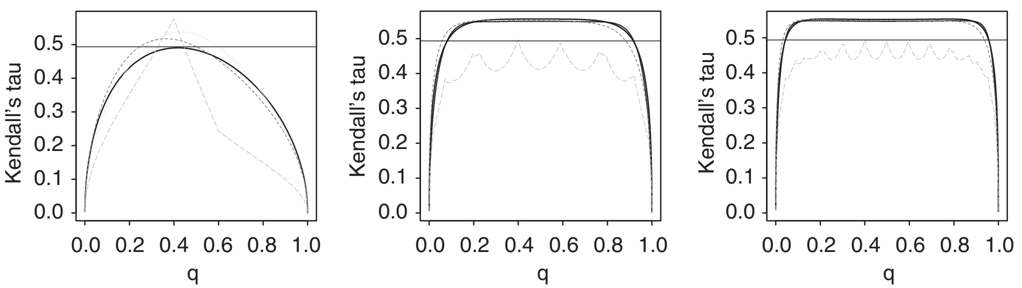


Fig. 8. Kendall's tau for binomial marginals and a Gauss copula with parameter 0.7 (solid line), Frank copula with parameter 5.6 (short-dashed line), Gumbel copula with parameter 1.97 (dotted line) and Fréchet copula with parameter 0.746 (long-dashed line).

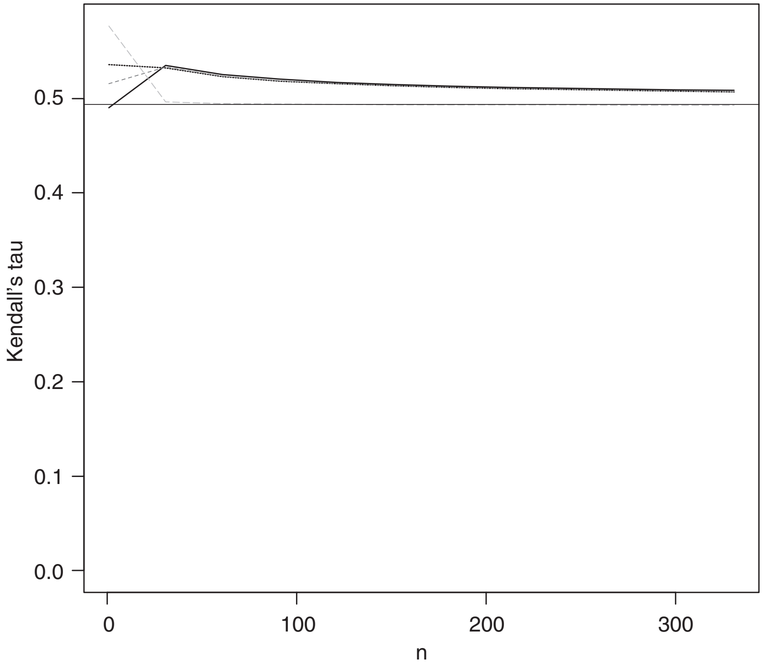


Fig. 9. Kendall's tau for binomial marginals and a Gauss copula with parameter 0.7 (solid line), Frank copula with parameter 5.6 (short-dashed line), Gumbel copula with parameter 1.97 (dotted line) and Fréchet copula with parameter 0.746 (long-dashed line) for $n \to \infty$.

Kendall's tau for binomial marginals that have been joined together by four different copulas: Gauss, Frank, Gumbel and Fréchet. The copula parameter is in each case chosen in a way that Kendall's tau of the copula is approximatively 0.49. The graphics reveal the ambiguity of this modeling approach: while Kendall's tau of the copula remains constant while altering the marginal parameters, Kendall's tau of the so-created bivariate binomial distribution does not. In fact, its value is quite different from 0.49, especially when the parameter n of the binomial marginals is small. With increasing n, the values of Kendall's tau converge to Kendall's tau of the corresponding copula. The convergence seems, however, rather slow, except for the Fréchet case, as is shown in Fig. 9.

Acknowledgements

This work was partly supported by RiskLab, ETH Zürich. The author is grateful to Professors Dietmar Pfeifer and Paul Embrechts for many fruitful discussions and support. She would also like to thank Prof. Marco Scarsini and Dr. Alfred Müller as well as the referees for several useful comments.

References

[1] P. Capéraà, C. Genest, Spearman's ρ is larger than Kendall's τ for positively dependent random variables, J. Nonparametric Statist. 2 (1993) 183–194.

[2] P. Deheuvels, Non parametric tests of independence. in: J.P. Raoult (Ed.), Statistique Non Paramétrique Asymptotique Proceedings, Lecture Notes in Mathematics 821, Springer, Berlin, 1979, pp. 95–107.
[3] M. Denuit, P. Lambert, Constraints on concordance measures in bivariate discrete data, J. Multivariate Anal. 93 (2005) 40–57.
[4] P. Embrechts, A.J. McNeil, D. Straumann, Correlation and dependence in risk management: properties and pitfalls, in: M.A.H. Dempster (Ed.), Risk Management: Value at Risk and Beyond, Cambridge University Press, Cambridge, 2002, pp. 176–223.
[5] T.S. Ferguson, Mathematical Statistics: A Decision Theoretic Approach, Academic Press, New York, 1967.
[6] W. Hoeffding, Maßstabinvariante Korrelationstheorie, Schrift. Math. Seminars Inst. Angew. Math. Univ. Berlin 5 (3) (1940) 181–233.
[7] W. Hoeffding, Maßstabinvariante Korrelationstheorie für diskontinuierliche Verteilungen, Arch. Math. Wirtschafts- und Sozialforschung VII (2) (1940) 4–70.
[8] H. Joe, Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.
[9] K. Jogdeo, Concepts of dependence, in: S. Kotz, N.L. Johnson (Eds.), Encyclopedia of Statistical Sciences, Wiley, New York, 1982, pp. 324–334.
[10] W.H. Kruskal, Ordinal measures of association, J. Amer. Statist. Assoc. 53 (1958) 814–861.
[11] E.L. Lehmann, Some concepts of dependence, Ann. of Math. Statist. 37 (III) (1966) 1137–1153.
[12] E.L. Lehmann, Nonparametrics: Statistical Methods Based on Ranks, Holden-Day, Inc., San Francisco, 1975.
[13] A. Lindner, A. Szimayer, A limit theorem for copulas, preprint, Technische Universität München, 2004.
[14] A.W. Marshall, Copulas, marginals and joint distributions, in: L. Rüschendorf, B. Schweizer, M.D. Taylor (Eds.), Distributions with Fixed Marginals and Related Topics, Institute of Mathematical Statistics, Hayward, CA, 1996, pp. 213–222.
[15] A.J. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management: Concepts, Techniques and Tools, Princeton University Press, Princeton, 2005.
[16] M. Mesfioui, A. Tajar, On the properties of some nonparametric concordance measures in the discrete case, J. Nonparametric Statist. 17 (2005) 541–554.
[17] R.B. Nelsen, Copulas and association, in: G. Dall'Aglio, S. Kotz, G. Salinetti (Eds.), Advances in Probability Distributions with Given Marginals, Kluwer Academic Publishers, Dordrecht, 1991, pp. 51–74.
[18] R.B. Nelsen, An Introduction to Copulas, Springer, New York, 1999.
[19] J. Nešlehová, Dependence of Non-Continuous Random Variables, Ph.D. Thesis, Carl von Ossietzky Universität Oldenburg, 2004.
[20] M. Scarsini, On measures of concordance, Stochastica 8 (1984) 201–218.
[21] B. Schweizer, A. Sklar, Operation on distribution functions not derivable from operations on random variables, Studia Math. 52 (1974) 43–52.
[22] B. Schweizer, E.F. Wolff, On nonparametric measures of dependence for random variables, Ann. Statist. 9 (1981) 870–885.
[23] A. Tajar, M. Denuit, P. Lambert, Copula-type representations for random couples with Bernoulli margins, Discussion paper #0119, Institute of Statistics, Université Catholique de Louvain, Belgium, 2001.
[24] A. Tajar, M. Mesfioui, M. Denuit, On the monotonicity of some nonparametric dependence measures with respect to concordance ordering. Discussion paper #0142, Institute of Statistics, Université Catholique de Louvain, Belgium, 2001.
[25] A.W. Van der Vaart, J.A. Wellner, Weak Convergence and Empirical Processes: With Applications to Statistics, Springer, New York, 1996.