Lecture Notes of
Seminario Interdisciplinare di Matematica
Vol. 12(2015), pp. 221 – 242.

A survey of copula–based measures of association

Carlo Sempi[1]

Abstract. We survey the measures of association that are based on bivariate copulas. Almost no proof will be reported, although an exception is made in the case of the Schweizer–Wolff measure, since the details of the proof are mainly contained in Wolff's Ph.D. dissertation, which is not readily available.

1. Introduction

"Measure of association" is a broad term that denotes the class of all the measures that have been constructed with the aim of quantifying specific relationships between two or more random variables. The term may include, for instance, measures that want to capture functional relationship (i.e., linear relationships) among random variables, as well as measures of dependence that aim at quantifying the "degree of non–independence" in a set of variables. In this paper we review those among these measures that may expressed in terms of the copulas of two random variables.

In the next section we recall, without any proof, the properties of (bivariate) copulas that will be needed in the sequel. References on copulas are the following books and surveys [9, 14, 4, 5].

2. Copulæ

A (bivariate) copula is a distribution function (=d.f.) on $\mathbb{I}^2$ whose univariate marginals are uniformly distributed on $\mathbb{I}$. Equivalently, a copula C is a function $C : \mathbb{I}^2 \to \mathbb{I}$ such that

(C1) $C(t, 1) = C(1, t) = t$ and $C(t, 0) = C(0, t) = 0$ for every $t \in \mathbb{I}$;
(C2) for all u, u', v and v' in $\mathbb{I}$ with $u \le u'$ and $v \le v'$,

$$C(u', v') - C(u', v) - C(u, v') + C(u, v) \ge 0 \, .$$

The set of copulas will be denoted by $\mathcal{C}_2$.

The following three examples of copulas are essential: the *comonotonicity* copula $M_2(u, v) = \min\{u, v\}$, the *independence* copula $\Pi_2(u, v) = uv$ and the *countermonotonicity* copula $W_2(u, v) = \max\{0, u + v - 1\}$.

---

[1]C. Sempi, Università del Salento, Dipartimento di Matematica e Fisica "Ennio De Giorgi", Via per Arnesano, 73100 Lecce, Italy; `carlo.sempi@unisalento.it`

Keywords. Measures of association, copulæ, Sperman's rho, Kendall's tau, Schweizer–Wolff measure.

AMS Subject Classification. Primary 62H20, Secondary 60E05.

Other important families of copulas are listed below.

- *Archimedean*

(2.1) $$C(u,v) = f^{(-1)}\left(f(u) + f(v)\right) ,$$

where $f : \mathbb{I} \to [0, +\infty]$ is continuous, convex, strictly decreasing and such that $f(1) = 0$ and $f^{(-1)}(t) = f^{-1}(t)$ if $t \in [0, f(0)]$, while $f^{(-1)}(t) = 0$ for $t \geq f(0)$.

- *Eyraud–Farlie-Gumbel–Morgenstern* (briefly EFGM)

(2.2) $$C_\alpha^{\mathbf{EFGM}}(u,v) = uv\ (1 + \alpha(1-u)\,(1-v)) , \quad \alpha \in [-1,1] ;$$

- *Marshall–Olkin* for $\alpha, \beta \in \,]0,1[$

(2.3) $$C_{\alpha,\beta}^{\mathbf{MO}}(u,v) := \min\left\{u^{1-\alpha}v, uv^{1-\beta}\right\} = \begin{cases} u^{1-\alpha}v & , \quad u^\alpha \geq v^\beta , \\ uv^{1-\beta} & , \quad u^\alpha \leq v^\beta , \end{cases}$$

- *Gumbel–Hougaard*

(2.4) $$C_\alpha^{\mathbf{GH}}(u,v) = \exp\left(-\left((-\ln u)^\alpha - (\ln v)^\alpha\right)^{1/\alpha}\right) , \ \alpha \geq 1 .$$

For $\alpha = 1$ one obtains the independence copula as a special case, and the limit of $C_\alpha^{\mathbf{GH}}$ for $\alpha \to +\infty$ is the comonotonicity copula M_2.

- *Bivariate Gaussian*

(2.5) $$C_\rho^{\mathbf{Ga}}(u,v) =$$

$$= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left(-\frac{s^2 - 2\rho st + s^2}{2(1-\rho^2)}\right) ds\, dt ,$$

where ρ is in $]-1,1[$, and Φ^{-1} denotes the inverse of the standard Gaussian distribution $N(0,1)$. For more details, see [12].

- *Mardia–Takahashi–Clayton*

(2.6) $$C_\alpha^{\mathbf{MTC}}(u,v) =$$

$$= \max\left\{0, \left(\frac{1}{u^\alpha} + \frac{1}{v^\alpha} - 1\right)^{-1/\alpha}\right\} \quad , \quad \alpha \in [-1 + \infty[\setminus \{0\} .$$

- *Frank*

(2.7) $$C_\alpha^{\mathbf{Frank}}(u,v) =$$

$$= -\frac{1}{\alpha} \ln\left(1 + \frac{(e^{-\alpha u} - 1)\ (e^{-\alpha v} - 1)}{(e^{-\alpha} - 1)}\right) \quad , \quad \alpha \in \mathbb{R} .$$

The limiting case $\alpha = 0$ corresponds to Π_2. Copulas of this type have been introduced by [6] in relation with a problem about associative functions on $\mathbb{I}$. They are absolutely continuous.

- *Extreme value*

(2.8) $$C_A(u,v) =$$

$$= \exp\left(A\left(\frac{\ln v}{\ln u + \ln v}\right) \ln(uv)\right) \quad , \quad (u,v) \in \,]0,1[^2 ,$$

where $A : \mathbb{I} \to [1/2, 1]$ is convex and satisfies the inequality $\max\{1-t, t\} \le A(t) \le 1$.

The importance of copulas stems form Sklar's theorem [23].

Theorem 2.1 (Sklar's theorem)**.** *Let F be a d.f. with univariate margins F_1, F_2. Let A_j denote the range of F_j, $A_j := F_j(\overline{\mathbb{R}})$ $(j = 1, 2)$. Then there exists a copula C such that for all $(x_1, x_2) \in \overline{\mathbb{R}}^2$,*

$$F(x_1, x_2) = C\left(F_1(x_1), F_2(x_2)\right) . \tag{2.9}$$

Such a C is uniquely determined on $A_1 \times A_2$ and, hence, it is unique when both F_1 and F_2 are continuous. Conversely, if C is a copula, then (2.9) defines a d.f..

If the marginals F_1 and F_2 are continuous the existence of the (unique) copula asserted by Sklar's theorem can be easily proved. Under the assumptions of Theorem 2.1, if F_1 and F_2 are continuous, then there exists a unique copula C associated with $\mathbf{X} = (X_1, X_2)$. It is determined, for every $(u, v) \in \mathbb{I}^d$, via the formula

$$C(u, v) = H\left(F_1^{(-1)}(u), F_2^{(-1)}(v)\right) , \tag{2.10}$$

where, for $j = 1, 2$, $F_j^{(-1)}(t) = \inf\{x \in \mathbb{R} : F_j(x) \ge t\}$ is the right–continuous quasi–inverse of F_j.

Example 2.1 (Copula of the extreme order statistics)**.** In the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $X_1, \cdots, X_n$ be indipendent and identically distributed random variables with a common continuous d.f. F and let $X_{(j)}$ be the j–th order statistic so that $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$. We wish to determine the copula of the vector $(X_{(1)}, X_{(n)})$. It is known (see, e.g., [3]) that the joint d.f. of the j–th and the k–th order statistics is

$$H_{j,k}(x, y) = \sum_{h=k}^{n} \sum_{i=j}^{h} \frac{n!}{i!(h-i)!(n-h)!} F^i(x) \left[F(y) - F(x)\right]^{h-i} \left[1 - F(y)\right]^{n-h} ,$$

for $x < y$, and $F_{(k)}(y)$, for $x \ge y$, where $F_{(k)}(t) := \mathbb{P}(X_{(k)} \le t)$ is the d.f. of the k–th order statistic

$$F_{(k)}(t) = \sum_{i=k}^{n} \binom{n}{i} F^i(t) \left[1 - F(t)\right]^{n-i} .$$

Hence, setting first $j = 1$ and then $k = n$, one has, for x and y in $\mathbb{R}$,

$$F_{(1)}(x) = 1 - \left[1 - F(x)\right]^n , \qquad F_{(n)}(y) = F^n(y) ;$$

now the joint d.f. of $X_{(1)}$ and $X_{(n)}$ is given by

$$H_{1,n}(x, y) = \begin{cases} \displaystyle\sum_{i=1}^{n} \binom{n}{i} F^i(x) \left[F(y) - F(x)\right]^{n-i} & , \quad x < y , \\ F_{(n)}(y) & , \quad x \ge y , \end{cases}$$

$$= \begin{cases} F^n(y) - \left[F(y) - F(x)\right]^n & , \quad x < y , \\ F^n(y) & , \quad x \ge y . \end{cases}$$

By recourse to (2.10), one can now write the (unique) copula of $X_{(1)}$ and $X_{(n)}$

$$C_{1,n} = H_{1,n}\left(F_{(1)}^{-1}(u), F_{(n)}^{-1}(v)\right) = \tag{2.11}$$
$$= \begin{cases} v - \left[v^{1/n} + (1-u)^{1/n} - 1\right]^n & , \quad 1-(1-u)^{1/n} < v^{1/n} , \\ v & , \quad 1-(1-u)^{1/n} \geq v^{1/n} . \end{cases}$$

The copula of eq. (2.11) is related to a 2–copula from the Mardia–Takahasi–Clayton copulas family (2.6). In fact, by using the symmetry $\xi(u,v) = (1-u,v)$, one has

$$C_{1,n}^{\xi}(u,v) = \begin{cases} \left(v^{1/n} + u^{1/n} - 1\right)^n & , \quad v^{1/n} + u^{1/n} - 1 > 0 , \\ 0 & , \quad \text{elsewhere,} \end{cases}$$

which belongs to the Mardia–Takahasi–Clayton family with $\theta = -1/n$.

One often considers, along with a copula C its *survival copula* defined by

$$\widehat{C}(u,v) := u + v - 1 + C(1-u, 1-v) . \tag{2.12}$$

One of the important properties of copulas is contained in the following result.

Theorem 2.2. *Let X and Y be continuous random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and consider the continuous mappings $\varphi : \operatorname{Ran} X \to \overline{\mathbb{R}}$ and $\psi : \operatorname{Ran} Y \to \mathbb{R}$.*

(a) *If both φ and ψ are strictly increasing, then, for every $(u,v) \in \mathbb{I}^2$,*
$$C_{\varphi(X),\psi(Y)}(u,v) = C_{XY}(u,v) ;$$

(b) *if φ is strictly increasing while ψ is strictly decreasing, then, for every $(u,v) \in \mathbb{I}^2$,*
$$C_{\varphi(X),\psi(Y)}(u,v) = C^{\sigma_2}(u,v) := u - C_{XY}(u, 1-v) ;$$

(c) *if φ is strictly decreasing while ψ is strictly increasing, then, for every $(u,v) \in \mathbb{I}^2$,*
$$C_{\varphi(X),\psi(Y)}(u,v) = C^{\sigma_1}(u,v) := v - C_{XY}(1-u, v) ;$$

(d) *if both φ and ψ are strictly decreasing, then, for every $(u,v) \in \mathbb{I}^2$,*
$$C_{\varphi(X),\psi(Y)}(u,v) = C^{\sigma_1\sigma_2}(u,v) := u + v - 1 + C_{XY}(1-u, 1-v) .$$

There is a one–to–one correspondence between copulas and stochastic measures on $(\mathbb{I}^2, \mathcal{B}(\mathbb{I}^2))$: a measure μ on this measurable space is said to be a *stochastic measure* if $\mu(A \times \mathbb{I}) = \mu(\mathbb{I} \times A) = \lambda(A)$ for every Borel subsest A of $\mathbb{I}$; here and in the following λ_d denotes the d–dimensional Lebesgue measure $(d = 1, 2)$.

Given a copula C for every rectangle $R =]a,b] \times]c,d]$ one defines

$$\mu_C(R) := C(b,d) - C(a,d) - C(b,c) + C(a,c)$$

by the usual techniques of measure theory the definition of μ_C is then extended to the family $\mathcal{B}(\mathbb{I}^2)$ of Borel subsets of $\mathbb{I}^2$. Conversely, if a stochastic measure μ is given, a copula C is defined via $C(u,v) := \mu([0,u] \times [0,v])$. Because of this correspondence one may write $\int_{\mathbb{I}^2} \cdots dC$ in order to denote the integral $\int_{\mathbb{I}^2} \cdots d\mu_C$.

The integration–by–parts formula presented below (see [11]) is needed in the calculation of some of statistical quantities surveyed here.

Theorem 2.3. *Let A and B be 2–copulas, and let the function $\varphi : \mathbb{I} \to \mathbb{R}$ be continuously differentiable, i.e. , $\varphi \in C^1(\mathbb{I})$. Then*

$$\int_{\mathbb{I}^2} \varphi \circ A \, dB = \tag{2.13}$$

$$= \int_0^1 \varphi(t)\, dt - \int_{\mathbb{I}^2} \varphi'(A(u,v))\, \partial_1 A(u,v)\, \partial_2 B(u,v)\, du\, dv =$$

$$= \int_0^1 \varphi(t)\, dt - \int_{\mathbb{I}^2} \varphi'(A(u,v))\, \partial_2 A(u,v)\, \partial_1 B(u,v)\, du\, dv\ . \tag{2.14}$$

Now, let $\mathcal{P}$ be a property of association that may be assigned to a random vector $\mathbf{X} = (X_1, X_2)$. A (copula–based) measure of association is any functional $\mathcal{M}_{\mathcal{P}} : \mathcal{C}_2 \to D \subset \mathbb{R}$ that assigns to each vector $\mathbf{X}$ of *continuous* r.v.'s with copula C a real number that is interpreted as the $\mathcal{P}$–degree of association of $\mathbf{X}$. If no confusion arises, $\mathcal{M}(\mathbf{X})$ or, equivalently, $\mathcal{M}(C)$ will denote the value of $\mathcal{M}$ for a vector $\mathbf{X}$ with copula C. Usually, the range D of a measure of association $\mathcal{M}$ is either $D = [0,1]$, where $\mathcal{M}(\mathbf{X}) = 0$ represents the absence of property $\mathcal{P}$ in $\mathbf{X}$ or $D = [-1,1]$, where $\mathcal{M}(\mathbf{X}) = 1$ (respectively, $\mathcal{M}(\mathbf{X}) = -1$) represents the maximal positive (respectively, negative) presence of property $\mathcal{P}$ in $\mathbf{X}$.

3. Concordance

Loosely speaking, one may say that two random variables are concordant if they tend to take large values together or to take small values together. Negative concordance of two random variables means that one of them takes large values while the other one takes small values.

Let X and Y be two continuous random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Given two different observations (x_j, y_j) and (x_k, y_k) of the random vector (X, Y) these will be said to be

- *concordant*, if one has either $x_j < x_k$ and $y_j < y_k$, or, $x_j > x_k$ and $y_j > y_k$, or equivalently, if $(x_k - x_j)\,(y_k - y_j) > 0$;
- *discordant*, if $x_j < x_k$ and $y_j > y_k$, or $x_j > x_k$ and $y_j < y_k$, or, equivalently, if $(x_k - x_j)\,(y_k - y_j) < 0$.

Definition 3.1. Given two copulas A and B in $\mathcal{C}_2$, B will be said to be *more concordant* than A, and this will be denoted by $A \prec B$, if $A(u,v) \le B(u,v)$ for every $(u,v) \in \mathbb{I}^2$.

In order to investigate concordance between two random pairs, it is expedient, following [10] and [14, Section 5.1.1], to introduce a *concordance function* Q. Assume that the pair (X_1, Y_1) and (X_2, Y_2) of continuous random vectors have (not necessarily equal) joint d.f.'s H_1 and H_2, respectively, but common marginals F and G; thus, both H_1 and H_2 belong to the Fréchet class $\Gamma(F,G)$, which is the set of two–dimensional d.f.'s whose marginals are F and G. It will turn out that Q depends only on the copulas C_1 and C_2 of the two vectors.

Theorem 3.1. *Let X_1, Y_1, X_2, Y_2 be continuous random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let the random vectors (X_1, Y_1) and (X_2, Y_2) be independent and let H_1 and H_2 be their respective joint d.f.'s and let the marginals d.f.'s satisfy*

$$F_{X_1} = F_{X_2} = F\ , \quad \text{and} \quad F_{Y_1} = F_{Y_2} = G\ ,$$

so that H_1 *and* H_2 *both belong to the Fréchet class* $\Gamma(F,G)$*, and*

$$H_1(x,y) = C_1(F(x),G(y)) \quad \text{and} \quad H_2(x,y) = C_2(F(x),G(y)) \ ,$$

where C_1 *and* C_2 *are the (unique) copulas of* (X_1,Y_1) *and* (X_2,Y_2)*, respectively. Define*

$$Q := \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right] \ . \tag{3.1}$$

Then Q *depends only on* C_1 *and* C_2 *and is given by*

$$Q(C_1,C_2) = 4\int_{\mathbb{I}^2} C_2(u,v)\, dC_1(u,v) - 1 \ . \tag{3.2}$$

Some of the properties of the concordance function Q will be useful in the sequel. They are collected in the following result, whose proof is immediate.

Theorem 3.2. *Let* C_1*,* C_2 *and* Q *have the same meaning as in Theorem 3.1; then:*

(a) Q *is symmetric:* $Q(C_1,C_2) = Q(C_2,C_1)$*;*
(b) Q *is increasing in each place with respect to the concordance order* $\prec$*: if* $C_1 \prec C_1'$ *and* $C_2 \prec C_2'$*, then* $Q(C_1,C_2) \le Q(C_1',C_2')$*;*
(c) Q *is invariant under the replacement of copulas by their survival copulas:*

$$Q(C_1,C_2) = Q(\hat{C}_1,\hat{C}_2) \ .$$

Example 3.1. The function Q will be evaluated for all possible pairs of the three fundamental copulas M_2, Π_2 and W_2. As will be seen below (see Section 6), one is at the same time calculating Kendall's τ for all pair of random variable having M_2, or Π_2 or, again, W_2 as their copula. In every case, use will be made of (2.13). Since

$$\partial_1 M_2(s,t) = \begin{cases} 1 & , \quad s < t \ , \\ 0 & , \quad s > t \ , \end{cases} \qquad \text{and} \quad \partial_2 M_2(s,t) = \begin{cases} 0 & , \quad s < t \ , \\ 1 & , \quad s > t \ , \end{cases}$$

one has

$$\partial_1 M_2(s,t) = \mathbf{1}_{(s,1)}(t) = \mathbf{1}_{(0,t)}(s) \ , \tag{3.3}$$

$$\partial_2 M_2(s,t) = \mathbf{1}_{(0,s)}(t) = \mathbf{1}_{(t,1)}(s) \ . \tag{3.4}$$

Similarly,

$$\partial_1 W_2(s,t) = \partial_2 W_2(s,t) = \begin{cases} 0 & , \quad s+t-1 < 0 \ , \\ 1 & , \quad s+t-1 > 0 \ , \end{cases}$$

so that

$$\partial_1 W_2(s,t) = \partial_2 W_2(s,t) = \mathbf{1}_{(t-1,1)}(s) = \mathbf{1}_{(1-s,1)}(t) \ . \tag{3.5}$$

As a consequence of Theorem 2.3 and because of (3.3), one has

$$Q(M_2,M_2) = 4\int_{\mathbb{I}^2} M_2(u,v)\, dM_2(u,v) - 1 = 2 - 1 = 1 \ .$$

Similarly

$$
\begin{aligned}
Q(M_2,\Pi_2) &= Q(\Pi_2,M_2) = 4\int_0^1 u^2\,du - 1 = \frac{4}{3} - 1 = \frac{1}{3}\ ,\\
Q(M_2,W_2) &= Q(W_2,M_2) = 4\int_{1/2}^1 (2u-1)\,du - 1 = \\
Q(W_2,\Pi_2) &= Q(\Pi_2,W_2) = 4\int_0^1 u\,(1-u)\,du = -\frac{1}{3}\ ,\\
Q(W_2,W_2) &= 4\int_0^1 W_2(u,1-u)\,du - 1 = -1\ .
\end{aligned}
$$

Finally $Q(\Pi_2,\Pi_2) = 0$.

Since Q is the difference of two probabilities (see eq. (3.1)), one has

$$Q(C,C) \in [-1,1]$$

for every copula C. In view of Theorem 3.2 (b) and of the previous example, one has also, for every copula C,

$$Q(C,M_2) \in [0,1] \quad , \quad Q(C,W_2) \in [-1,0] \quad , \quad Q(C,\Pi_2) \in [-1/3,1/3]\ .$$

4. Spearman's rank correlation

Charles Spearman, a psychologist, introduced his rank correlation coefficient, known as *Spearman's rho* in 1904 [24].

In our setting, his coefficient is defined as the normalised concordance between a copula C and the independence copula Π_2, so that it measures the discrepancy of concordance property of (X,Y) with respect to an independent coupling (X',Y) belonging to the same Fréchet class.

Definition 4.1. If C is the copula of two continuous random variables X and Y defined on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$, then the population version of *Spearman's rho* of X and Y, which will be denoted indifferently by $\rho_{X,Y}$ or by ρ_C or by $\rho(C)$, is given by

$$\rho_{X,Y} = \rho(C) = 3\,Q(C,\Pi_2) = 12\int_{\mathbb{I}^2} = 12\int_{\mathbb{I}^2} C(u,v)\,du\,dv - 3\ . \tag{4.1}$$

The coefficient 3 that appears in eq. (4.1) represents a normalisation: in fact, thanks to it, ρ_C is in $[-1,1]$ since $Q(C,\Pi_2)$ belongs to the interval $[-1/3,1/3]$, as was seen in the previous Section.

Notice also that

$$\int_{\mathbb{I}^2} \Pi_2(u,v)\,du\,dv = \int_{\mathbb{I}^2} u\,v\,du\,dv = \frac{1}{4}\ ,$$

so that Spearman's rho may be written in the form

$$\rho(X,Y) = 12\int_{\mathbb{I}^2} \{C(u,v) - u\,v\}\,du\,dv = \frac{\displaystyle\int_{\mathbb{I}^2} uv\,dC(u,v) - 1/4}{1/12}\ . \tag{4.2}$$

Below the values of Spearman's rho for a few copulas of Section 2 are reported.

- EFGM copulas (2.2): $\rho(C_\alpha^{\mathbf{EFGM}}) = \alpha/3$;

- Marshall–Olkin copulas (2.3): $\rho(C^{\mathbf{MO}}_{\alpha,\beta}) = (3\alpha\,\beta)\,(2\alpha - \alpha\,\beta + 2\beta)^{-1}$;
- Gaussian bivariate copula (2.5):

$$\rho(C^{\mathbf{Ga}}) = \frac{6}{\pi}\,\arcsin\left(\frac{\rho}{2}\right)\ ;$$

for a proof see [10];
- Frank copulas (2.7):

$$\rho(C^{\mathbf{Frank}}_{\alpha}) = 1 - \frac{12}{\alpha}\,(D_1(\alpha) - D_2(\alpha))\ ,$$

where D_n is the Debye function given, for any natural n by

$$D_n(x) = \frac{n}{x^n}\int_0^x \frac{t^n}{e^t - 1}\,dt\ ; \tag{4.3}$$

- Extreme–value copulas (2.8):

$$\rho(C_A) = 12\int_0^1 \frac{1}{(1 + A(t))^2}\,dt - 3\ .$$

Example 4.1. Spearman's rho $\rho(X_{(1)}, X_{(n)})$ for the copula of the order statistics $X_{(1)}$ and $X_{(n)}$ of the independent and identically distributed random variables X_1, $\cdots$, X_n, a copula that was determined in Example 2.1 is given by

$$\rho(X_{(1)}, X_{(n)}) = 3 - \frac{12n}{\binom{2n}{n}}\sum_{k=0}^{n}\frac{(-1)^k}{2n-k}\binom{2n}{n+k} + 12\frac{(n!)^3}{(3n)!}\,(-1)^n\ . \tag{4.4}$$

It follows from Theorem 4.1 that

$$\begin{aligned}
\rho_n &= \rho(X_{(1)}, X_{(n)}) = 12\int_{\mathbb{I}^2} C_n(u,v)\,du\,dv - 3 = \\
&= 12\int_{A_n}\left\{v - \left[v^{1/n} + (1-u)^{1/n} - 1\right]^n\right\}du\,dv + 12\int_{B_n} v\,du\,dv - 3 = \\
&= 12\int_{A_n\cup B_n} v\,du\,dv - 12\int_{A_n}\left[v^{1/n} + (1-u)^{1/n} - 1\right]^n du\,dv - 3\ ,
\end{aligned}$$

where

$$\begin{aligned}
A_n &:= \left\{(u,v)\in\mathbb{I}^2 : 1 - (1-u)^{1/n} < v^{1/n}\right\}\ , \\
B_n &:= \left\{(u,v)\in\mathbb{I}^2 : 1 - (1-u)^{1/n} > v^{1/n}\right\}\ .
\end{aligned}$$

The substitutions $s = (1-u)^{1/n}$ and $t = v^{1/n}$, and the use of the beta function yield

$$I_n = \int_0^1 dv \int_0^{1-(1-v^{1/n})^n} \left[v^{1/n} + (1-u)^{1/n} - 1\right]^n du =$$

$$= n^2 \int_0^1 t^{n-1}\, dt \int_{1-t}^1 (t+s-1)^n\, s^{n-1}\, ds =$$

$$= n^2 \sum_{k=0}^n \binom{n}{k} \int_0^1 (t-1)^k\, t^{n-1}\, dt \int_{1-t}^1 s^{n-k+n-1}\, ds =$$

$$= n^2 \sum_{k=0}^n \binom{n}{k} \int_0^1 (t-1)^k\, t^{n-1} \left[\frac{1}{2n-k}\left(1-(1-t)^{2n-k}\right)\right] dt =$$

$$= n^2 \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2n-k} \left[\int_0^1 t^{n-1}\,(1-t)^k\, dt - \int_0^1 t^{n-1}\,(1-t)^{2n}\, dt\right] =$$

$$= n^2 \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2n-k} \left\{B(n,k+1) - B(n,2n+1)\right\} .$$

Now

$$B(n,2n+1)\, n^2 \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2n-k} = \frac{(n-1)!(2n)!}{(3n)!}\; n^2 \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2n-k} =$$

$$= \frac{n(2n)!n!}{(3n)!}\; \frac{(-1)^n}{n\binom{2n}{n}} = (-1)^n \frac{(n!)^3}{(3n)!} .$$

Similarly

$$B(n,k+1) = \frac{(n-1)!\,k!}{(n+k)!} ,$$

which proves the assertion.

5. Gini's cograduation index

Spearman's rho of two continuous random variables having C as their copula may be expressed in the form

$$\rho(C) = 12 \int_{\mathbb{I}^2} \left\{(u+v-1)^2 - (u-v)^2\right\} dC(u,v) .$$

In fact, the r.h.s. reduces to

$$12 \int_{\mathbb{I}^2} uv\, dC(u,v) + 3 - 6 \int_{\mathbb{I}^2} u\, dC(u,v) - 6 \int_{\mathbb{I}^2} v\, dC(u,v) ;$$

but, since one integrates with respect to the stochastic measure μ_C induced by the copula C,

$$6 \int_{\mathbb{I}^2} u\, dC(u,v) = 6 \int_0^1 u\, du = 3 ,$$

so that the claimed equality holds. Gini's index replaces the square by the absolute value, and, of course, adopts a different normalizing constant

$$\gamma(C) := 2 \int_{\mathbb{I}^2} \{|u+v-1| - |u-v|\} \, dC(u,v) \, . \tag{5.1}$$

Gini's index may also be expressed in terms of the concordance function.

Theorem 5.1. *Gini's cograduation index for a copula $C \in \mathcal{C}_2$ may be expressed in either of the following forms*

$$\gamma(C) = Q(C, M_2) + Q(C, W_2) =$$

$$= 4 \left\{ \int_0^1 C(u, 1-u) \, du - \int_0^1 (u - C(u,u)) \, du \right\} \, .$$

6. Kendall's rank correlation

Kendall's tau is defined as the concordance between a random pair (X, Y) and an independent copy of it. Formally, it is defined as follows.

Definition 6.1. The Kendall's tau of two continuous random variables X and Y on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is defined by

$$\tau_{X,Y} = Q(C, C) = 4 \int_{\mathbb{I}^2} C(u,v) \, dC(u,v) - 1 \, . \tag{6.1}$$

In view of (2.13), the expression for τ may be written as

$$\tau_{X,Y} = 1 - 4 \int_{\mathbb{I}^2} \partial_1 C(u,v) \, \partial_2 C(u,v) \, du \, dv \, . \tag{6.2}$$

Since Kendall's tau for a pair of continuous random variables X and Y having copula C, depends only on the copula C, so that, in order to stress this dependence, we shall write $\tau_{X,Y} = \tau_C = \tau(C)$.

The result of Example 3.1 yield

$$\tau(M_2) = Q(M_2, M_2) = 1 \quad , \quad \tau(\Pi_2) = 0 \quad , \quad \tau(W_2) = -1 \, .$$

The calculation of Kendall's tau is easier for Archimedean copulas since in this case one has to evaluate the integral of a function of a single variable, the generator of the copula, rather than the integral of functions of two variables.

Theorem 6.1. *Kendall's tau $\tau(C_f)$ for an Archimedean 2–copula C_φ with additive generator f is given by*

$$\tau(C_f) = 1 + 4 \int_0^1 \frac{f(t)}{f'(t)} \, dt \, . \tag{6.3}$$

Below, the values of Kendall's tau for a few copulas are reported.

- EFGM copulas (2.2): $\tau(C_\alpha^{\mathbf{EFGM}}) = 2\,\alpha/9$;
- Marshall–Olkin copulas (2.3): $\tau(C_{\alpha,\beta}^{\mathbf{MO}}) = \dfrac{\alpha\,\beta}{\alpha - \alpha\,\beta + \beta}$;
- Gumbel–Hougaard copulas (2.4): $\tau(C_\alpha^{\mathbf{GH}}) = (\alpha - 1)/\alpha$;
- Mardia–Takahasi–Clayton copulas (2.6): $\tau(\mathcal{C}_\alpha^{\mathbf{MTC}}) = \alpha/(\alpha + 2)$;

- Frank copulas (2.7): $\tau(C_\alpha^{\mathbf{Frank}}) = 1 - \dfrac{4}{\alpha}\,(1 - D_1(\alpha))$, where D_1 is the Debye function of eq. (4.3);
- Extreme–value copulas (2.8):

$$\tau(C_A) = \int_0^1 \frac{t(1-t)}{A(t)}\, dA'(t)\ .$$

The EFGM copulas do not allow the modelling of a large spectrum of dependence among the random variables involved, since $\tau(C_\alpha^{\mathbf{EFGM}}) \in [-2/9, 2/9]$, while $\rho(C_\alpha^{\mathbf{EFGM}}) \in [-1/3, 1/3]$. Moreover, all the tail dependence coefficients, to be met later, see Section 9, associated with it are equal to 0.

Example 6.1. As was done in Example 4.1 for Spearman's rho, one may wish to calculate Kendall's tau $\tau(X_{(1)}, X_{(n)})$ for the copula of the order statistics $X_{(1)}$ and $X_{(n)}$ of the independent and identically distributed random variables $X_1, \cdots, X_n$, whose copula was determined in Example 2.1. However, the calculation is long and tedious; therefore we shall limit ourselves to quoting the results from the literature ([1]). Kendall's tau for the order statistics $X_{(j)}$ and $X_{(k)}$ $(j, k = 1, \cdots, n; j < k)$ is given by

$$\tau(X_{(j)}, X_{(k)}) = 1 - \frac{2\,(n-1)}{2n-1} \binom{n-2}{j-1} \binom{n-j-1}{k-j-1} \times$$

$$\times \sum_{h=0}^{n-k} \sum_{s=0}^{j-1} \binom{n}{s} \binom{n-s}{h} \binom{2n-2}{n-k+h,\, s+j-1}^{-1} . \tag{6.4}$$

Setting $j = 1$ and $k = n$ one has

$$\tau(X_{(1)}, X_{(n)}) = \frac{1}{2n-1}\ , \tag{6.5}$$

a result originally obtained by Schmitz [18, 19].

7. Blomqvist's rank correlation

If in the expression (3.1) for the concordance function one takes $X_2 = x_0$ and $Y = y_0$, namely constant random variables, one obtains

$$Q = \mathbb{P}\left[(X_1 - x_0)\,(Y_1 - y_0) > 0\right] - \mathbb{P}\left[(X_1 - x_0)\,(Y_1 - y_0) < 0\right]\ .$$

Blomqvist [2] chose $x_0 = \widetilde{x}$ and $y_0 = \widetilde{y}$, where $\widetilde{x}$ and $\widetilde{y}$ denote the medians of X and Y, respectively. The corresponding measure of concordance, which is usually called *Blomqvist's beta* or *medial correlation coefficient* is then

$$\begin{aligned} \beta &= \mathbb{P}\left[(X - \widetilde{x})\,(Y - \widetilde{y}) > 0\right] - \mathbb{P}\left[(X - \widetilde{x})\,(Y - \widetilde{y}) < 0\right] = \\ &= 2\,\mathbb{P}\left[(X - \widetilde{x})\,(Y - \widetilde{y}) > 0\right] - 1 = 4\,C(1/2, 1/2) - 1\ . \end{aligned} \tag{7.1}$$

8. On the definition of measure of concordance

In two papers [16, 17] Scarsini introduced axioms for a measure of concordance. These are collected in the following

Definition 8.1. Consider the family Δ_c^2 of continuous 2–dimensional distribution functions and let (X, Y) be a random pair having joint distribution function $H \in \Delta_c^2$. Then a mapping $\kappa : \Delta_c^2 \to [0, 1]$, denoted by $\kappa(X, Y)$ or by $\kappa(H)$ is said to be a *measure of concordance* if

(κ1) κ is defined for every pair of continuous random variables;
(κ2) κ is symmetric: $\kappa(X,Y) = \kappa(Y,X)$;
(κ3) κ is increasing in the sense that if $C_{X,Y} \geq C_{W,Z}$, where $C_{X,Y}$ and $C_{W,Z}$ are the copulas of the pairs (X,Y) and (W,Z), respectively, then $\kappa(X,Y) \geq \kappa(W,Z)$;
(κ4) $\kappa(X,Y) \in [-1,1]$;
(κ5) $\kappa(X,Y) = 0$ if X and Y are independent;
(κ6) $\kappa(-X,Y) = \kappa(X,-Y) = -\kappa(X,Y)$;
(κ7) weak continuity: if $\{(X_n,Y_n)\}_{n\in\mathbb{N}}$ is a sequence of continuous random variables and if the corresponding sequence of copulas converges pointwise to the copula C of the pair (X,Y), then
$$\lim_{n\to+\infty} \kappa(X_n,Y_n) = \kappa(X,Y) .$$

If the inequalities in (κ3) are replaced by $C_{X,Y} > C_{W,Z}$ and $\kappa(X,Y) > \kappa(W,Z)$, respectively, then κ is said to be a *strong* measure of concordance.

In terms of copulas it is possible to reformulate Scarsini's axioms in the following way.

Definition 8.2. A measure of concordance is a mapping $\kappa : \mathcal{C}_2 \to \mathbb{R}$ such that
(κ1) κ is defined for every copula $C \in \mathcal{C}_2$;
(κ2) for every $C \in \mathcal{C}_2$, $\kappa(C) = \kappa(C^T)$;
(κ3) $\kappa(C) \leq \kappa(C')$ whenever $C \leq C'$;
(κ4) $\kappa(C) \in [-1,1]$;
(κ5) $\kappa(\Pi_2) = 0$;
(κ6) $\kappa(C^{\sigma_1}) = \kappa(C^{\sigma_2}) = -\kappa(C)$, where C^{σ_1} and C^{σ_2} are defined in Theorem 2.2;
(κ7) weak continuity: if $C_n \xrightarrow[n\to+\infty]{} C$, then $\lim_{n\to+\infty} \kappa(C_n) = \kappa(C)$.

A measure of concordance is invariant under strictly increasing transformations.

Theorem 8.1. *The following statements hold for a measure of concordance κ:*
(a) *If the continuous functions $f, g : \mathbb{R} \to \mathbb{R}$ are simultaneously strictly increasing or strictly decreasing then $\kappa(f \circ X, g \circ Y) = \kappa(X,Y)$;*
(b) *if $f : \mathbb{R} \to \mathbb{R}$ is continuous and strictly increasing (respectively, decreasing), then*
$$\kappa(X, f\circ X) = 1 \quad , \quad \textit{respectively} \quad \kappa(X, f \circ X) = -1 .$$

Proof. (a) Set $W := f \circ X$ and $Z := g \circ Y$. If f and g are strictly increasing, then, by Theorem 2.2, one has $C_{X,Y} = C_{W,Z}$ so that the assertion follows from (κ3). If, on the other hand, f and g are both strictly decreasing, apply what has just been proved to the strictly increasing functions $-f$ and $-g$. Then, it follows from (κ6) that
$$\kappa(X,Y) = \kappa\left((-f)\circ X, (-g)\circ Y\right) = \\ = (-1)^2\,\kappa(f\circ X, g\circ Y) = \kappa(f \circ X, g \circ Y) .$$
(b) If f is strictly increasing, then $C_{X,f\circ X} = M_2$; then (κ4) yields $\kappa(X, f\circ X) = 1$, while if f is strictly decreasing, then $C_{X,f\circ X} = W_2$ so that $\kappa(X, f \circ X) = -1$. □

The quantities introduced in the previous sections are measures of concordance according to Scarsini's definition.

Theorem 8.2. *Spearman's rho, Gini's index, Kendall's tau and Blomqvist's beta are all measures of concordance.*

Proof. The proof of this result presents no difficulty with one exception: in proving that Kendall's tau satisfies the convergence property ($\kappa 7$) a special argument is needed. Recall that a sequence of finite measures (μ_n) on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is said to *converge vaguely* to μ if

$$\lim_{n\to+\infty} \int f\, d\mu_n = \int f\, d\mu \ ,$$

for every continuous function f with compact support. Then the following statements are equivalent (see [21]):

(a) $\lim_{n\to+\infty} \int_\Omega f_n\, d\mu_n = \int_\Omega f\, d\mu < +\infty$;

(b) $\lim_{a\to+\infty} \sup_{n\in\mathbb{N}} \int_{\{f_n>a\}} f_n\, d\mu_n = 0$.

It suffices now to choose $\Omega = \mathbb{I}^2$, $f_n = C_n$ and $\mu_n = \mu_{C_n}$. □

9. Tail dependence

For bivariate probability distributions it is possible to introduce the notion of tail dependence (see [8]); this is related just to the amount of dependence in the upper right quadrant tail or in the lower left quadrant tail. Their measure is via the tail dependence coefficients that were introduced by Sibuya [22] and which are defined below.

Definition 9.1. Let $\mathbf{X} = (X, Y)$ be a vector with continuous components and let F and G be the d.f.'s of X and Y, respectively. The *upper tail dependence coefficient* λ_U of $\mathbf{X}$ is defined by

$$\lambda_U := \lim_{\substack{t\to 1\\ t<1}} \mathbb{P}\left(X > F^{(-1)}(t) \mid Y > G^{(-1)}(t)\right) \ , \tag{9.1}$$

if this limit exists. $\mathbf{X}$ is said to be *upper tail dependent* when $\lambda_U > 0$ and *upper tail independent* when $\lambda_U = 0$.

The *lower tail dependence coefficient* λ_L of $\mathbf{X}$ is defined by

$$\lambda_L := \lim_{\substack{t\to 0\\ t>0}} \mathbb{P}\left(X \le F^{(-1)}(t) \mid Y \le G^{(-1)}(t)\right) \ , \tag{9.2}$$

if this limit exists. $\mathbf{X}$ is said to be *lower tail dependent* when $\lambda_U > 0$ and *lower tail independent* when $\lambda_L = 0$.

These coefficients, and, hence, tail dependence, depend only on the copula of $\mathbf{X}$.

Theorem 9.1. *Let $\mathbf{X}$ have copula $C \in \mathcal{C}_2$; if the limits of Definition 9.1 exist, and if $\delta_C(t) := C(t,t)$ denotes the diagonal of C; then*

$$\lambda_U = \lim_{\substack{t\to 0\\ t>0}} \frac{1 - 2t + C(t,t)}{1-t} = \lim_{\substack{t\to 0\\ t>0}} \frac{1 - 2t + \delta_C(t)}{1-t} \ , \tag{9.3}$$

and

$$\lambda_L := \lim_{\substack{t\to 0\\ t>0}} \frac{C(t,t)}{t} = \lim_{\substack{t\to 0\\ t>0}} \frac{\delta_C(t)}{t} \ . \tag{9.4}$$

Proof. By recourse to the survival copula of (2.12) one has, since $F(X)$ and $G(Y)$ are uniformly distributed on $\mathbb{I}$,

$$\mathbb{P}\left(X > F^{(-1)}(t) \mid Y > G^{(-1)}(t)\right) = \mathbb{P}\left(F(X) > t \mid G(Y) > t\right) = \\ = \frac{\widehat{C}(t,t)}{1-t} = \frac{1-2t+\delta_C(t)}{1-t} ,$$

which proves (9.3) if the limit exists. In a similar manner one proves (9.4).

□

We report the tail dependence coefficients of a few bivariate copulas

- Marshall–Olkin (2.3): $\lambda_L(C^{\mathbf{MO}}_{\alpha,\beta}) = 0$, $\lambda_U(C^{\mathbf{MO}}_{\alpha,\beta}) = \min\{\alpha, \beta\}$;
- bivariate Gaussian (2.5): $\lambda_L(C^{\mathbf{Ga}}) = \lambda_U(C^{\mathbf{Ga}}) = 0$;
- Gumbel–Hougaard copulas (2.4)): $\lambda_L(C^{\mathbf{GH}}_\alpha) = 0$, $\lambda_U(C^{\mathbf{GH}}_\alpha) = 2 - 2^{1/\alpha}$;
- Mardia–Takahasi–Clayton copulas (2.6):
$$\lambda_L(\mathcal{C}^{\mathbf{MTC}}_\alpha) = \begin{cases} 2^{-1\alpha} & , \quad \alpha > 0 , \\ 0 & , \quad \alpha \in [-1, 0] \end{cases} \quad , \quad \lambda_U(\mathcal{C}l) = 0 ;$$
- Frank copulas (2.7)): $\lambda_L(C^{\mathbf{Frank}}_\alpha) = \lambda_U(C^{\mathbf{Frank}}_\alpha) = 0$ (for more details, see [13, 7]);
- Extreme–value copulas (2.8): $\lambda_L(C_A) = 0$, $\lambda_U(C_A) = 2(1 - A(1/2))$.

10. The Schweizer–Wolff measure of dependence

Let X and Y be continuous random variables and let F and G be their d.f.'s, H their joint d.f., and C their (unique) connecting copula. The graph of C is a surface over the unit square, which is bounded above by the surface $z = M_2(u, v)$, and is bounded below by the surface $z = W_2(u, v)$. If X and Y happen to be independent, then the surface $z = C(u, v)$ is the hyperbolic paraboloid $z = u\,v$. The volume between the surfaces $z = C(u, v)$ and $z = u\,v$ can be used as a measure of dependence. Notice that

$$\int_{\mathbb{I}^2} \{M_2(u,v) - u\,v\}\, d\lambda_2 = \frac{1}{12} \quad \text{and} \quad \int_{\mathbb{I}^2} \{u\,v - W_2(u,v)\}\, d\lambda_2 = \frac{1}{12} ,$$

which, by normalizing, leads to the quantity

(10.1) $$\sigma(X,Y) := 12 \int_{\mathbb{I}^2} |C(u,v) - u\,v|\, d\lambda_2 = 12 \int_{\mathbb{I}^2} |C - \Pi_2|\, d\lambda_2 .$$

This quantity will be called the *Schweizer–Wolff measure of dependence* [20, 26]: it depends only on the copula C of the continuous random variables X and Y and represents the L^1 distance between the surfaces $z = \Pi_2(u, v)$ and $z = C(u, v)$. We proceed to establish its properties.

By a standard change of variables this measure of dependence may be expressed in the form

(10.2) $$\sigma(X,Y) = 12 \int_{\mathbb{I}^2} |H(u,v) - F(u)\,G(v)|\, dF(u)\, dG(v) .$$

Theorem 10.1. *For every copula C, $\sigma(X,Y)$ takes values in $[0,1]$; moreover, $\sigma(X,Y)=1$, if, and only if, either $C=M_2$ or $C=W_2$.*

Proof. Since, by definition it is obvious that $\sigma(X,Y)$ takes values in $[0,1]$, our attention will be devoted to proving the remaining assertion. Fix v_0 in $\mathbb{I}$ and introduce the vertical sections from $\mathbb{I}$ into $[0,v_0]$ defined by

$$C_0(u) := C(u,v_0) \quad , \quad W_0(u) := W_2(u,v_0) \ ,$$

$$P_0(u) := u\,v_0\,m \quad , \quad M_0(u) := M_2(u,v_0) \ .$$

Put

$$A(C_0) := \int_0^1 |C_0(u) - u\,v_0| \; du \ .$$

It will be proved that $A(C_0)$ attains its maximum when either $C=M_2$ or $C=W_2$. Since the graphs of W_0 and M_0 form a parallelogram contained in $\mathbb{I}^2$, of which the graph of P_0 is the diagonal, the equality

$$A(W_0) = A(M_0) \ ,$$

is proved. Now we shall establish the inequality

$$A(C_0) \le A(W_0) = A(M_0) \ . \tag{10.3}$$

Case 1: $C_0 \ge P_0$. Since $C_0 \le M_0$ and the three functions C_0, P_0 and M_0 are all continuous, inequality (10.3) follows at once. A similar argument holds when $C_0 \le P_0$.

Case 2: C_0 is neither everywhere greater nor everywhere smaller than P_0. Because of the continuity of both C_0 and P_0, the set $\{u \in \mathbb{I} : C_0(u) < P_0(u)\}$ is open, and is, therefore the union of at most countably many disjoint open intervals, say $\cup_{i\in I}\,]r_i, s_i[$.

For every $i \in I$, let p_i denote the point $(r_i, C_0(r_i))$ and q_i the point $(s_i, C_0(s_i))$; notice that both p_i and q_i lie on the diagonal of the parallelogram described above. Through p_i and q_i draw parallels to the side of the parallelogram and denote by m_i and n_i the points of intersection. Next, let f_i and g_i the piecewise linear functions defined on $[r_i, s_i]$ and determined by p_i, m_i, q_i and p_i, n_i, q_i, respectively. Since C_0 is increasing and satisfies the Lipschitz condition, the graph of f_i bounds below the portion of the graph of C_0 that lies between r_i and s_i. Therefore, since $C_0 < P_0$ on $]r_i, s_i[$, one has

$$\int_{r_i}^{s_i} |C_0(u) - u\,v_0| \; du \le \int_{r_i}^{s_i} |f_i(u) - u\,v_0| \; du = \tag{10.4}$$

$$= \int_{r_i}^{s_i} |g_i(u) - u\,v_0| \; du \ .$$

Define a function ∂_0 on $\mathbb{I}$ via

$$\partial_0(u) := \begin{cases} C_0(u) & , \quad \text{if } C_0(u) \ge P_0(u) \ , \\ g_i(u) & , \quad \text{if } C_0(u) \le P_0(u) \text{ and } r_i < u < s_i \ . \end{cases}$$

Then

$$A(C_0) = \sum_{i\in I} \int_{r_i}^{s_i} |C_0(u) - u\,v_0| \; du + \sum_{i\in I} \int_{s_i}^{r_{i+1}} |C_0(u) - u\,v_0| \; du \le$$

$$\le A(\partial_0) \le A(M_0) \ ,$$

where the last inequality follows from Case 1.

Finally, since every copula is continuous, the function $v \mapsto A(C_v)$ is also continuous on $\mathbb{I}$. As a consequence, since

$$\int_{\mathbb{I}^2} |C - \Pi_2| \, d\lambda_2 = \int_0^1 A(C_v) \, dv ,$$

inequality (10.3) implies that $\int_{\mathbb{I}^2} |C - \Pi_2| \, d\lambda_2$ attains its maximum if, and only if, either $C = M_2$ or $C = W_2$.

□

Theorem 10.2. *For every pair of continuous random variables X and Y on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\sigma(X,Y) = \sigma(Y,X)$.*

Proof. For every point (u,v) in $\mathbb{I}^2$, one has $C_{X,Y}(u,v) = C_{Y,X}(v,u)$.

□

Theorem 10.3. *For a pair of continuous random variables X and Y on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the following statements are equivalent:*

(a) *$\sigma(X,Y) = 0$;*

(b) *X and Y are independent.*

Proof. Since C and Π_2 are continuous, $\int_{\mathbb{I}^2} |C - \Pi_2| \, d\lambda_2 = 0$ if, and only if, $C = \Pi_2$.

□

Theorem 10.4. *For a pair of continuous random variables X and Y on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the following statements are equivalent:*

(a) *$\sigma(X,Y) = 1$;*

(b) *there exist two strictly monotonic functions $\varphi, \psi : \mathbb{R} \to \mathbb{R}$, such that $X = \varphi \circ Y$ a.e., and $Y = \psi \circ X$ a.e..*

Proof. By Theorem 10.1, $\sigma(X,Y) = 1$, if, and only if, either $C_{X,Y} = M_2$ or $C_{X,Y} = W_2$. But $C_{X,Y} = M_2$, if, and only if, X and Y are a strictly increasing function of each other, while $C_{X,Y} = W_2$ (see, e.g., [5]), if, and only if, X and Y are a strictly decreasing function of each other.

□

Theorem 10.5. *Let X and Y be random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\varphi : \operatorname{Ran} X \to \mathbb{R}$ and $\psi : \operatorname{Ran} Y \to \mathbb{R}$ be strictly monotonic and such that $\varphi \circ X$ and $\psi \circ Y$ have continuous d.f.'s. Then*

$$\sigma(\varphi \circ X, \psi \circ Y) = \sigma(X,Y) . \tag{10.5}$$

Proof. If both φ and ψ are strictly increasing, then the assertion is an immediate consequence of Theorem 2.2 (a).

If φ is strictly increasing while ψ is strictly decreasing, Theorem 2.2 (b) yields, for every $(u,v) \in \mathbb{I}^2$,

$$C_{\varphi\circ X, \psi\circ Y}(u,v) - uv = u - C_{X,Y}(u, 1-v) - uv = u\,(1-v) - C_{X,Y}(u,1-v) ,$$

so that by the change of variables $s = u$, $t = 1 - v$

$$\begin{aligned} \sigma(\varphi \circ X, \psi \circ Y) &= 12 \int_{\mathbb{I}^2} |C_{\varphi\circ X, \psi\circ Y}(u,v) - uv| \, du \, dv = \\ &= 12 \int_{\mathbb{I}^2} |st - C_{X,Y}(s,t)| \, ds \, dt = \sigma(X,Y) . \end{aligned}$$

If φ is strictly decreasing, while ψ is strictly increasing, one has, because of Theorem 10.2 and of what has just been proved

$$\sigma(\varphi \circ X, \psi \circ Y) = \sigma(\psi \circ Y, \varphi \circ X) = \sigma(Y, X) = \sigma(X, Y) .$$

Finally if both φ and ψ are strictly decreasing, then Theorem 2.2 (d) yields, for every $(u, v) \in \mathbb{I}^2$,

$$\begin{aligned} C_{\varphi \circ X, \psi \circ Y}(u, v) - uv &= u + v - 1 + C_{X,Y}(1-u, 1-v) - uv = \\ &= C_{X,Y}(1-u, 1-v) - (1-u)\,(1-v) , \end{aligned}$$

so that by the change of variables $s = 1 - u$, $t = 1 - v$,

$$\begin{aligned} \sigma(\varphi \circ X, \psi \circ Y) &= 12 \int_{\mathbb{I}^2} |C_{\varphi \circ X, \psi \circ Y}(u, v) - uv| \; du\, dv = \\ &= 12 \int_{\mathbb{I}^2} |C_{X,Y}(s, t) - st| \; ds\, dt = \sigma(X, Y) , \end{aligned}$$

which concludes the proof.

□

In order to evaluate $\sigma(X, Y)$ when the joint d.f. of X and Y is bivariate normal, one needs an auxiliary result, a differential relation for the bivariate normal density.

Lemma 10.1. *Let φ be the standard bivariate normal density, namely*

$$\varphi_\rho(x, y) = \frac{1}{2\,\pi\,\sqrt{1-\rho^2}} \; \exp\left\{ -\frac{x^2 - 2\,\rho\, xy + y^2}{2\,(1-\rho^2)} \right\} .$$

Then

(10.6)
$$\frac{\partial \varphi_\rho}{\partial \rho} = \frac{\partial^2 \varphi_\rho}{\partial x\, \partial y} .$$

Proof. The characteristic function f of φ_ρ is given by

$$f(t_1, t_2) = \exp\left(\frac{t_1^2 - 2\,\rho\, t_1 t_2 + t_2^2}{2} \right) .$$

The multivariate inversion formula, e.g., [25, p. 120] yields

$$\begin{aligned} \varphi_\rho(x, y) &= \frac{1}{(2\,\pi)^2} \int_{\mathbb{R}^2} \exp(-i\,(x t_1 + y t_2))\, f(t_1, t_2)\, dt_1\, dt_2 = \\ &= \frac{1}{(2\,\pi)^2} \int_{\mathbb{R}^2} \exp(-i\,(x t_1 + y t_2)) \, \exp\left(\frac{t_1^2 - 2\,\rho\, t_1 t_2 + t_2^2}{2} \right) dt_1\, dt_2 , \end{aligned}$$

whence

$$\begin{aligned} &\frac{\partial \varphi_\rho(x, y)}{\partial \rho} = \frac{\partial^2 \varphi_\rho(x, y)}{\partial x\, \partial y} = \\ &= \frac{1}{(2\,\pi)^2} \int_{\mathbb{R}^2} (-t_1 t_2) \, \exp(-i\,(x t_1 + y t_2)) \, \exp\left(\frac{t_1^2 - 2\,\rho\, t_1 t_2 + t_2^2}{2} \right) dt_1\, dt_2 , \end{aligned}$$

which proves the assertion.

□

From the definition of d.f. one immediately has

Theorem 10.6. *The d.f.* Φ_ρ *of the standard bivariate normal satisfies the relation*

$$\frac{\partial \Phi_\rho(x,y)}{\partial \rho} = \frac{\partial^2 \Phi_\rho(x,y)}{\partial x \partial y} = \varphi_\rho(x,y) \quad , \quad (x,y) \in \mathbb{R}^2 \ . \tag{10.7}$$

Corollary 10.1. *The d.f.* Φ_ρ *of the standard bivariate normal satisfies for all* $(x,y) \in \mathbb{R}^2$*, the inequalities:*

$$\Phi_\rho(x,y) > \Phi_0(x,y) = F_X(x)\, F_Y(y) \quad if\, \rho > 0 \ ,$$
$$\Phi_\rho(x,y) < \Phi_0(x,y) = F_X(x)\, F_Y(y) \quad if\ \rho < 0 \ .$$

Theorem 10.7. *Let* X *and* Y *be random variables whose joint d.f. is standard bivariate normal, with correlation coefficient* ρ*. Then*

$$\sigma(X,Y) = \frac{6}{\pi} \, \arcsin \left(\frac{|\rho|}{2} \right) \ . \tag{10.8}$$

Proof. Consider first the case in which $\rho > 0$, and X and Y have zero mean and standard deviation equal to 1.

By (10.2) and Corollary 10.1, one has

$$\sigma(X,Y) = \frac{12}{2\,\pi} \int_{\mathbb{R}^2} \{F_{X,Y}(x,y) - F_X(x)\, F_Y(y)\} \, \exp\left(-\frac{x^2+y^2}{2}\right) \, dx\, dy \ ,$$

and, in view of Theorem 10.6,

$$\begin{aligned}
\frac{\partial \sigma(X,Y)}{\partial \rho} &= \frac{6}{\pi} \int_{\mathbb{R}^2} \frac{\partial F_{X,Y}(x,y)}{\partial \rho} \, \exp\left(-\frac{x^2+y^2}{2}\right) \, dx\, dy = \\
&= \frac{3}{\pi^2 \sqrt{1-\rho^2}} \int_{\mathbb{R}^2} \exp\left(\frac{x^2 - 2\,\rho\, xy + y^2}{2\,(1-\rho^2)}\right) \, \exp\left(-\frac{x^2+y^2}{2}\right) \, dx\, dy = \\
&= \frac{3}{\pi^2 \sqrt{1-\rho^2}} \int_{\mathbb{R}^2} \exp\left(\frac{(\rho^2-2)\, x^2 + 2\,\rho\, xy + (\rho^2-2)\, y^2}{2\,(1-\rho^2)}\right) \, dx\, dy \ .
\end{aligned}$$

Consider the change of variables

$$s = \left(\frac{4-\rho^2}{2\,(2-\rho^2)}\right)^{1/2} y \quad , \quad t = \frac{(\rho^2-2)\, x + \rho\, y}{(2\,(1-\rho^2)\,(2-\rho^2))^{1/2}} \ ,$$

so that

$$s^2 + t^2 = -\frac{(\rho^2-2)\, x^2 + 2\,\rho\, xy + (\rho^2-2)\, y^2}{2\,(1-\rho^2)} \ .$$

The Jacobian J of this transformation is

$$J = 2\sqrt{\frac{1-\rho^2}{4-\rho^2}}$$

so that

$$\frac{\partial \sigma(X,Y)}{\partial \rho} = \frac{6}{\pi^2 \sqrt{4-\rho^2}} \int_{\mathbb{R}^2} \exp\left(-(s^2+t^2)\right) \, ds\, dt = \frac{6}{\pi \sqrt{4-\rho^2}} \ .$$

Integrating yields

$$\sigma(X,Y) = \frac{6}{\pi} \, \arcsin\left(\frac{\rho}{2}\right) + k \ .$$

The constant k is zero, since for the correlation coefficient ρ of a pair of random variables X and Y with bivariate normal distribution is zero if, and only if, they are independent, namely, if, and only if, $\sigma(X,Y)=0$. Thus, for $\rho \geq 0$,

$$\sigma(X,Y)=\frac{6}{\pi}\ \arcsin\left(\frac{\rho}{2}\right) .$$

Assume now $\rho < 0$. Then the joint d.f. of X and $-Y$ is bivariate normal with correlation coefficient $-\rho > 0$; thus, from the last expression and from Theorem 10.5 one has

$$\sigma(X,Y)=\sigma(X,-Y)=\frac{6}{\pi}\ \arcsin\left(\frac{-\rho}{2}\right)=\frac{6}{\pi}\ \arcsin\left(\frac{|\rho|}{2}\right) .$$

In the general case, let X and Y have means m_1 and m_2 and standard deviations σ_1 and σ_2; by Theorem 10.5 with

$$\varphi(t)=\frac{t-m_1}{\sigma_1} \quad \text{and} \quad \sigma(t)=\frac{t-m_2}{\sigma_2} ,$$

one has

$$\sigma(X,Y)=\sigma(\varphi\circ X,\psi\circ Y)=\frac{6}{\pi}\ \arcsin\left(\frac{|\rho|}{2}\right) ,$$

which completes the proof.

□

Finally we study the behaviour of the Schweizer–Wolff measure of dependence with respect to weak convergence.

Theorem 10.8. *Let $\{(X_n,Y_n)\}_{n\in\mathbb{N}}$ be a sequence of bivariate random vectors in the probability space $(\Omega,\mathcal{F},\mathbb{P})$ and, for every $n\in\mathbb{N}$, let the joint d.f. H_n of (X_n,Y_n) be continuous. If the sequence $\{(X_n,Y_n)\}_n$ converges in law to the random vector (X,Y) with continuous joint d.f. H_0, then*

$$\lim_{n\to+\infty}\sigma\left(X_n,Y_n\right)=\sigma(X,Y) .$$

Proof. Let C_n be the unique copula of H_n, and C_0 the unique copula of H_0. Then, by dominated convergence,

$$\lim_{n\to+\infty}\sigma\left(X_n,Y_n\right)=\lim_{n\to+\infty} 12\int_{\mathbb{I}^2}|C_n-\Pi_2|\ d\lambda_2=12\int_{\mathbb{I}^2}|C_0-\Pi_2|\ d\lambda_2=\sigma(X,Y) ,$$

whence the assertion.

□

It is now possible to list the properties of the Schweizer–Wolff measure of dependence σ:

(SW1) σ is defined for every pair of continuous random variables X and Y defined on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$;
(SW2) σ is symmetric, $\sigma(X,Y)=\sigma(Y,X)$;
(SW3) for every pair of random variables X and Y defined on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$, $\sigma(X,Y)$ belongs to $[0,1]$;
(SW4) $\sigma(X,Y)=0$ if, and only if, X and Y are independent;
(SW5) $\sigma(X,Y)=1$ if either $X=\varphi\circ Y$ or $Y=\psi\circ X$ for some strictly monotonic functions $\varphi,\psi:\mathbb{R}\to\mathbb{R}$;
(SW6) if $\varphi,\psi:\mathbb{R}\to\mathbb{R}$ are strictly monotonic and such that $\varphi\circ X$ and $\psi\circ Y$ have continuous d.f.'s. Then $\sigma(\varphi\circ X,\psi\circ Y)=\sigma(X,Y)$;

(SW7) if the joint distribution of X and Y is a bivariate normal distribution with correlation coefficient ρ, then $\sigma(X,Y) = 6/\pi \arcsin(|\rho|/2)$;

(SW8) if (X_n, Y_n) has joint continuous d.f. H_n and converges in law to the random vector (X,Y) with continuous joint d.f. H_0, then $\sigma(X_n, Y_n) \longrightarrow \sigma(X,Y)$.

The Schweizer–Wolff measure of dependence is the normalized L^1 distance on the standard probability space $(\mathbb{I}^2, \mathcal{B}(\mathbb{I}^2), \lambda_2)$ between the copulas C and Π_2. Of course, other norms can be taken into consideration for a pair of continuous random variables X and Y on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$:

- the L^∞ norm:

$$\sigma_\infty(X,Y) := k_\infty \, \| C - \Pi_2 \|_\infty = k_\infty \sup_{(u,v)\in\mathbb{I}^2} |C(u,v) - \Pi_2(u,v)| \; ; \tag{10.9}$$

- the L^p norm:

$$\sigma_p(X,Y) := k_p \left(\int_{\mathbb{I}^2} |C(u,v) - \Pi_2(u,v)|^p \, d\lambda_2 \right)^{1/p} \; ;$$

here k_∞ and k_p are normalizing constants.

Next σ_∞ will be briefly considered. First, we determine the constant k_∞ of (10.9).

Lemma 10.2. *One has*

$$\sup_{(u,v)\in\mathbb{I}^2} |M_2(u,v) - uv| = 1/4 \, ,$$

and

$$\sup_{(u,v)\in\mathbb{I}^2} |W_2(u,v) - uv| = 1/4 \, .$$

Proof. Since $M_2 \geq \Pi_2$, it suffices to study the difference $M_2 - \Pi_2$. If $u \leq v$, then

$$M_2(u,v) - \Pi_2(u,v) = u - uv \leq u - u^2 = M_2(u,u) - \Pi_2(u,u) \, ,$$

and this quantity assumes its maximum $1/4$ at $u = 1/2$.

Similarly if $u \leq v$, then

$$0 \leq \Pi_2(u,v) - W_2(u,v) \leq uv - u - v + 1 = (1-u)\,(1-v) \leq (1-u)^2$$

that takes its maximum $1/4$ at $u = 1/2$.

□

Thus

$$\sigma_\infty(X,Y) := 4 \, \| C - \Pi_2 \|_\infty = 4 \sup_{(u,v)\in\mathbb{I}^2} |C(u,v) - \Pi_2(u,v)| \; .$$

Since all copulas are continuous on $\mathbb{I}^2$, the supremum is actually a maximum, so that a point (u_0, v_0) exists in $\mathbb{I}^2$ such that

$$\sigma_\infty(X,Y) = 4 \, |C_{X,Y}(u_0, v_0) - \Pi_2(u_0, v_0)| \; .$$

Theorem 10.9. *For all random variables X and Y defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\sigma_\infty(X,Y)$ satisfies properties* (SW1)–(SW6) *and* (SW8).

Now let X and Y have bivariate normal joint d.f.. We shall need the following lemmata.

Lemma 10.3. *For the the standard normal d.f.* Φ_ρ *one has,*

$$\sup_{(x,y)\in\mathbb{R}^2} |\Phi_\rho(x,y) - \Phi(x)\,\Phi(y)| = \sup_{(x,y)\in\mathbb{R}^2} |\Phi_\rho(x,y) - \Phi_0(x,y)| = \\ = |\Phi_\rho(0,0) - \Phi_0(0,0)| \ .$$

Proof. The marginals of Φ_ρ are both standard normal d.f.'s Φ, so that $\Phi_0(0,0) = (\Phi(0))^2$. From Corollary 10.1, one has $\Phi_\rho(x,y) > \Phi_0(x,y)$ for $\rho > 0$, and $\Phi_\rho(x,y) < \Phi_0(x,y)$ for $\rho < 0$ and from Theorem 10.6

$$\frac{\partial \Phi_\rho(x,y)}{\partial \rho} = \frac{\partial^2 \Phi_\rho(x,y)}{\partial x\,\partial y} = \tag{10.10}$$

$$= \frac{1}{2\,\pi\,\sqrt{1-\rho^2}}\ \exp\left(-\frac{1}{2}\ \frac{x^2 - 2\rho xy + y^2}{2\,(1-\rho^2)}\right)\ ,$$

whose maximum is attained at $(0,0)$.

□

Lemma 10.4. *The d.f.* Φ_ρ *satisfies*

$$\Phi_\rho(0,0) = \frac{1}{4} + \frac{1}{2\pi}\,\arcsin(\rho)\ .$$

Proof. One has from (10.10)

$$\frac{\partial \Phi_\rho(0,0)}{\partial \rho} = \frac{1}{2\,\pi\,\sqrt{1-\rho^2}}\ ,$$

whence, integrating

$$\Phi_\rho(0,0) = \frac{1}{2\pi}\ \arcsin(\rho) + k\ .$$

The constant k is determined by considering that

$$\Phi_0(0,0) = (\Phi(0))^2 = 1/4\ ,$$

so that $k = 1/4$, which establishes the assertion.

□

It is now possible to state

Theorem 10.10. *Let* X *and* Y *be random variables with standard bivariate normal joint d.f. and with correlation coefficient* ρ*. Then*

$$\sigma_\infty(X,Y) = \frac{2}{\pi}\ \arcsin(|\rho|)\ . \tag{10.11}$$

In 1959, Alfred Rényi [15] proposed a list of axioms for a measure of dependence R of two random variables X and Y defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The properties of the Schweizer–Wolff measure of dependence are fairly close to Rényi's axioms. The main differences are that Rényi defined R for any pair of random variables X and Y that are not a.e. constant, that in (SW5) the function φ and ψ were assumed to be Borel–measurable rather than strictly monotonic, that in $(SW6)$ φ and ψ were assumed to be one–to–one and Borel measurable; finally, Rényi required $R(X,Y) = |\rho|$, while in (SW7) $\sigma(X,Y)$ is a function of ρ.

The relationship between Spearman's ρ and the Schweizer–Wolff measure of dependence follows from the definition of this latter measure and the expression (4.2):

for a pair of continuous random variables X and Y defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$,

$$|\rho(X,Y)| \leq \sigma(X,Y) \; .$$

References

[1] J. Avérous, C. Genest & S.C. Kochar, *On the dependence structure of order statistics*, J. Multivariate Anal., 94(2005), 159–171.

[2] N. Blomqvist, *On a measure of dependence between two random variables*, Ann. Math. Statist., 21(1950), 593–600.

[3] H.A. David, *Order statistics*, Wiley, New York, 1981.

[4] F. Durante & C. Sempi, *Copula theory: an introduction*, In P. Jaworski, F. Durante, W. Härdle and T. Rychlik, eds., Workshop on *Copula Theory and its Applications*, Lecture Notes in Statistics – Proceedings, Springer, Berlin–Heidelberg, 198(2010), 3–31 (2010).

[5] F. Durante & C. Sempi, *Principles of copula theory*, Chapman & Hall/CRC, Boca Raton FL, 2015.

[6] M.J. Frank, *On the simultaneous associativity of* $F(x, y)$ *and* $x+y-F(x, y)$, Aequationes Math., 19(1979), 194–226.

[7] C. Genest, *Frank's family of bivariate distributions*, Biometrika, 74(1987), 549–555.

[8] H. Joe, *Parametric families of multivariate distributions with given marginals*, J. Multivariate Anal., 46(1993), 262–282.

[9] H. Joe, *Multivariate models and dependence concepts*, Chapman & Hall, London, 1997.

[10] W.H. Kruskal, *Ordinal measures of association*, J. Amer. Statist. Soc., 53(1958), 814–861.

[11] X. Li, P. Mikusiński & M.D. Taylor, *Some integration–by–parts formulas involving* 2–*copulas*, in *C.M. Cuadras*, J. Fortiana, J.A. Rodríguez Lallena, eds., *Distributions with given marginals and statistical modelling*, Kluwer, Dordrecht, 153–159.

[12] C. Meyer, *The bivariate normal copula*, Comm. Statist. Theory Methods, 42(2013), 2402–2422.

[13] R.B. Nelsen, *Properties of a one-parameter family of bivariate distributions with specified marginals*, Comm. Statist. A –Theory Methods, 15(1986), 3277–3285.

[14] R.B. Nelsen, *An introduction to copulas*, 2nd Edition, Springer, New York, 2006.

[15] A. Rényi, *On measures of dependence*, Acta Math. Acad. Sci. Hung., 10(1959), 441–451; see also P. Turán, ed., *Selected papers of Alfréd Rényi*, Akadémiai Kiadó, Budapest, 1(1976), 402–412.

[16] M. Scarsini, *On measures of concordance*, Stochastica, 8(1984), 201–218.

[17] M. Scarsini, *Strong measures of concordance and convergence in probability*, Riv. Mat. Sci. Econom. Social., 7(1984), 39–44.

[18] V. Schmitz, *Copulas and stochastic processes*, Ph.D. Dissertation, Rheinische–Westfälische Technische Hochschule, Aachen, Germany, 2003.

[19] V. Schmitz, *Revealing the dependence structure between* $X_{(1)}$ *and* $X_{(n)}$, J. Statist. Plann. Inference, 123(2004), 41-47.

[20] B. Schweizer & E.F. Wolff, *On nonparametric measures of dependence for random variables*, Ann. Statist., 9(1981), 879–885.

[21] R. Serfozo, *Convergence of Lebesgue integrals with varying measures*, Sankhyā, 44(1982), 380–402.

[22] M. Sibuya, *Bivariate extreme statistics. I*, Ann. Inst. Statist. Math. Tokyo, 11(1960), 195–210.

[23] A. Sklar, *Fonctions de répartition à n dimensions et leurs marges*, Publ. Inst. Statist. Univ. Paris, 8(1959), 229–231.

[24] C. Spearman, *The proof and measurement of association between two things*, Amer. J. Psycol., 15(1904), 72–101.

[25] S.S. Wilks, *Mathematical statistics*, Wiley, New York, 1962.

[26] E.F. Wolff, *Measures of dependence derived from copulas*, Ph.D. Dissertation, University of Massachusetts, Amherst MA, 1977.