Copula Theory and Its Applications

Chapter 10
Copula-Based Measures of Multivariate Association

Friedrich Schmid, Rafael Schmidt, Thomas Blumentritt, Sandra Gaißer and Martin Ruppert

Abstract This chapter constitutes a survey on copula-based measures of multivariate association – i.e. association in a $d$-dimensional random vector $X = (X_1, \ldots, X_d)$ where $d \geq 2$. Some of the measures discussed are multivariate extensions of well-known bivariate measures such as Spearman’s rho, Kendall’s tau, Blomqvist’s beta or Gini’s gamma. Others rely on information theory or are based on $L_p$-distances of copulas. Various measures of multivariate tail dependence are derived by extending the coefficient of bivariate tail dependence. Nonparametric estimation of these measures based on the empirical copula is further addressed.

10.1 Introduction and Definitions

The measurement of bivariate association is well established and measures such as Spearman’s rho, Kendall’s tau, Blomqvist’s beta, Gini’s gamma, Spearman’s...
footrule, and some lesser known are widely used in economics and social sciences. All these measures share one important property: For continuous random variables they are invariant with respect to the two marginal distributions, i.e. they can be expressed as a function of their copula. This property is also known as 'scale-invariance'. Note that not all measures of association satisfy this property, e.g. Pearson’s linear correlation coefficient (see [26] for related discussions).

It is natural to generalize these bivariate copula-based measures to the multivariate case, i.e. to try to measure the amount of association in a \(d\)-dimensional random vector \(X = (X_1, \ldots, X_d)\) where \(d > 2\). This is of interest in many fields of application, e.g. in risk management or in the multivariate analysis of financial asset returns. In a multivariate setting, a number of additional problems and questions occur which are not present in the bivariate case. In dimension \(d = 3\) e.g., three perfectly negatively associated variables do not exist. This is also expressed by the fact that the lower Fréchet-Hoeffding bound of a copula is not a copula itself for \(d \geq 3\), implying that a natural lower bound for the measures does not exist in this case. While desirable analytical properties of a bivariate measure of association are fairly clear and well investigated, this is different for \(d \geq 3\). Indeed, there might be differing views concerning the normalization of the multivariate measure or its preferred behaviour regarding the addition, deletion or transformation of one or several components of \(X = (X_1, \ldots, X_d)\). We do not think that a best measure of multivariate association, satisfying all of the desirable features, has already been found or even exists. We therefore give a survey and a short discussion of some of the measures which have been suggested in the past. There is, however, room for further contributions. Note that we focus on multivariate versions that take into account the multivariate association structure as represented by the \(d\)-dimensional copula of \(X\). We thus do not consider the type of multivariate measures which is given by the average of pairwise bivariate measures with respect to all distinct bivariate margins of the copula. We further do not address measures of complete or functional dependence (see [50, 62, 68, 101]).

Throughout this chapter, we assume that the \(d\)-dimensional random vector \(X\) has distribution function \(F\) with continuous marginal distribution functions \(F_i, i = 1, \ldots, d\). The associated copula \(C\) of \(X\) is thus uniquely defined, which allows for the definition of well-defined copula-based measures of multivariate association. Regarding the case of non-continuous marginal distributions, we refer to Vanden-hende and Lambert [112], Nešlehová [80, 81], Denuit and Lambert [19], Mesfioui and Tajar [75], Genest and Nešlehová [37] as well as Feidt et al. [28].

We further address the statistical estimation of the multivariate measures, which, in our opinion, has not been sufficiently treated in the literature yet but needs further consideration. To do so, we introduce additional notation and definitions in the following, which are not given in Durante and Sempi [22]. Note that, in order to ease notation, we omit the subscript referring to the dimension in the notation of the copula.
Let \((X_j)_{j=1,...,n}\) be a random sample of \(X\) and assume that the distribution function \(F\), the marginal distribution functions \(F_i, i = 1,\ldots,d\), and the copula \(C\) of \(X\) are completely unknown. The marginal distribution functions \(F_i\) are estimated by their empirical counterparts

\[
\hat{F}_{i,n}(x) = \frac{1}{n} \sum_{j=1}^{n} I\{X_{ij} \leq x\} \text{ for } i = 1,\ldots,d \text{ and } x \in \mathbb{R}.
\]

Further, set \(\hat{U}_{ij,n} := \hat{F}_{i,n}(X_{ij})\) for \(i = 1,\ldots,d, j = 1,\ldots,n\), and \(\hat{U}_{j,n} = (\hat{U}_{1,j,n},\ldots,\hat{U}_{d,j,n})\). Since \(\hat{U}_{ij,n} = \frac{1}{n} (\text{rank of } X_{ij} \text{ in } X_{i1},\ldots,X_{in})\), we consider rank order statistics. The copula \(C\) is then estimated by the empirical copula which is defined as

\[
\hat{C}_n(u) = \frac{1}{n} \sum_{j=1}^{n} \prod_{i=1}^{d} I\{\hat{U}_{ij,n} \leq u_i\} \text{ for } u = (u_1,\ldots,u_d) \in [0,1]^d. \tag{10.1}
\]

Empirical copulas were introduced by Rüschendorf [88] and Deheuvels [18]. The asymptotic statistical theory for the related estimators of the multivariate measures is based on the following proposition concerning the asymptotic behaviour of the empirical copula process \(C_n = \sqrt{n} \{\hat{C}_n(u) - C(u)\}\), which has been discussed e.g. by Rüschendorf [88], Gänßler and Stute [33], Fermanian et al. [29], and Tsukahara [110].

**Proposition 10.1.1.** Let \(F\) be a continuous \(d\)-dimensional distribution function with copula \(C\). Under the additional assumption that the \(i\)th partial derivatives \(D_i C(u)\) exist and are continuous for \(i = 1,\ldots,d\), we have

\[
C_n = \sqrt{n} \{\hat{C}_n(u) - C(u)\} \xrightarrow{w} \mathbb{G}_C(u).
\]

Weak convergence takes place in \(L^\infty([0,1]^d)\) and

\[
\mathbb{G}_C(u) = \mathbb{B}_C(u) - \sum_{i=1}^{d} D_i C(u) \mathbb{B}_C(u^{(i)}). \tag{10.2}
\]

The vector \(u^{(i)}\) denotes the vector where all coordinates, except the \(i\)th coordinate of \(u\), are replaced by 1. The process \(\mathbb{B}_C\) is a tight centered Gaussian process on \([0,1]^d\) with covariance function

\[
E\{\mathbb{B}_C(u)\mathbb{B}_C(v)\} = C(u \wedge v) - C(u)C(v),
\]

i.e., \(\mathbb{B}_C\) is a \(d\)-dimensional Brownian bridge.

A similar result can be obtained for the survival function \(\overline{C}\) (cf. [94]). Consider the estimator

\[
\hat{C}_n(u) = \frac{1}{n} \sum_{j=1}^{n} \prod_{i=1}^{d} I\{\hat{U}_{ij,n} > u_i\} \text{ for } u = (u_1,\ldots,u_d) \in [0,1]^d. \tag{10.3}
\]
Under the assumptions of Proposition 10.1.1, weak convergence of the process \( C_n = \sqrt{n} \{ \tilde{C}_n(u) - C(u) \} \) in \( \mathcal{C}^\infty([0,1]^d) \) to the Gaussian process \( G_{\tilde{C}} \) can be established, where \( G_{\tilde{C}} \) has the form

\[
G_{\tilde{C}}(u) = \mathbb{B}_{\tilde{C}}(u) - \sum_{i=1}^{d} D_i \tilde{C}(u) B_C(u^{(i)})
\] (10.4)

with \( d \)-dimensional Brownian bridges \( \mathbb{B}_C, \mathbb{B}_{\tilde{C}} \).

10.2 Aspects of Multivariate Association

The works by Rényi [86], Scarsini [91] as well as Schweizer and Wolff [99] introduce various axioms to characterize bivariate measures of association. However, the derivation of a comparable set of axioms to comprehensively describe multivariate measures of association is not straightforward. We thus concentrate on providing an overview of existing criteria in the literature that are considered to be relevant for distinguishing measures of multivariate association.

A measure of association is a functional

\[
A : \mathcal{C}_d \rightarrow D \subseteq \mathbb{R},
\]

which we denote by \( \mathcal{M}(C) \) or equivalently by \( \mathcal{M}(X) = \mathcal{M}(X_1, \ldots, X_d) \). The following criteria summarize and extend those presented in Wolff [114], Taylor [109], and Dolati and Ubeda-Flores [21]:

**W** Well-definedness: The measure \( \mathcal{M} \) is well-defined for every random vector \( X = (X_1, \ldots, X_d) \) with continuous marginals and is a function of the copula \( C \in \mathcal{C}_d \), i.e. \( \mathcal{M}(X_1, \ldots, X_d) = \mathcal{M}(C) \).

A measure \( \mathcal{M} \) satisfying **W** is invariant with respect to its marginal distributions; in particular, moment assumptions are not required for \( \mathcal{M}(X) \) to be defined.

**P** Invariance with respect to permutations: For every permutation \( \pi \) we have \( \mathcal{M}(X_1, \ldots, X_d) = \mathcal{M}(X_{\pi(1)}, \ldots, X_{\pi(d)}) \).

In general, the measures further vary regarding their range and maximal and minimal arguments. We differentiate the following normalization attributes:

**N** Normalization:

- **N1** If \( \Pi \) is the copula of \( X \) then \( \mathcal{M}(X) = \mathcal{M}(\Pi) = 0 \).
- **N2** If \( \mathcal{M}(X) = 0 \) then \( X \) has copula \( \Pi \).
- **N3** If \( M \) is the copula of \( X \) then \( \mathcal{M}(X) = \mathcal{M}(M) = 1 \).
- **N4** If \( \mathcal{M}(X) = 1 \) then \( X \) has copula \( M \) or \( W \) in dimension \( d = 2 \). If \( \mathcal{M}(X) = 1 \) then \( X \) has copula \( M \) in higher dimension.
If the joint distribution of $X$ is multivariate normal and all pairwise correlations $\rho_{ij}$ of $X_i$ and $X_j$ are either nonnegative or nonpositive, then $\mathcal{M}(X)$ is a strictly increasing function of the absolute value of each of the pairwise correlations.

Note that $\textbf{N4}$ considers the lower Fréchet-Hoeffding bound $W$ in order to cover those measures that are based on notions of distance to independence. It does not impose a lower bound for the measure’s range.

Multivariate measures of association further support different notions of orderings in the set of copulas. Here, we consider the partial order $\preceq$, where $C_1 \preceq C_2$ if and only if $C_1(u) \leq C_2(u)$ for all $u \in [0, 1]^d$. Further, $C_1$ is smaller than $C_2$ according to the concordance (partial) order, denoted by $C_1 \preceq_c C_2$, if and only if $C_1(u) \leq C_2(u)$ and $\overline{C}_1(u) \leq \overline{C}_2(u)$ for all $u \in [0, 1]^d$.

**M Monotonicity and concordance:**
- $\textbf{M1}$ For $I \preceq C_1 \preceq C_2 \preceq M$ we have $\mathcal{M}(C_1) \leq \mathcal{M}(C_2)$.
- $\textbf{M2}$ For $C_1 \preceq C_2$ we have $\mathcal{M}(C_1) \leq \mathcal{M}(C_2)$.
- $\textbf{M3}$ For $C_1 \preceq_c C_2$ we have $\mathcal{M}(C_1) \leq \mathcal{M}(C_2)$.

Note that $\textbf{M3}$ implies $\textbf{M2}$ which itself implies $\textbf{M1}$. Criteria $\textbf{M2}$ and $\textbf{M3}$ are equivalent for bivariate measures, cf. Joe [58]. $\textbf{M1}$ is relevant for all measures relying on some notion of distance between an arbitrary copula and the independence copula. $\textbf{M3}$ is important in the context of measures of concordance which are defined later in this section.

If one or several components of a random vector $X$ are transformed strictly monotonously, then the copula either stays invariant or changes in a well-known way. The behaviour of multivariate measures of association under strictly monotonic transformations of the random vector can be characterized by:

**T Behaviour under transformations:**
- $\textbf{T1}$ For strictly increasing and continuous transformations $I_i$ we have $\mathcal{M}(X_1, \ldots, X_d) = \mathcal{M}(I_1(X_1), \ldots, I_d(X_d))$.
- $\textbf{T2}$ For strictly decreasing and continuous transformations $D_i$ of all components we have $\mathcal{M}(X_1, \ldots, X_d) = \mathcal{M}(D_1(X_1), \ldots, D_d(X_d))$.
- $\textbf{T3}$ For a strictly decreasing and continuous transformation $D_i$ of one arbitrary component $i$ we have $\mathcal{M}(X_1, \ldots, X_d) = \mathcal{M}(X_1, \ldots, D_i(X_i), \ldots, X_d))$.

Since the copulas of $(X_1, \ldots, X_d)$ and $(I_1(X_1), \ldots, I_d(X_d))$ are identical, $\textbf{T1}$ follows from $\textbf{W}$. Wolff [114] points out that $\textbf{T2}$ is equivalent to

$$\mathcal{M}(X_1, \ldots, X_d) = \mathcal{M}(-X_1, \ldots, -X_d),$$  \tag{10.5}$$

independent of the particular choice of transformations. The literature on concordance measures refers to Eq. (10.5) as the Duality axiom. Note that criterion $\textbf{T3}$ implies $\textbf{T2}$ whereas the converse does not hold.
The following criterion is technical and allows to consider sequences of random variables:

**C Continuity:** If \((X_n)_{n \in \mathbb{N}}\) is a sequence of random vectors and corresponding copulas \((C_n)_{n \in \mathbb{N}}\) and if \(\lim_{n \to \infty} C_n(u) = C(u)\) for all \(u \in [0, 1]^d\) and a copula \(C\), then \(\lim_{n \to \infty} \mathcal{M}(C_n) = \mathcal{M}(C)\).

To generalize the bivariate axiom

\[
\mathcal{M}(X, Y) = -\mathcal{M}(-X, Y) = -\mathcal{M}(X, -Y) = \mathcal{M}(-X, -Y),
\]

validity of \(T2\) as well as an additional symmetry property are required. Here, Taylor [109] considers the following: Assume that \(\delta = (\delta_1, \ldots, \delta_d)\) is a vector of independent Rademacher variables, i.e. \(\delta_i \in \{-1; +1\}\) where probability 0.5 is assigned to each value. Furthermore, the random vector \(X\) and \(\delta\) are assumed to be independent. For measures of concordance, Taylor [109] then assumes that \(\mathcal{M}(\delta_1 X_1, \ldots, \delta_d X_d) = 0\). Calculating the conditional expectation given \(X\) of the left-hand side of the latter equation yields the following criterion:

**R Reflection symmetry:** \(\sum_{\varepsilon_1 \in \{-1; +1\}} \cdots \sum_{\varepsilon_d \in \{-1; +1\}} \mathcal{M}(\varepsilon_1 X_1, \ldots, \varepsilon_d X_d) = 0\).

In contrast, Dolati and Übeda-Flores [21] argue that there is no analogous multivariate generalization of Eq. (10.6) and thus do not consider \(R\).

The following criterion relates \((d-1)\)- and \(d\)-dimensional measures of association in order to quantify changes in the measure that are solely caused by the transition to a higher dimension:

**TP Transition property:** For every \(X = (X_1, \ldots, X_d)\) a sequence \((r_d)_{d \geq 3}\) exists, such that \(r_{d-1}\mathcal{M}(X_2, \ldots, X_d) = \mathcal{M}(X_1, X_2, \ldots, X_d) + \mathcal{M}(-X_1, X_2, \ldots, X_d)\).

A measure satisfying the afore listed properties except \(N2, N4, N5\) and \(T3\) is called a measure of concordance. Whether or not \(R\) is required to hold depends on the respective definition of Taylor [109] or Dolati and Übeda-Flores [21]. For further discussions on multivariate measures of concordance, see Joe [58] and Nelsen [78].

The behaviour of multivariate measures of association may differ if an independent component is added to the random vector \(X\) (cf. [30]). This might be of interest in portfolio analysis, when an additional independent asset is incorporated into an existing portfolio.

**A Addition of an independent component:**

**A1** \(\mathcal{M}(X_1, \ldots, X_d) \geq \mathcal{M}(X_1, \ldots, X_d, X_{d+1})\) if \(X_{d+1}\) is independent of \((X_1, \ldots, X_d)\).

**A2** \(\mathcal{M}(X_1, \ldots, X_d) = \mathcal{M}(X_1, \ldots, X_d, X_{d+1})\) if \(X_{d+1}\) is independent of \((X_1, \ldots, X_d)\).

In order to justify the use of sophisticated multivariate measures of association, we need to investigate whether they can be expressed as a function of lower dimensional measures:
10 Copula-Based Measures of Multivariate Association

I Irreducibility: For every dimension $d$ and every copula $C$ the measure $\mathcal{M}(C)$ cannot be written as a function of lower dimensional measures $\{\mathcal{M}(C')\}_{C' \in \mathcal{F}}$, where $\mathcal{F}$ denotes the set of all marginal copulas $C'$ of $C$.

Note that even if $I$ applies, there can be exceptions in particular cases, e.g. for radially symmetric copulas (cf. [94, 114]).

10.3 Multivariate Generalizations of Spearman’s Rho, Kendall’s Tau, Blomqvist’s Beta, and Gini’s Gamma

This section describes how the well-known measures of bivariate association Spearman’s rho, Kendall’s tau, Blomqvist’s beta, and Gini’s gamma can be generalized to the multivariate case. In the bivariate case, these measures are often referred to as measures of concordance since they fulfill the set of axioms given by Scarsini [91] (cf. Sect. 10.2). As shown below, all multivariate versions can solely be expressed in terms of the copula $C$ of the random vector $X$ and satisfy properties W, P, T1, C, and I; further properties are stated separately next. For similar discussions regarding the measure of association Spearman’s footrule we refer to Genest et al. [39] and references therein.

10.3.1 Spearman’s Rho

Spearman’s rank correlation coefficient (or Spearman’s rho) represents one of the best-known measures to quantify the degree of association between two random variables and was first studied by Spearman [106]. For the two random variables $X_1$ and $X_2$ with bivariate distribution function $F$ and continuous univariate margins $F_1$, $F_2$, Spearman’s rho is defined as

$$
\rho(X_1, X_2) = \frac{\text{Cov}(F_1(X_1), F_2(X_2))}{\sqrt{\text{Var}(F_1(X_1))} \sqrt{\text{Var}(F_2(X_2))}}.
$$

Assuming that $X_1$ and $X_2$ have copula $C$, this is equivalent to

$$
\rho(C) = \frac{\int_0^1 \int_0^1 u_1 u_2 dC(u_1, u_2) - \left(\frac{1}{2}\right)^2}{\left(\frac{1}{12}\right)} = 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3
$$

because of $\int_{[0,1]^2} M(u_1, u_2) du_1 du_2 = 1/3$ and $\int_{[0,1]^2} \Pi(u_1, u_2) du_1 du_2 = 1/4$. Thus, $\rho$ can be interpreted as the normalized average difference between the copula $C$ and the independence copula $\Pi$. Several multivariate extensions of Spearman’s rho and their estimation have been discussed in the literature, we mention Ruymgaart and van Zuijlen [89], Wolff [114], Joe [57], Nelsen [76], Stepanova [107], and Schmid.
and Schmidt [94]. Further, Schmid and Schmidt [92] suggest a related class of multivariate measures of tail dependence, cf. Sect. 10.6. Based on Eq. (10.7), the following $d$-dimensional extension of $\rho$ is straightforward

$$
\rho_1(C) = \frac{\int_{[0,1]^d} C(u) du}{\int_{[0,1]^d} M(u) du} = \frac{\int_{[0,1]^d} \Pi(u) du}{\int_{[0,1]^d} \Pi(u) du} = h_\rho(d) \left\{ 2^d \int_{[0,1]^d} C(u) du - 1 \right\},
$$

with $h_\rho(d) = (d+1)/\left\{ 2^d - (d+1) \right\}$. In a similar way, another multivariate version of Spearman's rho can be derived, which is given by

$$
\rho_2(C) = h_\rho(d) \left\{ 2^d \int_{[0,1]^d} \Pi(u) dC(u) - 1 \right\}.
$$

Nelsen [76] further considers the average of the two versions, i.e. $\rho_3 = (\rho_1 + \rho_2)/2$. All three measures satisfy $N_1$, $N_3$, $N_4$, $M_3$, $R$, $TP$, and $A1$. In addition, $T2$ can be verified for $\rho_3$, which, thus represents a multivariate measure of concordance according to Taylor [109]. For $d = 2$, the three versions coincide and reduce to Spearman's rho as given in (10.7). For $d = 3$, Nelsen [76] points out that $\rho_3$ is equal to the average of the pairwise Spearman's rho coefficients, which is, for example, discussed in Kendall [61]. A lower bound for $\rho_i$, $i \in \{1,2,3\}$ is given by

$$
\frac{2^d - (d+1)!}{d! \left\{ 2^d - (d+1) \right\}}, \quad d \geq 2,
$$

see Nelsen [76]. However, to our knowledge, there exist no literature on the best-possible lower bound for $\rho_i$ (see e.g. Úbeda-Flores [111]). Consider further an index set $I \subset \{1,\ldots,d\}$ with cardinality $2 \leq |I| \leq d$ and denote by $C_I$ the $|I|$-dimensional marginal copula of $C$ corresponding to those components $X_i$ of $X$ where $i \in I$. Then, the following relationship between $\rho_1$ and $\rho_2$ holds (cf. Schmid and Schmidt [94]):

$$
\rho_2(C) = \frac{\sum_{k=2}^d (-1)^k h_\rho(d) 2^d}{h_\rho(k) 2^k} \sum_{I \subset \{1,\ldots,d\} \atop |I| = k} \rho_1(C_I).
$$

It immediately follows from this relationship that $\rho_1$ and $\rho_2$ coincide in case the copula $C$ is radially symmetric.

Statistical inference for $\rho_i$, $i = 1,2$, based on the empirical copula is discussed in Schmid and Schmidt [94]. By replacing the copula $C$ with its empirical counterpart $\hat{C}_n$, we obtain the following nonparametric estimators for $\rho_i$, $i = 1,2$:

$$
\rho_1(\hat{C}_n) = h_\rho(d) \left\{ 2^d \int_{[0,1]^d} \hat{C}_n(u) du - 1 \right\} = h_\rho(d) \left\{ \frac{2^d}{n} \sum_{j=1}^n \prod_{i=1}^d (1 - \hat{U}_{ij,n}) - 1 \right\},
$$

$$
\rho_2(\hat{C}_n) = h_\rho(d) \left\{ 2^d \int_{[0,1]^d} \Pi(u) d\hat{C}_n(u) - 1 \right\} = h_\rho(d) \left\{ \frac{2^d}{n} \sum_{j=1}^n \prod_{i=1}^d \hat{U}_{ij,n} - 1 \right\}.
$$
Under the assumptions of the Proposition 10.1.1 (cf. Sect. 10.1), it can be shown that
\[ \sqrt{n} \left( \rho_{i}(\hat{C}_{n}) - \rho_{i}(C) \right) \overset{w}{\rightarrow} Z_{i} \sim N(0, \sigma_{i}^{2}), \quad n \rightarrow \infty, \quad i = 1, 2. \]

The variances are given by
\[
\sigma_{1}^{2} = 2^{2d} h_{p}(d)^2 \int_{[0,1]} \int_{[0,1]} E \left\{ G_{C}(u)G_{C}(v) \right\} dudv, \\
\sigma_{2}^{2} = 2^{2d} h_{p}(d)^2 \int_{[0,1]} \int_{[0,1]} E \left\{ G_{\tilde{C}}(u)G_{\tilde{C}}(v) \right\} dudv,
\]
with the tight Gaussian processes \( G_{C} \) and \( G_{\tilde{C}} \) as stated in Eqs. (10.2) and (10.4). Asymptotic normality of \( \rho_{i} \) can analogously be established based on the weak convergence of the process \((C_{n}, \tilde{C}_{n})\). For an alternative derivation of the asymptotic distribution of similar rank order statistics for Spearman’s rho, see also Stepanova [107]. If the copula \( C \) is radially symmetric, it follows that \( \sigma_{1}^{2} = \sigma_{2}^{2} \). The asymptotic variances can only be explicitly computed for a few copulas of simple form. For example in case of stochastic independence (i.e. \( C = \Pi \)), we obtain (cf. [94])
\[ \sigma_{1}^{2} = \sigma_{2}^{2} = \frac{(d + 1)^{2}(3/4)^{d} - d - 3}{3(1 + d - 2d)^{2}}. \]

As Schmid and Schmidt [92] show, the asymptotic variances can consistently be estimated by a nonparametric bootstrap method otherwise. Tests for stochastic independence based on various multivariate versions of Spearman’s rho with regard to their asymptotic relative efficiency are considered by Stepanova [107] and Quessy [85].

### 10.3.2 Kendall’s Tau

Let \((X_{1}, X_{2})\) and \((Y_{1}, Y_{2})\) be independent and identically distributed random vectors with distribution function \( F \). In the bivariate case, the population version of Kendall’s tau is defined as the probability of concordance minus the probability of discordance (see [60]):
\[ \tau(X_{1}, X_{2}) = P \{(X_{1} - Y_{1})(X_{2} - Y_{2}) > 0\} - P \{(X_{1} - Y_{1})(X_{2} - Y_{2}) < 0\}. \quad (10.8) \]

If \( F \) has the bivariate copula \( C \), this is equal to
\[ \tau(C) = 4 \int_{[0,1]} C(u, v) dC(u, v) - 1, \quad (10.9) \]
see e.g. Nelsen [79]. For (bivariate) Archimedean copulas, Kendall’s tau can directly be calculated from the generator \( \phi_{C} \) of the copula through [35, 36]
\[ \tau(C) = 1 + 4 \int_{0}^{1} \frac{\phi_{C}(t)}{\phi'_{C}(t)} dt. \]
For the relationship between Kendall’s tau and Spearman’s rho in the bivariate case, see Genest and Nešlehová [38] and references therein. Multivariate versions of Kendall’s tau are considered in Nelsen [76, 78], Joe [57], and Taylor [109]. Let \( \mathbf{X} \) and \( \mathbf{Y} \) be two independent \( d \)-dimensional random vectors with distribution function \( F \) and let \( D_j = X_j - Y_j, \ j = 1, \ldots, d \). Joe [57] suggests the following family of generalizations of Kendall’s tau

\[
\tau_1(\mathbf{X}) = \sum_{k=d'}^d w_k P\{(D_1, \ldots, D_d) \in B_{k,d-k}\},
\]

with \( d' = \lceil (d + 1)/2 \rceil \) and \( B_{k,d-k} \) being the subset of \( \mathbf{x} = (x_1, \ldots, x_d) \) in \( \mathbb{R}^d \) with \( k \) positive components and \( d - k \) negative or \( k \) negative components and \( d - k \) positive. Some technical conditions on the coefficients \( w_k \) such that \( \tau_1 \) satisfies \( N1, N3, M3, T2, R, \) and \( TP \) are given in Joe [57] and Taylor [109]. Hence, for certain choices of \( w_k \), the above generalization of Kendall’s tau is a multivariate measure of concordance according to Taylor [109], who also gives an alternative representation of \( \tau_1 \) in terms of the copula \( C \) of \( F \). Note that the family studied by Joe [57] includes both the average pairwise Kendall’s tau and the following generalization, given by

\[
\tau_2(C) = \frac{1}{2^{d-1} - 1} \left\{ 2^d \int_{[0,1]^d} C(u) dC(u) - 1 \right\},
\]

which is also considered in Nelsen [76, 78]. For dimension \( d = 2 \), the latter reduces to Kendall’s tau as given in (10.9). According to Nelsen [76], a lower bound for \( \tau_2 \) is \(-1/(2^{d-1} - 1)\), which is also best possible and attained if at least one of the bivariate margins of the copula \( C \) equals \( W \) as shown by Ubeda-Flores [111]. The measure \( \tau_2 \) equals the average of the pairwise Kendall’s tau for dimension \( d = 3 \) (cf. [76]).

Based on a random sample \( (\mathbf{X}_j)_{j=1,\ldots,n} \) from \( \mathbf{X} \) with distribution function \( F \), the sample version of (10.10) is

\[
\hat{\tau}_1 := \tau_1(\mathbf{X}_1, \ldots, \mathbf{X}_n) = \sum_{k=d'}^d \frac{2w_k}{n(n-1)} \sum_{i<j} I_{B_{k,d-k}}(\mathbf{X}_i - \mathbf{X}_j).
\]

In case \( C = II \), \( \hat{\tau}_1 \) is asymptotically normally distributed. Joe [57] calculates the asymptotic variance of \( \hat{\tau}_1 \) in this case and calculates corresponding asymptotic relative efficiencies for different families of copulas when the \( \hat{\tau}_1 \)’s are used as test statistics for multivariate independence, see also Stepanova [107]. Note that a natural estimator for \( \tau_2 \) is given by

\[
\tau_2(\hat{C}_n) = \frac{1}{2^{d-1} - 1} \left\{ 2^d \int_{[0,1]^d} \hat{C}_n(u) d\hat{C}_n(u) - 1 \right\},
\]
with empirical copula \( \hat{C}_n \). According to Gänßler and Stute [33], \( \tau_2(\hat{C}_n) \) is asymptotically normally distributed for dimension \( d = 2 \); for a discussion regarding \( d \geq 2 \) see Barbe et al. [3]. Other multivariate (sample) versions of Kendall's tau are discussed in Simon [104, 105], Chop and Marden [11], El Maache and Lepage [25], and Taskinen et al. [108], mainly in the context of tests for stochastic independence. For further nonparametric statistical analysis of Kendall's tau and related tests for (serial) independence, see Genest et al. [41] and references therein.

10.3.3 Blomqvist's Beta

Blomqvist [5] suggested a simple measure of association which is commonly referred to as Blomqvist's beta or the medial correlation coefficient. If \( X_1 \) and \( X_2 \) are two continuous random variables with medians \( \bar{x}_1 \) and \( \bar{x}_2 \), the population version of Blomqvist's beta is given by

\[
\beta = P \{(X_1 - \bar{x}_1)(X_2 - \bar{x}_2) > 0\} - P \{(X_1 - \bar{x}_1)(X_2 - \bar{x}_2) < 0\}.
\]

It can be expressed in terms of the copula \( C \) of \( (X_1, X_2) \) via

\[
\beta(C) = 2P \{(X_1 - \bar{x}_1)(X_2 - \bar{x}_2) > 0\} - 1 = 4C(1/2,1/2) - 1
\]

\[
= \frac{C(1/2,1/2) - \Pi(1/2,1/2) + \bar{C}(1/2,1/2) - \bar{\Pi}(1/2,1/2)}{\bar{M}(1/2,1/2) - \Pi(1/2,1/2) + \bar{M}(1/2,1/2) - \bar{\Pi}(1/2,1/2)}.
\]

As Eq. (10.11) implies, Blomqvist's beta can be interpreted as a normalized difference between the copula \( C \) and the independence copula at \( (1/2,1/2) \). Various extensions of Blomqvist's beta to the multivariate case have been considered in Joe [57], Nelsen [78], Taskinen et al. [108], Übeda-Flores [111], and Schmid and Schmidt [93]. The following multivariate version is motivated by Eq. (10.11):

\[
\beta(C) = \frac{C(1/2) - \Pi(1/2) + \bar{C}(1/2) - \bar{\Pi}(1/2)}{\bar{M}(1/2) - \Pi(1/2) + \bar{M}(1/2) - \bar{\Pi}(1/2)}.
\]

\[
= h_\beta(d) \left\{ C(1/2) + \bar{C}(1/2) - 2^{1-d} \right\},
\]

with \( h_\beta(d) := 2^{d-1}/(2^{d-1} - 1) \) and \( 1/2 := (1/2, \ldots, 1/2) \). It satisfies the properties N1, N3, and M3. Übeda-Flores [111] shows that the lower bound \(-1/(2^{d-1} - 1)\), which is attained if at least one of the bivariate margins of \( C \) equals \( W \), is best-possible. Further, \( \beta \) equals the average of pairwise Blomqvist's beta in dimension \( d = 3 \). Note that if the copula \( C \) is radially symmetric (i.e. \( C = \hat{C} \)), the expression in (10.12) reduces to

\[
\frac{2^d C(1/2) - 1}{2^{d-1} - 1},
\]

which coincides with the multivariate version originally introduced in Nelsen [78]. According to Taylor [109], this version also satisfies the properties R and TP. Schmid and Schmidt [93] studied more general extensions of Blomqvist's beta,
which measure the association in the tail region of the copula (cf. Sect. 10.6) and which include $\beta$ as defined in (10.12).

A natural estimator for $\beta$ is obtained by replacing the copula $C$ and the survival function $\overline{C}$ in the defining Eq. (10.12) with their empirical counterparts, i.e.

$$\hat{\beta}_n := \beta(\hat{C}_n) = h_\beta(d) \left\{ \hat{C}_n(1/2) + \hat{C}_n(1/2) - 2^{1-d} \right\},$$

where $\hat{C}_n$ denotes the empirical survival function as defined in Eq. (10.3). Under weak assumptions on the copula $C$ and the survival function $\overline{C}$, Schmid and Schmidt [93] establish asymptotic normality and consistency of $\hat{\beta}_n$. Namely, if the $i$-th partial derivatives $D_iC$ and $D_i\overline{C}$ exist and are continuous at the point $1/2$, we have

$$\sqrt{n} (\beta(\hat{C}_n) - \beta(C)) \overset{w}{\rightarrow} Z \quad \text{with} \quad Z \sim N(0, \sigma^2).$$

The variance $\sigma^2$ is given by $\sigma^2 = h_\beta(d)^2 E[\{G_C(1/2) + G_{\overline{C}}(1/2)\}^2]$ with the tight Gaussian processes $G_C$ and $G_{\overline{C}}$ as stated in Eqs. (10.2) and (10.4). One main advantage of Blomqvist's beta over other copula-based measures such as Spearman's rho or Kendall's tau is that the asymptotic variance of its estimator can explicitly be calculated whenever the copula and its partial derivatives are of explicit form (see Schmid and Schmidt [93] for related examples). For example if $C = \Pi$, we have

$$\sigma^2 = \frac{1}{2^{d-1} - 1}.$$

In case the copula is of more complicated form, it can be shown that a nonparametric bootstrap method can be applied to estimate the asymptotic variance. This makes it possible to use (standardized) Blomqvist's beta as test statistic for testing stochastic independence or more general dependence structures.

### 10.3.4 Gini's Gamma

Another measure of association is Gini's gamma (or Gini's rank association coefficient), which was proposed by Gini [43]. Its population version is quite similar to Spearman's rho, which can be rewritten in the bivariate case as (cf. [78])

$$\rho(C) = 3 \int_{[0,1]^2} \{(u + v - 1)^2 - (u - v)^2\} dC(u, v).$$

Gini's gamma now focuses on absolute values rather than on squares:

$$\gamma(C) = 2 \int_{[0,1]^2} (|u + v - 1| - |u - v|) dC(u, v)$$

$$= 4 \int_{[0,1]^2} \{M(u, v) + W(u, v)\} dC(u, v) - 2,$$

(10.13)
see Nelsen [77, 78]. A multivariate extension of Gini’s gamma has recently been considered by Behboodian et al. [4]. By defining the function \( A(u) = \{M(u) + W(u)\}/2, u \in [0, 1]^d \), with corresponding survival function \( \overline{A} \), the expression in Eq. (10.13) is equal to

\[
\gamma(C) = 4 \int_{[0,1]^2} \{A(u,v) + \overline{A}(u,v)\}dC(u,v) - \int_{[0,1]^2} \{A(u,v) + \overline{A}(u,v)\}d\Pi(u,v),
\]

as \( A(u,v) + \overline{A}(u,v) = 1 - u - v + 2A(u,v) \) for every \((u,v) \in [0,1]^2\). A multivariate version of Gini’s gamma is then defined as

\[
\gamma(C) = \frac{1}{b(d) - a(d)} \left[ \int_{[0,1]^d} \{A(u) + \overline{A}(u)\}dC(u) - a(d) \right], \tag{10.14}
\]

with normalization constants \( a(d) \) and \( b(d) \) of the form

\[
a(d) = \int_{[0,1]^d} \{A(u) + \overline{A}(u)\}d\Pi(u) = \frac{1}{d + 1} + \frac{1}{2(d + 1)!} + \sum_{i=0}^d (-1)^i \binom{d}{i} \frac{1}{2i+1}!
\]

and

\[
b(d) = \int_{[0,1]^d} \{A(u) + \overline{A}(u)\}dM(u) = 1 - \sum_{i=1}^{d-1} \frac{1}{4i}.
\]

It immediately follows from the above definition that \( \gamma = 0 \) if \( C = \Pi \) and \( \gamma = 1 \) if \( C = M \); thus, \( N1 \) and \( N3 \) hold. For dimension \( d = 3 \), \( \gamma \) equals the average of pairwise Gini’s gamma. Another multivariate generalization is discussed by Taylor [109] in the context of multivariate measures of concordance. Behboodian et al. [4] also provide a sample version for \( \gamma \) as defined in (10.14). In the bivariate case, a sample version based on the empirical copula is considered in Nelsen [77] which coincides with the traditional sample version of Gini’s gamma. The latter plays an important role in the context of tests for stochastic independence and has been discussed by many authors. We refer to Genest et al. [39], Cifarelli and Regazzini [13] and references therein (see also [12], who establish asymptotic normality of a generalized class of bivariate statistics including Gini’s gamma under suitable conditions). An asymptotic theory for \( d \geq 3 \) is not yet available to our knowledge.

### 10.4 Information-Based Measures of Multivariate Association

Relative entropy (also known as Kullback-Leibler divergence, see [66, 67]) is a measure of multivariate association that originated from information theory. This section focuses on a solely copula-based representation that is therefore independent of the marginal distributions. We will review theoretical aspects and consider nonparametric estimation techniques.

Joe [54, 56] introduced relative entropy as a measure of multivariate association in a random vector \( X = (X_1, ..., X_d) \). It is defined as
\[ \delta(X) = \int_{\mathbb{R}^d} \log \left[ \frac{f(x)}{\prod_{i=1}^{d} f_i(x_i)} \right] f(x) dx, \quad (10.15) \]

where \( f \) is the density of the distribution of \( X \) (which is assumed to exist) and \( f_i \) are the densities of the respective marginal distributions.

It is easy to prove that

\[ \delta(X) = \delta(C) = \int_{[0,1]^d} \log[c(u)] c(u) du, \]

where \( c \) is the density of the copula \( C \) of \( X \). \( \delta \) therefore does not depend on the marginal distributions of \( X \) but only on its copula via its density \( c \). If a density of \( X \) does not exist \( \delta \) is usually set to infinity and thus satisfies \( W \) and \( P \). It is well known that \( \delta = 0 \) if and only if \( c(u) \equiv 1 \), i.e. if \( C = \Pi \). Therefore properties \( N1 \) and \( N2 \) are satisfied.

The invariance of copulas under increasing and continuous transformations implies \( T1 \), because the respective densities are invariant under these transformations as well. It is also easy to prove that properties \( T2 \) and \( T3 \) as well as \( A2 \) and \( I \) are satisfied. For a sequence of copula densities \( (c_n)_{n \in \mathbb{N}} \) converging uniformly to a copula density \( c \) one can see that \( C \) holds as well.

Relative entropy can be calculated explicitly for selected distributions. For the Gaussian distribution it is given by

\[ \delta(X) = -\frac{1}{2} \log[|\Sigma|], \]

with \( |\Sigma| \) being the determinant of the correlation matrix \( \Sigma \). In case of an equicorrelated Gaussian distribution (where \(-1 < \rho < 1\) and \( \Sigma = \rho (I_d) + (1 - \rho) I_d \)) we have

\[ \delta(X) = -\frac{1}{2} \log \left[ (1 - \rho)^{d-1} (1 + (d - 1) \rho) \right], \quad (10.16) \]

which reduces to \( \delta = -(\log[1 - \rho^2])/2 \) in the bivariate case. One can see that \( N5 \) is satisfied for \( 0 \leq \rho < 1 \) for a general \( d \). As \( \delta \) is \([0, \infty]\)-valued a normalization \( \delta^* \) is introduced by solving Eq. (10.16) for \( \rho \); therefore \( \delta^* = |\rho| \) (in case of an equicorrelated Gaussian copula). For \( d > 2 \) this has to be done numerically; in the bivariate case, the normalization function is given explicitly as \( \delta^* = [1 - \exp(-2\delta)]^{1/2} \). \( N3 \) is satisfied asymptotically for the normalized relative entropy.

\( \delta \) can be calculated not only for the Gaussian, but also for the Student’s \( t \) distribution with \( \nu \) degrees of freedom (see [44, 45]).

If we expand the function \( g(x) = x \log x \) into a Taylor series at the point \( x^* = 1 \), we get under suitable regularity conditions

\[ \delta(C) = \int_{[0,1]^d} g(c(u)) du = \sum_{t=2}^{\infty} \frac{(-1)^t}{t(t-1)} \int_{[0,1]^d} (c(u) - 1)^t du. \]
The integral in the first summand is \( \int_{[0,1]^d} (c(u) - 1)^2 \, du \) and can be regarded as a measure of the deviation of the copula density from the density of the independence copula \( \Pi \). It is easy to see that

\[
\int_{[0,1]^d} (c(u) - 1)^2 \, du = \int_{[0,1]^d} c^2(u) \, du - 1 = \int_{[0,1]^d} \frac{f^2(x)}{\prod_{i=1}^d f_i(x_i)} \, dx - 1
\]

by substituting the densities on \( \mathbb{R}^d \) for the copula density. This is the multivariate version of Pearson’s Phi-Square as given in Joe [56] (see also [82]).

Estimation of \( \delta \) can be based on \( n^{-1} \sum_{j=1}^n \log[c(U_j)] \) or \( \int_{[0,1]^d} \log[c(u)] \, c(u) \, du \). In the latter case we have \( \delta = \int_{[0,1]^d} \log[c(u)] \, c(u) \, du = E_C(\log[c(U)]) \), where \( E_C \) denotes the expectation with respect to the copula \( C \) with corresponding density \( c \). An estimator for \( \delta \) is therefore given in both cases by \( \hat{\delta}_n = n^{-1} \sum_{j=1}^n \log[c(\tilde{U}_{j,n})] \), where \( \tilde{c} \) is an estimate of the copula density \( c \) based on pseudo-observations \( \tilde{U}_{j,n} = (\tilde{U}_{1,j,n}, \ldots, \tilde{U}_{d,j,n}) \) for \( j = 1, \ldots, n \). As copula densities have compact support, conventional kernel density estimators are subject to boundary bias and thus have to be complemented by boundary correction schemes. It would be preferable using estimators that have compact support themselves.

Probably the best known estimator with compact support is the histogram (cf. [100])

\[ \hat{\epsilon}_h(u) = \frac{N_k}{nh^d} \]

for \( u \in B_k \) with hyper-rectangular bins \( B_k \) \((k = 1, \ldots, m; m \in \mathbb{N})\). For the histogram we have the equality \( n^{-1} \sum_{j=1}^n \log[\hat{\epsilon}(U_j)] = \int_{[0,1]^d} \log[\hat{\epsilon}(u)] \, \hat{\epsilon}(u) \, du \) and thus the equivalence of both estimation approaches.

Another possible estimator is the \( k \)-nearest neighbour estimator

\[ \hat{\epsilon}_{\text{knn}}(u) = \frac{k/n}{\epsilon}, \]

where \( \epsilon = (2dk)^d \) and \( dk \) denotes the distance in the maximum norm from \( u \) to its \( k \)-nearest neighbour (cf. [103]). However, this estimator is not restricted to the unit cube especially if \( k \) is large and if \( u \) is near the boundary. We therefore suggest truncating the neighbourhood of \( u \) at the boundary. The modified estimator denoted by \( \hat{\epsilon}_{\text{trunc}} \) differs from \( \hat{\epsilon}_{\text{knn}} \) by the definition of the denominator, which is given for the truncated estimator by

\[
\epsilon = \prod_{i=1}^d \left\{ d_k + d_k 1_{\{u_i - d_k \geq 0\}} (u_i) 1_{\{u_i + d_k \leq 1\}} (u_i) \right\} + u_i 1_{\{u_i - d_k < 0\}} (u_i) + (1 - u_i) 1_{\{u_i + d_k > 1\}} (u_i).
\]

The histogram and the nearest neighbour estimator suffer from the disadvantage of being discontinuous. Additionally, the latter integrates to infinity (cf. [103]).
There are, however, other estimators which combine the properties of continuity and compact support with finite integral such as the Beta estimator, developed by Chen [10]. It is given as
\[
\hat{c}_{\text{beta}}(u) = \frac{1}{n} \sum_{j=1}^{n} \prod_{i=1}^{d} K\left(\hat{U}_{ij,n}, \frac{u_i}{h} + 1, \frac{1-u_i}{h} + 1\right),
\]
where
\[
K(x, \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}
\]
for some \(x \in [0, 1]\) denotes the univariate p.d.f. of the Beta distribution. Sancetta and Satchell [90] proposed using the density of the empirical Bernstein copula as this estimator is itself a copula density; it is given in Bouezmarni et. al. [8] as
\[
\hat{c}_{\text{bsa}}(u) = \frac{1}{n} \sum_{j=1}^{n} K_h(u, \hat{U}_{j,n}),
\]
where
\[
K_h(u, \hat{U}_{j,n}) = h^d \sum_{v_1=0}^{h-1} \ldots \sum_{v_d=0}^{h-1} \prod_{i=1}^{d} \left(\frac{1}{v_i + 1}\right) u_i^{v_i} (1-u_i)^{h-v_i},
\]
with \(B_\nu = \left[\frac{v_1}{h}, \frac{v_1+1}{h}\right] \times \ldots \times \left[\frac{v_d}{h}, \frac{v_d+1}{h}\right]\).

The performance of the different estimators with regard to the unnormalized relative entropy is compared in Blumentritt and Schmid [7]. The results indicate a good performance of the truncated nearest neighbour estimator with respect to bias and standard deviation.

Other simulation studies based on Kullback's and Leibler's original definition (10.15) of \(\delta\) are due to Kraskov et al. [65] and Darbellay and Vajda [15]. Joe [55] as well as Hall and Morton [47] give results for the estimation of the Shannon entropy.

### 10.5 Measures of Multivariate Association Based on \(L_p\)-Distances

Hoeffding [49] was the first to consider measures of association based on a \(L_p\)-type distance between a copula \(C\) and the independence copula \(\Pi\). His work focuses on \(p = 2\) and was extended by Schweizer and Wolff [99] who introduce \(L_1\)- and \(L_\infty\)-based measures of bivariate association. We first outline the multivariate generalizations of these measures and describe their properties. Secondly, we discuss their estimation and asymptotic behaviour.
10.5.1 $\Phi^2$ as a L$_2$-Distance-Based Measure

Gaißer et al. [34] define a generalized multivariate version of Hoeffding’s $\Phi^2$ by

$$L_2^2(C) = \Phi^2(C) := h_2(d) \int_{[0,1]^d} (C(u) - \Pi(u))^2 \, du.$$  

The normalization factor $h_2(d)$ is given by

$$h_2(d) = \left( \int_{[0,1]^d} (M(u) - \Pi(u))^2 \, du \right)^{-1} = \left( \frac{2}{(d+1)(d+2)} - \frac{1}{2d} \frac{d!}{\prod_{i=0} d \left( i + \frac{1}{2} \right)} + \left( \frac{1}{3} \right)^d \right)^{-1}.$$  

The latter explicit expression for $h_2(d)$ is derived in Gaißer et al. [34]. Note that for dimension $d = 2$, $h_2(2) = 90$ and $\Phi^2(C)$ reduces to the (bivariate) measure originally considered by Hoeffding [49]. Extracting the square root, we obtain $L_2(C) = \Phi(C) := + \sqrt{\Phi^2(C)}$. This measure allows for an interpretation as the normalized distance between the copula $C$ and the independence copula $\Pi$ with respect to the $L_2$-norm.

Due to their structure, all $L_p$-distance-based measures share a set of common properties. Irrespective of the particular choice of $p$, the measures satisfy W and P. They further possess the strong property that they are zero if and only if $\Pi$ is the copula of $X$, thus N1 and N2 hold. Normalizing by means of the upper Fréchet-Hoeffding bound, N3 is assured. Consider a multivariate normal random vector $X$ for which all pairwise correlations $\rho_{ij}$ of $X_i$ and $X_j$ are either nonnegative or nonpositive. Analogously to Wolff [114], it can be shown that all $L_p$-distance-based measures are a strictly increasing function of the absolute value of each of the pairwise correlations. Thus N5 is valid. In general, the $L_p$-distance-based measures further satisfy M1, C, I and T1. For dimension $d \geq 3$, T2 usually does not hold except in case the copula $C$ is radially symmetric, i.e. $C = \bar{C}$.

We discuss some important analytical properties of $\Phi^2$ next; analogous results hold for $\Phi$ (the respective proofs are given in Gaißer et al. [34]). The measure satisfies N4. With regard to N5, it is an open problem to determine the explicit form of the function, cf. Schweizer and Wolff [99]. However, in the bivariate case a power series expansion for $\Phi^2$ given $\rho$ is provided by Hoeffding [49].

Regarding property T, $\Phi^2$ satisfies T2 and T3 in dimension $d = 2$. In higher dimensions, $\Phi^2$ is invariant under strictly decreasing transformations of one component $X_i$, if one of the following two conditions holds: the remaining $(d - 1)$ components are either independent, i.e. their copula is $\Pi$, or they are independent of the transformed component.
In the particular case that an independent component $X_{d+1}$ is added to a $d$-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with copula $C$, $\Phi^2(X_1, \ldots, X_{d+1})$ can be expressed as a function of the $d$-dimensional measure:

$$
\Phi^2(X_1, \ldots, X_{d+1}) = \frac{1}{3} \frac{h_2(d+1)}{h_2(d)} \Phi^2(X_1, \ldots, X_d) < \Phi^2(X_1, \ldots, X_d).
$$

Thus, criterion A1 is satisfied, meaning that an independent variable $X_{d+1}$ reduces overall association in the enlarged vector.

Based on a random sample $(\mathbf{X}_j)_{j=1, \ldots, n}$ from $\mathbf{X}$, the estimation of $\Phi^2(C)$ can be performed by replacing the copula $C$ with the empirical copula $\hat{C}_n$:

$$
\Phi^2(\hat{C}_n) = h_2(d) \int_{[0,1]^d} (\hat{C}_n(u) - \Pi(u))^2 \, du
$$

$$
= h_2(d) \left\{ \left( \frac{1}{n} \right)^2 \sum_{j=1}^n \sum_{k=1}^n \prod_{l=1}^d \left( 1 - \max \{ \hat{U}_{ij}, \hat{U}_{ik} \} \right) \right. 

- \frac{2}{n} \left( \frac{1}{2} \right)^d \sum_{j=1}^n \prod_{l=1}^d \left( 1 - \hat{U}_{ij}^2 \right) \left\} + \left( \frac{1}{3} \right)^d \right\}.
$$

The estimate is therefore easy to calculate even for large $d$. A bias reduction for $\Phi^2(\hat{C}_n)$ has been suggested in Gaißer et al. [34]. Simulations have shown that the estimator works well for various copula families. Obviously, we obtain an estimator for the alternative measure $\Phi$ by $\Phi(\hat{C}_n) = +\sqrt{\Phi^2(\hat{C}_n)}$.

The asymptotic theory for $\Phi^2(\hat{C}_n)$ is derived from the asymptotic behaviour of the empirical copula process $\sqrt{n}(\hat{C}_n(u) - C(u))$ as provided by Proposition 10.1.1. Then, asymptotic normality of the estimator $\Phi^2(\hat{C}_n)$ can be derived by means of the functional delta method (see e.g. [113], p. 389). Under the assumptions of Proposition 10.1.1 and the additional presumption that $C \neq \Pi$ it follows that

$$
\sqrt{n} \{ \Phi^2(\hat{C}_n) - \Phi^2(C) \} \xrightarrow{w} Z_{\Phi^2},
$$

where $Z_{\Phi^2} \sim N(0, \sigma_{\Phi^2}^2)$ and

$$
\sigma_{\Phi^2}^2 = \{2h_2(d)\}^2 \int_{[0,1]^d} \int_{[0,1]^d} E \left\{ \{C(u) - \Pi(u)\} C_C(u) C_C(v) \{C(v) - \Pi(v)\} \right\} \, du \, dv.
$$

Regarding the alternative measure $\Phi$ we have

$$
\sqrt{n} (\Phi(\hat{C}_n) - \Phi(C)) \xrightarrow{w} Z_{\Phi}
$$

with $Z_{\Phi} \sim N(0, \sigma_{\Phi}^2)$ and
The proof is given in Gaißer et al. [34]. The above assumption \( C \neq \Pi \) guarantees that the limiting random variable is nondegenerate as implied by the form of the variance \( \sigma_{\Phi^2}^2 \); the limiting behaviour of \( \Phi^2(\hat{C}_n) \) in case \( C = \Pi \) is considered in Gaißer et al. [34].

### 10.5.2 \( \sigma \) as a \( L_1 \)-Distance-Based Measure

Wolff [114] generalizes the \( L_1 \)-distance-based measure of Schweizer and Wolff [99] to the multivariate case. It is defined by

\[
L_1(C) = \sigma(C) := h_1(d) \int_{[0,1]^d} |C(u) - \Pi(u)| \, du,
\]

where the normalizing factor \( h_1(d) \) is given by

\[
h_1(d) = \left( \frac{1}{d+1} - \frac{1}{2^d} \right)^{-1}.
\]

The measure satisfies \( N4 \). With regard to \( N5 \), an explicit form of the function is derived in Schweizer and Wolff [99] for the bivariate case: \( \sigma(C_p) = \frac{6}{\pi} \text{arcsin}(\frac{\sqrt{2}}{2}) \).

Except for taking the absolute value, this functional form matches the one that can be derived for Spearman’s \( \rho \), illustrating that the two measures are closely related.

A similar calculation as before shows that \( \sigma \) satisfies \( A1 \), too:

\[
\sigma(X_1, \ldots, X_{d+1}) = \frac{1}{2} \frac{h_1(d+1)}{h_1(d)} \sigma(X_1, \ldots, X_d) < \sigma(X_1, \ldots, X_d).
\]

The estimation of \( L_1(C) \) has not yet been considered in detail. Various estimators for this measure can be obtained by replacing \( C \) in the defining formulas with the empirical copula \( \hat{C}_n \). However, no explicit expressions (as e.g. for \( \Phi^2(\hat{C}_n) \)) are available and the estimate must be determined numerically, which can be demanding for large dimension \( d \).

### 10.5.3 \( \kappa \) as a \( L_\infty \)-Distance-Based Measure

A \( L_\infty \)-distance-based multivariate measure is derived in Wolff [114] and investigated in detail by Fernández-Fernández and González-Barrios [30]. The measure is defined by

\[
L_\infty(C) = \kappa(C) := h_\infty(d) \sup_{u \in [0,1]^d} |C(u) - \Pi(u)|.
\]
Fernández-Fernández and González-Barrios [30] do not normalize the population version of the measure. We add a normalization factor \( h_\infty(d) \) in order to assure comparability with alternative measures, which is given by

\[
h_\infty(d) := \left( \left( \frac{1}{d} \right)^{\frac{1}{d-1}} \left( 1 - \frac{1}{d} \right) \right)^{-1}.
\]

Wolff [114] proves that the measure satisfies all normalization criteria except for N4. This is due to the fact that there exist other copulas than the upper Fréchet-Hoeffding bound for which the measure attains its maximal value. With regard to N5, an explicit form of the function is derived in Schweizer and Wolff [99] for the bivariate case: \( \kappa(C_{p}) = \frac{2}{\pi} \arcsin(|\rho|) \). With respect to the addition of further components, the measure behaves differently than the measures discussed before. It generally holds that

\[
0 \leq \kappa(X_1, X_2) \leq \kappa(X_1, X_2, X_3) \leq \ldots \leq \kappa(X_1, \ldots, X_d).
\]

In particular, the measure satisfies A2 if an independent component is added to a \( d \)-dimensional random vector \( X \), i.e.

\[
\kappa(X_1, \ldots, X_{d+1}) = \kappa(X_1, \ldots, X_d).
\]

Estimation of \( \kappa(C) \) from a sample \( (X_j)_{j=1,\ldots,n} \) from \( X \) can analogously be performed by replacing all distribution functions with their empirical counterparts:

\[
\kappa(\hat{C}_n) = \frac{\sup_{u \in [0,1]^d} \left| \hat{C}_n(u) - \prod_{j=1}^{d} U_n(u_j) \right|}{\max_{0 \leq i \leq n} \left( \frac{i}{n} - \left( \frac{i}{n} \right)^d \right)},
\]

where \( U_n \) denotes the (univariate) distribution function of a uniformly distributed random variable on the set \( \{ \frac{1}{n}, \ldots, \frac{n}{n} \} \). In order to reduce bias, the independence copula is replaced by its discretized version \( \prod_{j=1}^{d} U_n(u_j) \). Fernández-Fernández and González-Barrios [30] prove a strong law of large numbers for the unnormalized statistic. An explicit asymptotic theory for this estimator is not available.

The measures introduced in this section offer a range of applications, whereas a substantial strand of literature considers tests of stochastic independence: Hoeffding [51] defines a test of independence based on \( \Phi^2 \) in the bivariate case. Blum et al. [6], Genest and Rémillard [40] as well as Genest et al. [42] define related statistics for testing multivariate independence.

### 10.6 Multivariate Tail Dependence

This section gives an overview of various measures of multivariate tail dependence. Here, tail dependence quantifies the degree of dependence in the joint tail of a multivariate distribution function, i.e. the dependence between extreme events. For a
bivariate distribution, tail dependence is commonly defined as the limiting proportion of exceedance of one margin over a certain threshold given that the other margin has already exceeded that threshold. More precisely, the coefficient of lower tail dependence $\lambda_L$ is defined by

$$
\lambda_L(C) := \lim_{u \downarrow 0} \frac{C(u,u)}{u} = \lim_{u \downarrow 0} P(X_1 \leq F_1^{-1}(u) \mid X_2 \leq F_2^{-1}(u)) = \lim_{u \downarrow 0} P(U_1 \leq u \mid U_2 \leq u) \leq \lim_{u \downarrow 0} P(U_2 \leq u \mid U_1 \leq u).
$$

where $X = (X_1, X_2)$ is a bivariate random vector with distribution function $F$ and inverse marginal distribution functions $F_1^{-1}$, $F_2^{-1}$. Further, $U_i = F_i(X_i)$, $i = 1, 2$. Equivalently, the coefficient of upper tail dependence $\lambda_U$ is

$$
\lambda_U(C) := \lim_{u \uparrow 1} \frac{1 - 2u + C(u,u)}{1 - u} = \lim_{u \uparrow 1} P(U_1 > u \mid U_2 > u)
$$

if the above limits exist. Observe that $0 \leq \lambda_L, \lambda_U \leq 1$. We say $C$ is lower (orthant) tail dependent if $\lambda_L > 0$ or is upper (orthant) tail dependent if $\lambda_U > 0$. Similarly $C$ is called lower and upper tail independent if $\lambda_L = 0$ and $\lambda_U = 0$, respectively.

Joe [58] derives the coefficient of tail dependence for various families of bivariate distributions. Tail dependence of elliptically contoured distributions and copulas is discussed in Hult and Lindskog [53], Schmidt [96], Abdous et al. [1], Klüppelberg et al. [63], and Chan and Li [9]. Other copulas are for example considered in Schmidt [97], Li [71, 72], Joe et al. [59], see also reference therein. The natural nonparametric estimator for $\lambda_L$ from a random sample $(X_j)_{j=1}^{n}$ of $X$ is

$$
\hat{\lambda}_{L,n,k} = \frac{C_n(k,k)}{k(n)}
$$

with suitably chosen parameter $k = k(n)$. The statistical properties of $\hat{\lambda}_{L,n,k}$ have been investigated by several authors using techniques from extreme value theory; we mention Huang [52], Ledford and Tawn [70], Dobrić and Schmid [20], Frahm et al. [31], and Schmidt and Stadtmüller [98]. Coles et al. [14] and Draisma et al. [23] investigate the case of tail independence. For an overview and background reading see also Falk et al. [27], and de Haan and Ferreira [16, Chap. 7].

A natural way to model and analyze tail dependence is by considering extreme value distributions which arise as the limiting distribution of linearly normalized (sample) componentwise maxima, as the sample size tends to infinity; we refer to the monograph by de Haan and Ferreira [16] for a detailed treatment. In particular, a $d$-dimensional random vector $X$ with distribution function $F$ is in the domain of attraction of a $d$-dimensional extreme value distribution $G$, if there exist constants $a_{mi} > 0$ and $b_{mi} \in \mathbb{R}$, $i = 1, \ldots, d$, such that for all $(x_1, \ldots, x_d) \in \mathbb{R}^d$

$$
\lim_{m \to \infty} F_m(a_{m1}x_1 + b_{m1}, \ldots, a_{md}x_d + b_{md}) = G(x_1, \ldots, x_d).
$$

The copula function of a $d$-dimensional extreme value distribution $G$ is given by ([84])
where the function $V$ is homogeneous of order $-1$ and called the exponent measure function. For a comprehensive discussion regarding extreme value copulas see Gudendorf and Segers [46]. It can be shown that the following relationship holds between the coefficient of upper tail dependence $\lambda_U$ and a bivariate extreme value distribution $G$ with marginal distribution functions $G_1$ and $G_2$:

$$
\lambda_U = 2 + \log G \left\{ \left( \frac{1}{-\log G_1} \right)^{-1}(1), \left( \frac{1}{-\log G_2} \right)^{-1}(1) \right\}.
$$

Equation (10.19) can be rewritten as follows

$$
\lambda_U = 2 - \lim_{t \to 0} \left\{ 1 - C_F \left( 1 - \frac{1}{t}, 1 - \frac{1}{t} \right) \right\} = 2 + \log \left\{ C_G \left( \frac{1}{e}, \frac{1}{e} \right) \right\} = 2 - V(1, 1),
$$

where $C_F$ and $C_G$ denote the copula of $F$ and $G$. Note that $1 \leq V(1, 1) \leq 2$.

Equations (10.19) and (10.20) yield various possibilities to generalize the coefficient of bivariate tail dependence to a multidimensional tail-dependence measure. For example, the findings of Eq. (10.20) suggest to consider the copula $C_G$ of a multivariate extreme value distribution $G$, which is defined in (10.18), and evaluate it at a particular point such as $(1/e, \ldots, 1/e)$. Alternatively we may consider the multivariate version of the homogeneous function $V$ in (10.18) and evaluate it at $(1, \ldots, 1)$. Appropriate normalization then yields a multivariate measure of tail dependence (or extremal dependence) with values between 0 and 1. Similarly to considering extreme value distributions, alternatively one may consider so-called tail-dependence functions which are e.g. discussed in Huang [52], Schmidt and Stadtmüller [98], de Haan et al. [17], Einmahl et al. [24], Klüppelberg et al. [64], and Joe et al. [59], see also reference therein.

In the following, we focus on the lower tail-dependence coefficient $\lambda_L$, noting however that analogue definitions and results can be established for $\lambda_U$. In particular, the copula $C$ is upper (or lower) tail dependent if and only if the survival copula $\bar{C}$ is lower (or upper) tail dependent.

Suppose again that $X = (X_1, \ldots, X_d)$ is a $d$-dimensional random vector with distribution function $F$ and copula $C$. Set $U_i = F_i(X_i)$. An evident generalization of $\lambda_L$, as defined in the bivariate case (10.17), is given by (cf. [72, 96])

$$
\lambda_{L,I}(C) = \lim_{u \to 0} P(U_j \leq u, j \notin I \mid U_i \leq u, i \in I) = \lim_{u \to 0} \frac{C(u1)}{C(u_{(I)})}
$$

for every $I \subset \{1, \ldots, d\}$, $I \neq \emptyset$ and $C$ is said to be lower tail dependent if $\lambda_{L,I} > 0$ for some $I$. The vector $u_{(I)}$ denotes the vector where all coordinates, except the $i$th coordinate ($i \in I$) of $u1$, are replaced by 1. In the case of lower tail independence, i.e. $\lambda_{d,I} = 0$, the following multivariate measure $\eta_{L,I}$ is useful

$$
C(u1) = P(U_1 \leq u, \ldots, U_d \leq u)
$$
for $u \downarrow 0$. The function $\mathcal{L}(u)$ is slowly varying as $u \downarrow 0$. This type of tail-dependence measure has been considered in Ledford and Tawn [69], Coles et al. [14], and Hefernan [48] in the bivariate case. Corresponding statistical estimation is addressed in Peng [83]. For an alternative multivariate measure of tail dependence of similar type, we refer to Martins and Ferreira [73].

The following multivariate generalization of $\lambda_L$ is considered in Frahm [32]:

$$\lambda_L(C) = \lim_{u \downarrow 0} P \left( \max \{ U_1, \ldots, U_d \} \leq u \mid \min \{ U_1, \ldots, U_d \} \leq u \right) = \lim_{u \downarrow 0} \frac{C(u1)}{1 - \overline{C}(u1)},$$

where $\overline{C}(u, \ldots, u) = P(U_1 > u, \ldots, U_d > u)$ denotes the survival function of $C$. Note that the relationship between the survival copula $\hat{C}$ and the survival function is as follows: $\hat{C}(u, \ldots, u) = \hat{C}(1 - u, \ldots, 1 - u)$, cf. Durante and Sempi [22] for related discussions.

Schmid and Schmidt [93, 95] define multivariate generalizations of $\lambda_L$ which are based on conditional versions of Spearman’s rho and Blomqvist’s beta. Given the following $d$-dimensional conditional version of Spearman’s rho

$$\rho_p(C) = \frac{\int_{[0,p]^d} C(u) \, du - (p^2/2)^d}{p^{d+1}/(d+1) - (p^2/2)^d} \quad \text{with } 0 < p \leq 1,$$

a coefficient of multivariate lower tail dependence $\rho_L$ can be defined by

$$\rho_L(C) := \lim_{p \downarrow 0} \rho_p(C) = \lim_{p \downarrow 0} \frac{d + 1}{p^{d+1}} \int_{[0,p]^d} C(u) \, du$$

in case the limit exists. Obviously $0 \leq \rho_L \leq 1$. A possible estimator for $\rho_L$ is

$$\rho_L(\hat{C}_n) = \rho_n^L(\hat{C}_n)$$

with appropriate value $k = k(n)$, chosen by the statistician, and

$$\rho_p(\hat{C}_n) := \left\{ \frac{1}{n} \sum_{i=1}^d \prod_{j=1}^d (p - \hat{U}_{i,j,n})^+ + \left(\frac{p^2}{2}\right)^d \right\} \left/ \left\{ \frac{p^{d+1}}{d+1} - \left(\frac{p^2}{2}\right)^d \right\} \right\}.$$

Asymptotic normality of $\sqrt{n} \left( \rho_L(\hat{C}_n) - \rho_L(C) \right)$ can be established if $k = k(n) \to \infty$ and $k/n \to 0$ as $n \to \infty$. The asymptotic variance can be estimated using bootstrap techniques.

In a similar spirit, a $d$-dimensional conditional version of Blomqvist’s beta is defined by

$$\beta_{u,v}(C) := h_{u,v}(d) \left[ \{ C(u) + \hat{C}(v) \} - g_{u,v}(d) \right]$$

for $u, v \in [0,1]^d$ where $u \leq 1/2 \leq v$ and normalization is assured by $h_{u,v}(d)$ and $g_{u,v}(d)$. A coefficient of lower tail dependence can now be defined by
\[
\beta_L(C) := \lim_{p \to 0} \beta_{p,1}(C) = \lim_{p \to 0} \frac{C(p1) - p^d}{p + p^d} 
\]
if the limit exists.

Since tail dependence is limit-based, comparisons to the measures introduced in previous sections are only possible with constraints. The tail dependence measures presented generally satisfy \( W, N_1, N_3, T_1, \) and \( I \).

References

10 Copula-Based Measures of Multivariate Association