On the covariance of the asymptotic empirical copula process

Christian Genest [a,*], Johan Segers [b]

[a] *Département de mathématiques et de statistique, Université Laval, 1045, avenue de la Médecine, Québec, Canada G1V 0A6*

[b] *Institut de statistique, biostatistique et sciences actuarielles (ISBA), Université catholique de Louvain, Voie du Roman Pays 20, B–1348 Louvain-la-Neuve, Belgium*

A R T I C L E   I N F O

A B S T R A C T

Conditions are given under which the empirical copula process associated with a random sample from a bivariate continuous distribution has a smaller asymptotic covariance function than the standard empirical process based on observations from the copula. Illustrations are provided and consequences for inference are outlined.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Consider a pair (X, Y) of continuous random variables whose joint and marginal cumulative distribution functions are defined for all $x, y \in \mathbb{R}$ by

$$H(x, y) = \Pr(X \leqslant x, Y \leqslant y), \qquad F(x) = \Pr(X \leqslant x), \qquad G(y) = \Pr(Y \leqslant y),$$

respectively. The transformed variables $U = F(X)$ and $V = G(Y)$ are then uniform on $[0, 1]$ and their joint distribution function, defined at every $u, v \in [0, 1]$ by

$$C(u, v) = \Pr(U \leqslant u, V \leqslant v),$$

is the unique copula C associated with H. The two functions are related through the equation $C(u, v) = H\{F^{-1}(u), G^{-1}(v)\}$, where $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geqslant u\}$ and $G^{-1}(v) = \inf\{y \in \mathbb{R} : G(y) \geqslant v\}$ for all $u, v \in (0, 1)$. Inference on C is of interest, as it characterizes the dependence in the pair (X, Y); see, e.g., [1,2]. In particular, all margin-free concepts and measures of association such as Kendall's tau, Spearman's rho, Blomqvist's beta, Gini's gamma and Spearman's footrule depend only on C.

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample from H and write $U_i = F(X_i), V_i = G(Y_i)$ for all $i \in \{1, \ldots, n\}$. When F and G are known, a natural estimate of C is then given by the empirical distribution function of the sample $(U_1, V_1), \ldots, (U_n, V_n)$, defined at every $u, v \in [0, 1]$ by

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(U_i \leqslant u, V_i \leqslant v).$$

---

* Corresponding author.
E-mail addresses: Christian.Genest@mat.ulaval.ca (C. Genest), Johan.Segers@uclouvain.be (J. Segers).

In fact, standard results from the theory of empirical processes [3] imply that $\mathbb{C}_n = n^{1/2}(C_n - C)$ converges weakly, as $n \to \infty$, to a centered Gaussian process $\mathbb{C}$ on $[0, 1]^2$ with continuous trajectories and covariance function given by

$$\mathrm{cov}\{\mathbb{C}(u, v), \mathbb{C}(s, t)\} = C(u \wedge s, v \wedge t) - C(u, v)C(s, t),$$

for all $u, v, s, t \in [0, 1]$, with $a \wedge b = \min(a, b)$ for arbitrary $a, b \in \mathbb{R}$.

When the margins F and G are unknown, as is generally the case in practice, they can be estimated by their empirical counterparts, F_n and G_n. A surrogate sample from C is then given by the pairs $(\hat{U}_1, \hat{V}_1), \ldots, (\hat{U}_n, \hat{V}_n)$, where $\hat{U}_i = F_n(X_i)$ and $\hat{V}_i = G_n(Y_i)$ for all $i \in \{1, \ldots, n\}$. The corresponding empirical distribution function, defined at every $u, v \in [0, 1]$ by

$$\hat{C}_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(\hat{U}_i \leqslant u, \hat{V}_i \leqslant v),$$

is traditionally called the empirical copula [4], although it is not a copula *stricto sensu*. The function $\hat{C}_n$ provides a rank-based, consistent estimate of C often used in practice for copula model selection and goodness-of-fit purposes; see, e.g., [5,6].

To be specific, let $\dot{C}_1(u, v) = \partial C(u, v)/\partial u$ and $\dot{C}_2(u, v) = \partial C(u, v)/\partial v$ denote the partial derivatives of an arbitrary copula C, known to exist almost everywhere [2, Chapter 2]. Now assume that they exist in fact everywhere and that they are continuous on $(0, 1)^2$. Under these mild regularity conditions, it is then well known [7,8] that the empirical copula process $\hat{\mathbb{C}}_n = n^{1/2}(\hat{C}_n - C)$ converges weakly, as $n \to \infty$, to a centered Gaussian process $\hat{\mathbb{C}}$ defined at every $u, v \in [0, 1]$ by

$$\hat{\mathbb{C}}(u, v) = \mathbb{C}(u, v) - \dot{C}_1(u, v)\,\mathbb{C}(u, 1) - \dot{C}_2(u, v)\,\mathbb{C}(1, v).$$

In fact, it is observed in [8, Theorem 4] that the assumption on the existence and continuity of the partial derivatives is necessary for the empirical copula process to converge on $[0, 1]^2$.

In the copula modeling literature, the difference between $\mathbb{C}$ and $\hat{\mathbb{C}}$ is often interpreted as "the price to pay for the fact that the margins are unknown". This suggests that if F and G were known, it would be preferable to base the inference on C_n rather than on $\hat{C}_n$. It is shown here, perhaps surprisingly, that the opposite is true under weak positive dependence conditions on C. When this happens, procedures based on $\hat{\mathbb{C}}$ are thus more efficient than the analogous procedures based on $\mathbb{C}$.

The key result is stated and illustrated in Section 2, along with a partial extension to the case of negative dependence. In Section 3, circumstances are delineated under which a dependence parameter, say $\theta = \mathcal{T}(C)$, can be estimated more efficiently by a rank-based estimate $\hat{\theta}_n = \mathcal{T}(\hat{C}_n)$ than by the analogous estimate $\theta_n = \mathcal{T}(C_n)$ which exploits the knowledge of the margins. Concluding remarks are given in Section 4, along with a partial multivariate extension of the main result. All technical arguments are collected in the Appendix.

2. Main result

Following [9], suppose that the two continuous random variables X and Y are such that the mappings $t \mapsto \Pr(X \leqslant x \mid Y \leqslant t)$ and $t \mapsto \Pr(Y \leqslant y \mid X \leqslant t)$ are both decreasing in t whatever $x, y \in \mathbb{R}$. These tail monotonicity conditions, jointly referred to as left-tail decreasingness (LTD), imply that the pair (X, Y) satisfies the concept of positive quadrant dependence (PQD). From [10], this means that for all $x, y \in \mathbb{R}$,

$$\Pr(X \leqslant x, Y \leqslant y) \geqslant \Pr(X \leqslant x)\Pr(Y \leqslant y).$$

Both PQD and LTD can be stated in terms of the underlying copula only. As shown, e.g., in [2, Chapter 5], PQD holds if and only if $C(u, v) \geqslant uv$ for all $u, v \in [0, 1]$, while the LTD property is verified if for almost all $u, v \in (0, 1)$,

$$\dot{C}_1(u, v) \leqslant \frac{C(u, v)}{u}, \qquad \dot{C}_2(u, v) \leqslant \frac{C(u, v)}{v}. \tag{1}$$

Many bivariate models with positive dependence meet Conditions (1), including the bivariate Normal, Beta, Gamma, Student and Fisher distributions. Other examples are provided by the Cook–Johnson bivariate Pareto, Burr and logistic distributions, Gumbel's bivariate exponential and logistic distributions, the Ali–Mikhail–Haq bivariate logistics, the Clayton, Frank, Plackett and Raftery families of copulas.

As it turns out, the LTD property implies a dominance relation between the asymptotic covariance functions of the empirical processes $\mathbb{C}$ and $\hat{\mathbb{C}}$. A formal statement of this fact is given below and proved in the Appendix.

Proposition 1. *Suppose that C is an LTD copula whose partial derivatives $\dot{C}_1$ and $\dot{C}_2$ exist everywhere and are continuous on $(0, 1)^2$. Then for all $u, v, s, t \in [0, 1]$,*

$$\mathrm{cov}\{\hat{\mathbb{C}}(u, v), \hat{\mathbb{C}}(s, t)\} \leqslant \mathrm{cov}\{\mathbb{C}(u, v), \mathbb{C}(s, t)\}. \tag{2}$$

Inequality (2) seems to have been intuited in [11] in the context of copula density estimation; a heuristic explanation was offered, but a formal result was neither stated nor proved. Proposition 4.2 in [12] is also a forerunner of Proposition 1 in the case of extreme-value copulas. As shown in [13], the latter are monotone regression dependent in the sense of [10]; this concept of dependence is stronger than the LTD property.
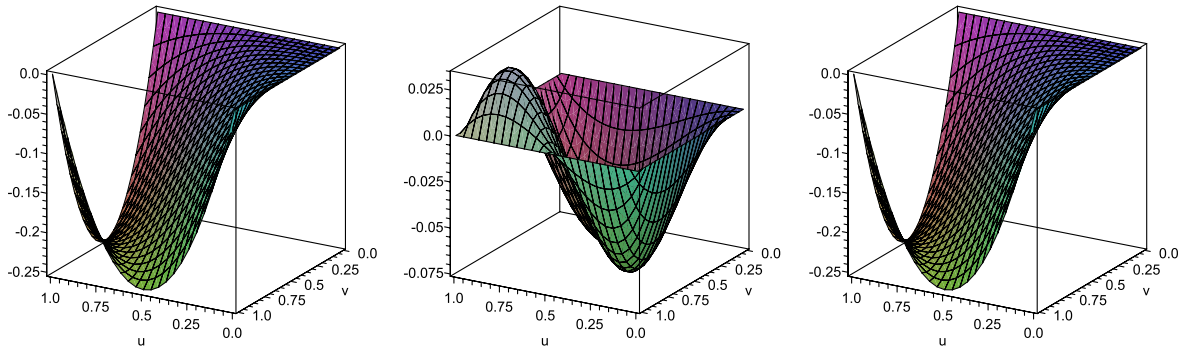
Fig. 1. Graph of $\mathrm{var}\{\hat{\mathbb{C}}(u, v)\} - \mathrm{var}\{\mathbb{C}(u, v)\}$ for the FGM copula with parameter $\theta = 1$ (left) and $\theta = -1$ (right); the middle graph shows the difference between the two surfaces (left – right).

A simple application of Proposition 1 is in the case of independence, where

$$\mathrm{cov}\{\hat{\mathbb{C}}(u, v), \hat{\mathbb{C}}(s, t)\} - \mathrm{cov}\{\mathbb{C}(u, v), \mathbb{C}(s, t)\} = 2uvst - us(v \wedge t) - vt(u \wedge s)$$

is readily seen to be negative for all $u, v, s, t \in [0, 1]$. Here is another illustration.

Example 1. Consider the Farlie–Gumbel–Morgenstern (FGM) copula with parameter $\theta \in [-1, 1]$, which is defined for all $u, v \in [0, 1]$ by $C_\theta(u, v) = uv + \theta uv(1 - u)(1 - v)$. It is easy to check that C_θ is LTD when $\theta \in [0, 1]$; hence Inequality (2) holds for all $u, v, s, t \in [0, 1]$. Analytic expressions for the asymptotic covariances can be derived using Maple but even in this relatively simple case, they are much too long to be displayed here. A graph of $\mathrm{var}\{\hat{\mathbb{C}}(u, v)\} - \mathrm{var}\{\mathbb{C}(u, v)\}$ is plotted in Fig. 1 for the cases $\theta = 1$ (left panel) and $\theta = -1$ (right panel); as the two surfaces look quite similar, their difference (left − right) is also shown in the middle panel.

The right panel of Fig. 1 suggests that Proposition 1 could possibly be extended to cases where the pair (X, Y) is negative quadrant dependent (NQD), i.e., such that $\Pr(X \leqslant x, Y \leqslant y) \leqslant \Pr(X \leqslant x) \Pr(Y \leqslant y)$ for all $x, y \in \mathbb{R}$. A partial finding along these lines is stated next for copulas that are "not too negatively dependent", in the sense that for all $u, v \in (0, 1)$, one has $C(u, v) \leqslant uv$ and

$$\dot{C}_1(u, v) \leqslant 2\frac{C(u, v)}{u}, \qquad \dot{C}_2(u, v) \leqslant 2\frac{C(u, v)}{v}. \tag{3}$$

Proposition 2. *Suppose that C is an NQD copula whose partial derivatives $\dot{C}_1$ and $\dot{C}_2$ exist everywhere and are continuous on $(0, 1)^2$. Further assume that Conditions (3) hold for all $u, v \in (0, 1)$. Then for all $u, v \in [0, 1]$,*

$$\mathrm{var}\{\hat{\mathbb{C}}(u, v)\} \leqslant \mathrm{var}\{\mathbb{C}(u, v)\}. \tag{4}$$

This result, which is proved in the Appendix, is considerably weaker than Proposition 1 because it only yields Inequality (2) in the case $u = s$ and $v = t$. Condition (3) does not correspond to any standard notion of negative dependence and it may be insufficient to establish (2) in full generality. As the following example shows, however, if Condition (3) does not hold, then the variance inequality (4) may be violated.

Example 2. Consider the Gumbel–Barnett copula with parameter $\theta \in (0, 1]$, which is defined for all $u, v \in (0, 1)$ by $C_\theta(u, v) = uv \exp(-\theta \log u \log v) \leqslant uv$. Then C_θ is NQD and for all $u, v \in (0, 1)$, one has both $u\dot{C}_1(u, v)/C(u, v) = 1 - \theta \log v$ and $v\dot{C}_2(u, v)/C(u, v) = 1 - \theta \log u$. It may easily be checked numerically that Condition (3) is not verified and that Inequality (4) fails for $u = v$ close to 0.

3. Consequences for inference

Proposition 1 has intriguing implications for inference about copula-based dependence parameters. To see why, consider the estimation of Blomqvist's medial correlation coefficient using a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from a continuous distribution H. A multivariate version of this problem was recently studied in [14].

When H is bivariate and has underlying copula C, Blomqvist's beta is given by

$$\mathscr{T}_1(C) = -1 + 4C\left(\frac{1}{2}, \frac{1}{2}\right).$$

If the margins F and G of H are known, one can then compute $U_i = F(X_i)$ and $V_i = G(Y_i)$ for each $i \in \{1, \ldots, n\}$ and the pairs $(U_1, V_1), \ldots, (U_n, V_n)$ form a random sample from C. A natural estimator of $\theta = \mathscr{T}_1(C)$ is then given by

$$\theta_n = \mathscr{T}_1(C_n) = -1 + 4C_n\left(\frac{1}{2}, \frac{1}{2}\right) = -1 + \frac{4}{n}\sum_{i=1}^{n}\mathbf{1}\left(U_i \leqslant \frac{1}{2}, V_i \leqslant \frac{1}{2}\right).$$
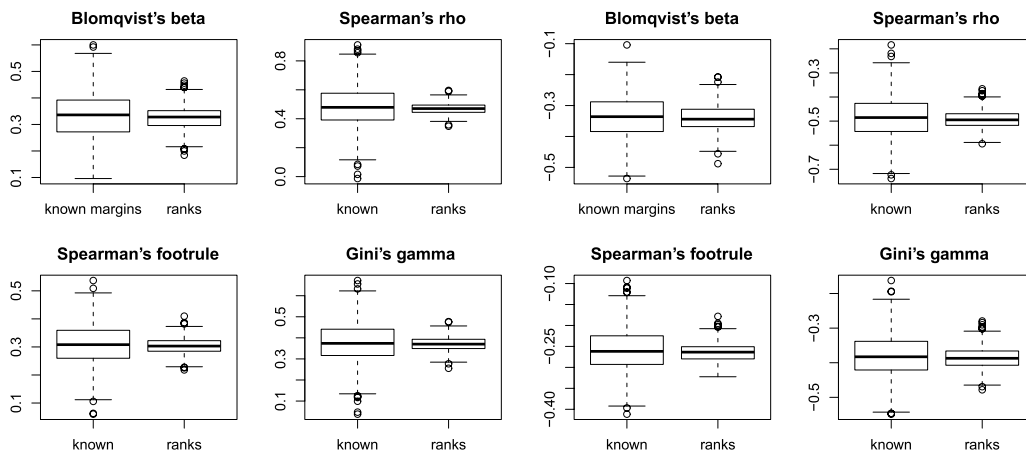
Fig. 2. Boxplots for estimates of Blomqvist's beta, Spearman's rho, Spearman's footrule, and Gini's gamma computed from C_n and $\hat{C}_n$, based on 1000 samples of size 500 of the bivariate Normal distribution with correlation $\rho = -0.5$ (left) and $\rho = 0.5$ (right).

When the margins are unknown, however, it is still possible to estimate θ using

$$\hat{\theta}_n = \mathcal{T}_1(\hat{C}_n) = -1 + 4\hat{C}_n\left(\frac{1}{2}, \frac{1}{2}\right) = -1 + \frac{4}{n} \sum_{i=1}^{n} \mathbf{1}\left(\frac{R_i}{n} \leqslant \frac{1}{2}, \frac{S_i}{n} \leqslant \frac{1}{2}\right),$$

where for fixed $i \in \{1, \ldots, n\}$, $R_i = n\hat{U}_i$ denotes the rank of X_i among $X_1, \ldots, X_n$ and $S_i = n\hat{V}_i$ denotes the rank of Y_i among $Y_1, \ldots, Y_n$.

It follows from the asymptotic behavior of the processes $\mathbb{C}_n$ and $\hat{\mathbb{C}}_n$ that the estimators θ_n and $\hat{\theta}_n$ are asymptotically unbiased and Gaussian. In other words, there exist centered Normal random variables Θ_1 and $\hat{\Theta}_1$ such that, as $n \to \infty$, $n^{1/2}(\theta_n - \theta) \rightsquigarrow \Theta_1$ and $n^{1/2}(\hat{\theta}_n - \theta) \rightsquigarrow \hat{\Theta}_1$, where $\rightsquigarrow$ denotes weak convergence.

Clearly, θ_n cannot be used if F and G are unknown. But if they are known, should θ_n be preferred to $\hat{\theta}_n$? Surprisingly perhaps, Proposition 1 implies that when C satisfies Conditions (1), the rank-based estimator is asymptotically more efficient than its competitor. In other words,

$$\text{var}(\hat{\Theta}_1) = 16\,\text{var}\left\{\hat{\mathbb{C}}\left(\frac{1}{2}, \frac{1}{2}\right)\right\} \leqslant 16\,\text{var}\left\{\mathbb{C}\left(\frac{1}{2}, \frac{1}{2}\right)\right\} = \text{var}(\Theta_1).$$

For example if C is the FGM copula with parameter $\theta \geqslant 0$, one gets

$$\text{var}(\hat{\Theta}_1) = \left(1 + \frac{\theta}{4}\right)\left(1 - \frac{\theta}{4}\right) \leqslant \left(1 + \frac{\theta}{4}\right)\left(3 - \frac{\theta}{4}\right) = \text{var}(\Theta_1).$$

The inequality remains valid for $\theta < 0$, as per Proposition 2. In particular, $\text{var}(\hat{\Theta}) = 1$ and $\text{var}(\Theta) = 3$ at independence. The difference is substantial!

Similar conclusions can be drawn for other copula functionals, such as Spearman's footrule, Spearman's rho and Gini's gamma, respectively defined by

$$\mathcal{T}_2(C) = -2 + 6\int_0^1 C(t, t)\,\mathrm{d}t,$$

$$\mathcal{T}_3(C) = -3 + 12\int_0^1 \int_0^1 C(u, v)\,\mathrm{d}u\,\mathrm{d}v,$$

$$\mathcal{T}_4(C) = -2 + 4\int_0^1 \{C(t, t) + C(t, 1 - t)\}\,\mathrm{d}t.$$

In each case, the rank-based estimator $\hat{\theta} = \mathcal{T}(\hat{C}_n)$ is more efficient asymptotically than the plug-in estimator $\theta_n = \mathcal{T}(C_n)$, as long as C satisfies Conditions (1). The efficiency ratio at independence is 5 for $\mathcal{T}_2$ and $\mathcal{T}_4$, and 7 for $\mathcal{T}_3$.

To illustrate the extent of the improvement in cases of dependence, estimators $\mathcal{T}_i(C_n)$ and $\mathcal{T}_i(\hat{C}_n)$ of parameters $\mathcal{T}_i(C)$ were computed for each $i \in \{1, \ldots, 4\}$ from 1000 random samples of size 500 from the bivariate Normal distribution with correlation $\rho = \pm 0.5$. Boxplots showing the variation in the estimates are shown in Fig. 2. As one can see, the rank-based estimators are preferable in all cases.

This observation can be extended as follows by treating $\mathcal{T}$ as a functional on the space $\mathscr{D}$ of càdlàg functions $\xi : [0, 1]^2 \to \mathbb{R}$, equipped with the sup-norm.

Proposition 3. *Suppose that* $\mathscr{T} : \mathscr{D} \to \mathbb{R}$ *is non-decreasing and Hadamard differentiable at any copula C, tangentially to the subspace* $\mathscr{C} \subset \mathscr{D}$ *of continuous maps. Further assume that as* $n \to \infty$, $n^{1/2}\{\mathscr{T}(C_n) - \mathscr{T}(C)\} \rightsquigarrow \Theta$ *and* $n^{1/2}\{\mathscr{T}(\hat{C}_n) - \mathscr{T}(C)\} \rightsquigarrow \hat{\Theta}$. *If C satisfies the conditions in* (1), *then* $\mathrm{var}(\hat{\Theta}) \leqslant \mathrm{var}(\Theta)$.

The conditions of Proposition 3 are very general and easily verified for many concordance measures [15], including functionals $\mathscr{T}_1$ to $\mathscr{T}_4$. However, they do not extend beyond the comparison of plug-in estimators based on non-decreasing functionals $\mathscr{T}$.

To illustrate this point, consider the functional

$$\mathscr{T}_5(\xi) = 1 + 3 \int_0^1 \{2\xi(t, t) - \xi(t, 1) - \xi(1, t)\}\, \mathrm{d}t.$$

As $\mathscr{T}_5$ fails to be non-decreasing, one cannot conclude that $\hat{\theta}_n = \mathscr{T}_5(\hat{C}_n)$ has greater asymptotic efficiency than $\theta_n = \mathscr{T}_5(C_n)$ as an estimator of $\theta = \mathscr{T}_5(C)$. In other words, if Θ_5 and $\hat{\Theta}_5$ are the weak limits of $n^{1/2}(\theta_n - \theta)$ and $n^{1/2}(\hat{\theta}_n - \theta)$, respectively, one can then have either $\mathrm{var}(\hat{\Theta}_5) \leqslant \mathrm{var}(\Theta_5)$ or $\mathrm{var}(\hat{\Theta}_5) \geqslant \mathrm{var}(\Theta_5)$. For instance, if C is the FGM copula, then both inequalities occur for different choices of the parameter $\theta \geqslant 0$, as can be checked readily using the formulas

$$\mathrm{var}(\hat{\Theta}_5) = \frac{2}{5} + \frac{3}{70}\theta - \frac{11}{150}\theta^2, \qquad \mathrm{var}(\Theta_5) = \frac{1}{2} - \frac{1}{10}\theta - \frac{1}{25}\theta^2.$$

A subtlety arises in that although $\mathscr{T}_5$ is not monotone, its restriction to the class of copulas coincides with Spearman's footrule. This is because if C is a copula, $C(t, 1) = C(1, t) = t$ for all $t \in [0, 1]$. Accordingly, $\mathscr{T}_5(C_n)$ and $\mathscr{T}_5(\hat{C}_n)$ are estimators of $\theta = \mathscr{T}_2(C) = \mathscr{T}_5(C)$ which differ from $\mathscr{T}_2(C_n)$ and $\mathscr{T}_2(\hat{C}_n)$, respectively. Thus if Θ_2 and $\hat{\Theta}_2$ denote the weak limits of $n^{1/2}\{\mathscr{T}_2(C_n) - \theta\}$ and $n^{1/2}\{\mathscr{T}_2(\hat{C}_n) - \theta\}$, respectively, then $\mathrm{var}(\hat{\Theta}_2) \leqslant \mathrm{var}(\Theta_2)$ whereas the same inequality may not hold for $\mathscr{T}_5$. For instance if C is the FGM copula with parameter $\theta \geqslant 0$, one gets

$$\mathrm{var}(\hat{\Theta}_2) = \frac{2}{5} + \frac{3}{70}\theta - \frac{11}{150}\theta^2 \leqslant 2 + \frac{2}{5}\theta - \frac{1}{25}\theta^2 = \mathrm{var}(\Theta_2),$$

in accordance with Proposition 1.

The fact that $\mathrm{var}(\hat{\Theta}_2) = \mathrm{var}(\hat{\Theta}_5)$ is not a coincidence. It occurs because

$$\int_0^1 \hat{C}_n(t, 1)\, \mathrm{d}t = \int_0^1 \hat{C}_n(1, t)\, \mathrm{d}t = \frac{n-1}{2n},$$

so the rank-based estimators $\mathscr{T}_5(\hat{C}_n)$ and $\mathscr{T}_2(\hat{C}_n)$ are asymptotically equivalent. To see that $\mathscr{T}_5(C_n)$ and $\mathscr{T}_2(C_n)$ are not asymptotically equivalent, note that

$$\mathscr{T}_2(C_n) = 4 - \frac{6}{n} \sum_{i=1}^n \max(U_i, V_i)$$

while

$$\mathscr{T}_5(C_n) = 1 - \frac{6}{n} \sum_{i=1}^n \max(U_i, V_i) + \frac{3}{n} \sum_{i=1}^n U_i + \frac{3}{n} \sum_{i=1}^n V_i.$$

Therefore, $n^{1/2}\{\mathscr{T}_5(C_n) - \mathscr{T}_2(C_n)\}$ converges weakly to a non-degenerate centered Normal random variable as $n \to \infty$.

More generally if $\mathscr{T}$ and $\mathscr{T}^*$ are smooth functionals that coincide on the class of copulas, then the rank-based plug-in estimators $\mathscr{T}(\hat{C}_n)$ and $\mathscr{T}^*(\hat{C}_n)$ are asymptotically equivalent, while $\mathscr{T}(C_n)$ and $\mathscr{T}^*(C_n)$ are not necessarily so. Indeed if C_n^* is the checkerboard copula associated with $\hat{C}_n$ (see, e.g., [16]), then for all $u, v \in [0, 1]$,

$$|\hat{C}_n(u, v) - C_n^*(u, v)| \leqslant \frac{1}{n}.$$

Furthermore, as C_n^* is a *bona fide* copula, $\mathscr{T}(C_n^*) = \mathscr{T}^*(C_n^*)$ for every integer $n \geqslant 1$. If the mappings $\mathscr{T}$ and $\mathscr{T}^*$ are differentiable in a neighborhood around C, one may conclude that the difference between $\mathscr{T}(\hat{C}_n)$ and $\mathscr{T}^*(\hat{C}_n)$ is $O_p(1/n)$.

Finally, note that Kendall's tau is an example of a statistic such that $\mathscr{T}(C_n) = \mathscr{T}(\hat{C}_n)$ for all choices of copula C. This is because the concordance or discordance status of pairs (X_i, Y_i) and (X_j, Y_j) is the same, whether it is determined from (U_i, V_i) and (U_j, V_j), or from $(\hat{U}_i, \hat{V}_i)$ and $(\hat{U}_j, \hat{V}_j)$. Accordingly, the use of ranks does not lead to any efficiency gain or loss in estimating Kendall's tau statistic, whatever C is.

4. Discussion

Proposition 1 provides weak and easy-to-check conditions on a copula C which ensure that the asymptotic covariance of the empirical copula process $\hat{\mathbb{C}}_n = n^{1/2}(\hat{C}_n - C)$ is uniformly smaller than the asymptotic covariance of the empirical process $\mathbb{C}_n = n^{1/2}(C_n - C)$ based on a random sample from C. As a consequence, it is shown in Proposition 3 that if a copula-based

dependence parameter $\mathcal{T}(C)$ is expressed in terms of a non-decreasing functional of C, the plug-in rank estimator $\mathcal{T}(\hat{C}_n)$ has a smaller asymptotic variance than the corresponding estimator $\mathcal{T}(C_n)$ which assumes knowledge of the marginal distributions. A numerical illustration further suggests that the gain in efficiency can be substantial, even in finite samples.

These findings are interesting from a theoretical perspective. They provide broad conditions under which inference on copula-based parameters can be improved when raw observations are replaced by ranks. The gains in efficiency come from the estimation of the marginal distributions, which are nuisance parameters in this context. The present results thus provide a new illustration of the paradoxical effect that nuisance parameters sometimes have on the efficiency of estimators; for additional discussion and examples arising in regression, see [17] and references therein.

At the moment, however, the practical implications of the results remain unclear. As mentioned by a referee, no procedures are currently available for testing that data arise from an LTD copula. The only contribution along these lines seems to be [18], where a test of the weaker condition PQD is considered. While the development of an LTD test would clearly be of interest, it is beyond the scope of the present work. But it may be worth pointing out that whatever the underlying copula, Inequality (2) cannot be reversed. For, the limit $\hat{\mathbb{C}}$ of the empirical copula process is a "tucked Brownian sheet" that vanishes everywhere on the border of $[0, 1]^2$, whereas the limit $\mathbb{C}$ of the empirical process is identically zero only on the set $\{(u, v) : u = 0 \text{ or } v = 0 \text{ or } u = v = 1\}$. Thus for an arbitrary copula C and for all $u, s \in (0, 1)$, one has

$$0 = \text{cov}\{\hat{\mathbb{C}}(u, 1), \mathbb{C}(s, 1)\} < \text{cov}\{\mathbb{C}(u, 1), \mathbb{C}(s, 1)\} = u \wedge s - us.$$

By continuity, one must also have $\text{cov}\{\hat{\mathbb{C}}(u, v), \mathbb{C}(s, t)\} < \text{cov}\{\mathbb{C}(u, v), \mathbb{C}(s, t)\}$ when v and t are sufficiently close to 1.

In future work, it would be of interest to find appropriate conditions under which Inequality (2) holds for negatively dependent dependence structures. The conclusion from Proposition 2 is considerably weaker. An extension of Proposition 1 to arbitrary dimension $d \geqslant 2$ would also be valuable but seems difficult. For, the proof detailed in the Appendix uses the fact that for all $u, v, s, t \in [0, 1]$,

$$\text{cov}\{\hat{\mathbb{C}}(u, v), \hat{\mathbb{C}}(s, t)\} - \text{cov}\{\mathbb{C}(u, v), \mathbb{C}(s, t)\} = \sum_{i=1}^{4} A_i(u, v, s, t) - \sum_{i=1}^{4} B_i(u, v, s, t) \tag{5}$$

for specific choices of functions $A_i, B_i, i \in \{1, \ldots, 4\}$. Under the conditions of Proposition 1, one can show that the right-hand side of (5) is non-negative by matching each A_i with a specific B_i, depending on the relative position of $u, v, s, t \in [0, 1]$. In the d-variate case, however, there are d^2 terms of type A but only $2d$ terms of type B, making it impossible to extend the technique used in the proof.

Nevertheless, it is shown in the Appendix that Inequality (2) continues to hold at independence in higher dimensions. This final result is formally stated below. By a continuity argument, one can thus expect that Proposition 1 can be extended to multivariate copulas in an appropriate neighborhood of independence.

Proposition 4. *Let $\mathbb{C}_d$ be the d-variate pinned π-Brownian sheet whose covariance function is given for all $\mathbf{u}, \mathbf{v} \in [0, 1]^d$ by $\text{cov}\{\mathbb{C}_d(\mathbf{u}), \mathbb{C}_d(\mathbf{v})\} = \pi(\mathbf{u} \wedge \mathbf{v}) - \pi(\mathbf{u})\pi(\mathbf{v})$, where for $\mathbf{u} = (u_1, \ldots, u_d)$, $\mathbf{v} = (v_1, \ldots, v_d) \in [0, 1]^d$, $\mathbf{u} \wedge \mathbf{v} = (u_1 \wedge v_1, \ldots, u_d \wedge v_d)$ and $\pi(\mathbf{u}) = u_1 \times \cdots \times u_d$. For arbitrary $\mathbf{u} \in [0, 1]^d$, let also*

$$\hat{\mathbb{C}}_d(\mathbf{u}) = \mathbb{C}_d(\mathbf{u}) - \sum_{k=1}^{d} \dot{\pi}_k(\mathbf{u})\mathbb{C}_d(\mathbf{u}_k),$$

where $\dot{\pi}_k(\mathbf{u}) = \partial \pi(\mathbf{u})/\partial u_k$ and where $\mathbf{u}_k$ denotes a vector whose jth coordinate is u_k if $j = k$ and 1 otherwise. Then for all $\mathbf{u}, \mathbf{v} \in [0, 1]^d$,

$$\text{cov}\{\hat{\mathbb{C}}_d(\mathbf{u}), \hat{\mathbb{C}}_d(\mathbf{v})\} \leqslant \text{cov}\{\mathbb{C}_d(\mathbf{u}), \mathbb{C}_d(\mathbf{v})\}.$$

Acknowledgments

Appendix. Proofs

Proof of Proposition 1. It is obvious from the definition of the limiting process $\hat{\mathbb{C}}$ that Eq. (5) holds with

$$A_1(u, v, s, t) = \dot{C}_1(u, v)\dot{C}_1(s, t)\text{cov}\{\mathbb{C}(u, 1), \mathbb{C}(s, 1)\},$$
$$A_2(u, v, s, t) = \dot{C}_1(u, v)\dot{C}_2(s, t)\text{cov}\{\mathbb{C}(u, 1), \mathbb{C}(1, t)\},$$
$$A_3(u, v, s, t) = \dot{C}_2(u, v)\dot{C}_1(s, t)\text{cov}\{\mathbb{C}(1, v), \mathbb{C}(s, 1)\},$$
$$A_4(u, v, s, t) = \dot{C}_2(u, v)\dot{C}_2(s, t)\text{cov}\{\mathbb{C}(1, v), \mathbb{C}(1, t)\}$$

and

$$B_1(u, v, s, t) = \dot{C}_1(u, v)\mathrm{cov}\{\mathbb{C}(u, 1), \mathbb{C}(s, t)\},$$
$$B_2(u, v, s, t) = \dot{C}_2(u, v)\mathrm{cov}\{\mathbb{C}(1, v), \mathbb{C}(s, t)\},$$
$$B_3(u, v, s, t) = \dot{C}_1(s, t)\mathrm{cov}\{\mathbb{C}(u, v), \mathbb{C}(s, 1)\},$$
$$B_4(u, v, s, t) = \dot{C}_2(s, t)\mathrm{cov}\{\mathbb{C}(u, v), \mathbb{C}(1, t)\}.$$

By symmetry, it can be assumed without loss of generality that $u \leqslant s$. Two cases must be distinguished, according to whether $v \leqslant t$ or $v > t$. If $v \leqslant t$, then

$$A_1(u, v, s, t) = \dot{C}_1(u, v)\,\dot{C}_1(s, t)\,u(1 - s),$$
$$A_2(u, v, s, t) = \dot{C}_1(u, v)\,\dot{C}_2(s, t)\{C(u, t) - ut\},$$
$$A_3(u, v, s, t) = \dot{C}_2(u, v)\,\dot{C}_1(s, t)\{C(s, v) - sv\},$$
$$A_4(u, v, s, t) = \dot{C}_2(u, v)\,\dot{C}_2(s, t)\,v(1 - t),$$

and

$$B_1(u, v, s, t) = \dot{C}_1(u, v)\{C(u, t) - u\,C(s, t)\},$$
$$B_2(u, v, s, t) = \dot{C}_2(u, v)\{C(s, v) - v\,C(s, t)\},$$
$$B_3(u, v, s, t) = \dot{C}_1(s, t)C(u, v)(1 - s),$$
$$B_4(u, v, s, t) = \dot{C}_2(s, t)C(u, v)(1 - t).$$

As it happens,

$$A_1 \leqslant B_3, \qquad A_2 \leqslant B_1, \qquad A_3 \leqslant B_2, \qquad A_4 \leqslant B_4, \tag{A.1}$$

where the dependence on u, v, s, t has been suppressed for clarity. Indeed, $A_1 \leqslant B_3$ occurs if and only if $\dot{C}_1(u, v)\dot{C}_1(s, t)\,u(1 - s) \leqslant \dot{C}_1(s, t)C(u, v)(1 - s)$, which is equivalent to (1). Similarly, $A_4 \leqslant B_4$.

To get $A_2 \leqslant B_1$, one must check that $\dot{C}_2(s, t)\{C(u, t) - ut\} \leqslant C(u, t) - uC(s, t)$. Using (1) and the fact that $C(u, t) \geqslant ut$, one finds

$$\dot{C}_2(s, t)\{C(u, t) - ut\} \leqslant \frac{C(s, t)}{t}\{C(u, t) - ut\}$$
$$= \frac{C(s, t)}{t}\,C(u, t) - uC(s, t) \leqslant C(u, t) - uC(s, t),$$

where the last inequality is justified because $C(s, t) \leqslant t$. A similar argument yields $A_3 \leqslant B_2$, and hence (A.1) is established.

Now assume that $v > t$. Then

$$A_1(u, v, s, t) = \dot{C}_1(u, v)\dot{C}_1(s, t)\,u(1 - s),$$
$$A_2(u, v, s, t) = \dot{C}_1(u, v)\dot{C}_2(s, t)\{C(u, t) - ut\},$$
$$A_3(u, v, s, t) = \dot{C}_2(u, v)\dot{C}_1(s, t)\{C(s, v) - sv\},$$
$$A_4(u, v, s, t) = \dot{C}_2(u, v)\dot{C}_2(s, t)\,t(1 - v),$$

and

$$B_1(u, v, s, t) = \dot{C}_1(u, v)\{C(u, t) - u\,C(s, t)\},$$
$$B_2(u, v, s, t) = \dot{C}_2(u, v)C(s, t)(1 - v),$$
$$B_3(u, v, s, t) = \dot{C}_1(s, t)C(u, v)(1 - s),$$
$$B_4(u, v, s, t) = \dot{C}_2(s, t)\{C(u, t) - tC(u, v)\}.$$

In that case, it turns out that

$$A_1 \leqslant B_1, \qquad A_2 \leqslant B_4, \qquad A_3 \leqslant B_3, \qquad A_4 \leqslant B_2. \tag{A.2}$$

Indeed, $A_1 \leqslant B_1$ is equivalent to $\dot{C}_1(s, t)\,u(1 - s) \leqslant C(u, t) - uC(s, t)$. But by (1),

$$\dot{C}_1(s, t)\,u(1 - s) \leqslant \frac{C(s, t)}{s}\,u - uC(s, t).$$

Thus one can see that

$$\frac{C(s, t)}{s}\,u - uC(s, t) \leqslant C(u, t) - uC(s, t)$$

whenever $C(s, t)/s \leqslant C(u, t)/u$. The latter holds true because $u < s$ and the mapping $u \mapsto C(u, t)/u$ is non-increasing by the LTD property. Similarly, $A_4 \leqslant B_2$.

Finally, $A_2 \leqslant B_4$ is equivalent to $\dot{C}_1(u, v)\{C(u, t) - ut\} \leqslant C(u, t) - tC(u, v)$. Given that $C(u, t) \geqslant ut$, one can invoke (1) to write

$$\dot{C}_1(u, v)\{C(u, t) - ut\} \leqslant \frac{C(u, v)}{u} C(u, t) - tC(u, v) \leqslant C(u, t) - tC(u, v),$$

where the last inequality is valid because $C(u, v) \leqslant u$. The proof that $A_3 \leqslant B_3$ is similar. The conjunction of (A.1) and (A.2) implies the desired conclusion. $\quad\square$

Proof of Proposition 2. Upon setting $s = u$ and $t = v$ in the formulas presented above, one finds

$$A_2 = A_3 = \dot{C}_1(u, v)\dot{C}_2(u, v)\{C(u, v) - uv\} \leqslant 0$$

for all $u, v \in [0, 1]$. Thus it suffices to see that $A_1 \leqslant B_1 + B_3$ and $A_4 \leqslant B_2 + B_4$. This is clearly the case, because $A_1 = (1 - u)u\dot{C}_1^2(u, v)$ and $B_1 = B_3 = (1 - u)C(u, v)\dot{C}_1(u, v)$, while $A_4 = (1 - v)v\dot{C}_2^2(u, v)$ and $B_2 = B_4 = (1 - v)C(u, v)\dot{C}_2(u, v)$. $\quad\square$

Proof of Proposition 3. As in [19, Chapter 20], Hadamard differentiability is taken to mean that there exists a continuous linear functional $\dot{\mathscr{T}}_C : \mathscr{C} \to \mathbb{R}$ such that for every $\xi \in \mathscr{C}$,

$$\lim_{n \to \infty} \frac{\mathscr{T}(C + h_n \xi_n) - \mathscr{T}(C)}{h_n} = \dot{\mathscr{T}}_C(\xi)$$

whenever $h_n \downarrow 0$ and $\xi_n \to \xi$ as $n \to \infty$ uniformly. An application of the Functional Delta Method thus implies that, as $n \to \infty$,

$$n^{1/2}\{\mathscr{T}(C_n) - \mathscr{T}(C)\} = n^{1/2}\{\mathscr{T}(C + n^{-1/2}\mathbb{C}_n) - \mathscr{T}(C)\} \rightsquigarrow \dot{\mathscr{T}}_C(\mathbb{C}).$$

Consequently, $\Theta = \dot{\mathscr{T}}_C(\mathbb{C})$ in distribution. Similarly, $\hat{\Theta} = \dot{\mathscr{T}}_C(\hat{\mathbb{C}})$ in distribution.

Now because the functional $\dot{\mathscr{T}}_C$ belongs to the dual of $\mathscr{C}$, the Riesz Representation Theorem implies the existence of a bounded Borel measure μ_C on $[0, 1]^2$ such that $\dot{\mathscr{T}}_C(\xi) = \int \xi \, d\mu_C$ for all $\xi \in \mathscr{C}$. Accordingly,

$$\mathrm{var}(\Theta) = \int_{[0,1]^2} \int_{[0,1]^2} \mathrm{cov}\{\mathbb{C}(u, v), \mathbb{C}(s, t)\} \, d\mu_C(u, v) \, d\mu_C(s, t),$$

$$\mathrm{var}(\hat{\Theta}) = \int_{[0,1]^2} \int_{[0,1]^2} \mathrm{cov}\{\hat{\mathbb{C}}(u, v), \hat{\mathbb{C}}(s, t)\} \, d\mu_C(u, v) \, d\mu_C(s, t).$$

Finally, the assumption that $\mathscr{T}$ is non-decreasing means that $\xi \leqslant \xi^* \Rightarrow \mathscr{T}(\xi) \leqslant \mathscr{T}(\xi^*)$, where the inequality between functions is understood to hold pointwise. It then follows that $\dot{\mathscr{T}}_C(\xi) \geqslant 0$ whenever $\xi \geqslant 0$. Thus the measure μ_C must be positive and $\mathrm{var}(\hat{\Theta}) \leqslant \mathrm{var}(\Theta)$ occurs as soon as Inequality (2) is verified for all $u, v, s, t \in [0, 1]$. $\quad\square$

Proof of Proposition 4. A simple calculation shows that

$$\left.\begin{array}{l} C_k(\boldsymbol{v})\mathrm{cov}\{\mathbb{C}_d(\boldsymbol{u}), \mathbb{C}_d(\boldsymbol{v}_k)\} \\ C_k(\boldsymbol{u})\mathrm{cov}\{\mathbb{C}_d(\boldsymbol{u}_k), \mathbb{C}_d(\boldsymbol{v})\} \\ C_k(\boldsymbol{u})C_\ell(\boldsymbol{v})\mathrm{cov}\{\mathbb{C}_d(\boldsymbol{u}_k), \mathbb{C}_d(\boldsymbol{v}_k)\} \end{array}\right\} = \pi(\boldsymbol{u})\pi(\boldsymbol{v}) \frac{u_k \wedge v_k - u_k v_k}{u_k v_k}.$$

Given that $\mathrm{cov}\{\mathbb{C}_d(\boldsymbol{u}_k), \mathbb{C}_d(\boldsymbol{v}_\ell)\} = 0$ whenever $k \neq \ell$, one gets

$$\mathrm{cov}\{\hat{\mathbb{C}}_d(\boldsymbol{u}), \hat{\mathbb{C}}_d(\boldsymbol{v})\} - \mathrm{cov}\{\mathbb{C}_d(\boldsymbol{u}), \mathbb{C}_d(\boldsymbol{v})\} = -\pi(\boldsymbol{u})\pi(\boldsymbol{v}) \sum_{k=1}^{d} \frac{u_k \wedge v_k - u_k v_k}{u_k v_k},$$

which is clearly negative for all $\boldsymbol{u}$ and $\boldsymbol{v} \in [0, 1]^d$. $\quad\square$

References

[1] H. Joe, Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.
[2] R.B. Nelsen, An Introduction to Copulas, 2nd ed., Springer, Berlin, 2006.
[3] G.R. Shorack, J.A. Wellner, Empirical Processes with Applications to Statistics, Wiley, New York, 1986.
[4] P. Deheuvels, La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendance, Acad. Roy. Belg. Bull. Cl. Sci. 65 (5) (1979) 274–292.
[5] D. Berg, Copula goodness-of-fit testing: an overview and power comparison, Europ. J. Finance 15 (2009) 675–701.
[6] C. Genest, B. Rémillard, D. Beaudoin, Goodness-of-fit tests for copulas: a review and a power study, Insurance: Math. Econom. 44 (2009) 199–213.
[7] L. Rüschendorf, Asymptotic distributions of multivariate rank order statistics, Ann. Statist. 4 (1976) 912–923.
[8] J.-D. Fermanian, D. Radulovic, M. Wegkamp, Weak convergence of empirical copula processes, Bernoulli 10 (2004) 847–860.
[9] J.D. Esary, F. Proschan, Relationships among some concepts of bivariate dependence, Ann. Math. Statist. 43 (1972) 651–655.
[10] E.L. Lehmann, Some concepts of dependence, Ann. Math. Statist. 37 (1966) 1137–1153.

[11] A. Charpentier, J.-D. Fermanian, O. Scaillet, The estimation of copulas: theory and practice, in: J. Rank (Ed.), Copulas: From Theory to Application in Finance, Risk Publications, London, 2007, pp. 35–60.
[12] C. Genest, J. Segers, Rank-based inference for bivariate extreme-value copulas, Ann. Statist. 37 (2009) 2990–3022.
[13] A.I. Garralda-Guillem, Structure de dépendance des lois de valeurs extrêmes bivariées, C. R. Acad. Sci. Paris Sér. I Math. 330 (2000) 593–596.
[14] F. Schmid, R. Schmidt, Nonparametric inference on multivariate versions of Blomqvist's beta and related measures of tail dependence, Metrika 66 (2007) 323–354.
[15] M. Scarsini, On measures of concordance, Stochastica 8 (1984) 201–218.
[16] C. Genest, J. Nešlehová, A primer on copulas for count data, Astin Bull. 38 (2007) 475–515.
[17] M. Henmi, A paradoxical effect of nuisance parameters on efficiency of estimators, J. Japan Statist. Soc. 34 (2005) 75–86.
[18] M. Denuit, O. Scaillet, Nonparametric tests for positive quadrant dependence, J. Financ. Econ. 2 (2004) 422–450.
[19] A.W. van der Vaart, Asymptotic Statistics, Cambridge University Press, Cambridge, 1998.