Local sensitivity analyses of goodness-of-fit tests for copulas

DANIEL BERG

University of Oslo and The Norwegian Computing Center

JEAN-FRANÇOIS QUESSY

Département de mathématiques et d'informatique,
Université du Québec à Trois-Rivières

ABSTRACT. The asymptotic behavior of several goodness-of-fit statistics for copula families is obtained under contiguous alternatives. Many comparisons between a Cramér–von Mises functional of the empirical copula process and new moment-based goodness-of-fit statistics are made by considering their associated asymptotic local power curves. It is shown that the choice of the estimator for the unknown parameter can have a significant influence on the power of the Cramér–von Mises test, and that some of the moment-based statistics can provide simple and efficient goodness-of-fit methods. The paper ends with an extensive simulation study that aims to extend the conclusions to small and moderate sample sizes.

Key words: contiguous alternatives, copula, Cramér–von Mises statistic, empirical copula process, goodness-of-fit test, local power curves, rank-based estimators.

1. Introduction

Copula functions contain all the information about the dependence structure of a random vector. Indeed, due to the representation theorem of Sklar (1959), every bivariate distribution function H can be written as $H(x, y) = C\{F(x), G(y)\}$, where F and G are the marginal distributions and $C : [0, 1]^2 \to [0, 1]$ is the copula. It turns out that C, which is unique when F and G are continuous, is a distribution function with uniform marginals on $[0, 1]$. This representation enables practitioners to model the marginal behaviors and the dependence structure in separate steps. While the adjustment of univariate distributions is well documented, the study of goodness-of-fit tests for copulas emerged only recently as a challenging inferential problem.

Let C be the underlying copula of a bivariate population with continuous marginals and suppose one wants to test the goodness-of-fit hypotheses $\mathcal{H}_0 : C \in \mathcal{F} = \{C_\theta; \theta \in \mathcal{M}\}$

and $\mathcal{H}_1 : C \notin \mathcal{F} = \{C_\theta; \theta \in \mathcal{M}\}$, where $\mathcal{M}$ is the parameter space. Test statistics that help discriminate between $\mathcal{H}_0$ and $\mathcal{H}_1$ have been proposed by Fermanian (2005), Genest *et al.* (2006a), Scaillet (2006) and Chen & Fan (2005), among others. A bayesian selection procedure has also been investigated by Huard *et al.* (2006). In most cases, the efficiency of these methods, i.e. the power, is approximated by simulating repeatedly from a fixed alternative copula $D \notin \mathcal{F}$. This is done, in particular, in the works of Genest *et al.* (2008) and Berg (2007), where many simulation results and recommendations are provided.

One of the most desirable property of a statistical procedure is its ability to detect small departures from the null hypothesis. In the context of testing the fit to a particular copula family, such perturbations from $\mathcal{H}_0$ are given by the sequence of distributions

$$\mathcal{Q}_{\delta_n}(x,y) = (1-\delta_n)\, C(x,y) + \delta_n D(x,y), \tag{1}$$

where $\delta_n = n^{-1/2}\delta$, $\delta > 0$ and C, D are bivariate copulas such that $C \in \mathcal{F}$. This mixture distribution is a copula for all $0 < \delta \leq n^{1/2}$. It is supposed throughout the paper that $\mathcal{Q}_{\delta_n}$ belongs to $\mathcal{F}$ only at the limit when $n \to \infty$. Moreover, in order to ensure that the departure from $\mathcal{H}_0$ increases as δ becomes larger (at least for large values of n), it is assumed that the copula D stochastically dominates C, i.e. $D(x,y) \geq C(x,y)$ for all $(x,y) \in [0,1]^2$. The skill of a goodness-of-fit test to reject $\mathcal{H}_0$ under (1) can easily be motivated from applications in finance, where it is often advisable to detect changes in the dependence pattern over time, e.g. regime shifts for commodity markets.

In this paper, the asymptotic non-degenerate distribution of some goodness-of-fit statistics is investigated under the sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$ of alternatives. The focus is put on a Cramér–von Mises type statistic computed from a version of the empirical copula process and on simple but efficient moment-based test statistics. The characterization of their limiting behavior enables to compute asymptotic local power curves from which comparisons between the goodness-of-fit statistics under investigation can be made.

In Section 2, the goodness-of-fit test statistics studied in this work are defined. In Section 3, their asymptotic distribution under alternatives of the form (1) are obtained. These results enable to compute, in Section 4, the local power curves of the statistics under study and hence to compare the latter under chosen scenarios of local distributions. In Section 5, a new measure of asymptotic relative efficiency generalizing that of Pitman is described and computed for many cases. This index is particularly useful for the Cramér–von Mises goodness-of-fit statistic whose local power curve has no explicit expression. An extensive simulation study that aim to investigate the local behavior of the testing procedures in small and moderate sample sizes and compare with the asymptotic results follows in Section 6. The paper ends with a discussion about ideas of future investigations.

2. Some goodness-of-fit statistics for copula families

Let $(X_1,Y_1),\ldots,(X_n,Y_n)$ be a random sample from a bivariate population with continuous marginal distributions F, G and whose underlying copula is C. In Subsections 2.1, 2.2 and 2.3, statistical procedures to determine if C belongs or not to a parametric family $\mathcal{F} = \{C_\theta; \theta \in \mathcal{M}\}$ are described. It is assumed throughout that $\mathcal{M}$ is a subset of the real line, so that θ can be estimated by an empirical version of a moment of C_θ. Since all statistics considered in this work are invariant under strictly increasing transformations of the variables, one can consider, for simplicity and without any loss of generality, that the marginal distributions are uniform on the interval $[0,1]$.

2.1. The empirical copula goodness-of-fit process

A consistent estimation of a copula is possible via the empirical copula, which Deheuvels (1979) described as the distribution function of the sample of normalized ranks, i.e. $(\tilde{R}_{1,n}, \tilde{S}_{1,n})$, …, $(\tilde{R}_{n,n}, \tilde{S}_{n,n})$, where $\tilde{R}_{i,n} = F_n(X_i)$ and $\tilde{S}_{i,n} = G_n(Y_i)$, with

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\,(X_i \leq x) \quad \text{and} \quad G_n(y) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\,(Y_i \leq y)$$

being the empirical marginal distributions. Explicitly, C is estimated by

$$C_n(x,y) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left(\tilde{R}_{i,n} \leq x, \tilde{S}_{i,n} \leq y\right). \tag{2}$$

The weak consistency of the empirical process $\mathcal{C}_{n,\theta} = \sqrt{n}(C_n - C_\theta)$ to a centered gaussian limit was obtained by Deheuvels (1979) under the hypothesis of independence, i.e. in the special case when $C_\theta(x,y) = xy$. This result was extended under general distributions by Gänssler & Stute (1987), Fermanian *et al.* (2004) and Tsukahara (2005). A suggestion made by Fermanian (2005) and exploited by Quessy (2005) and Genest *et al.* (2008) consists in basing a goodness-of-fit test on a modified version of $\mathcal{C}_{n,\theta}$, namely $\mathcal{C}_n = \sqrt{n}(C_n - C_{\hat{\theta}_n})$, where $\hat{\theta}_n$ consistently estimates θ. As shown by Quessy (2005), $\mathcal{C}_n$ is weakly consistent under $\mathcal{H}_0$ if the following assumptions are satisfied.

$\mathcal{A}_1$. For all $\theta \in \mathcal{M}$, the first order partial derivatives of C_θ exist and are continuous;

$\mathcal{A}_2$. $(\mathcal{C}_{n,\theta}, \Theta_n)$ converges jointly to a gaussian process $(\mathcal{C}_\theta, \Theta)$, where $\Theta_n = \sqrt{n}(\hat{\theta}_n - \theta)$. Moreover, for all $\theta \in \mathcal{M}$ and as $\varepsilon \downarrow 0$,

$$\sup_{\|\theta^\star - \theta\| < \varepsilon} \sup_{(x,y) \in [0,1]^2} \left| \dot{C}_{\theta^\star}(x,y) - \dot{C}_\theta(x,y) \right| \longrightarrow 0,$$

where $\dot{C}_\theta = \partial C_\theta / \partial \theta$.

Under $\mathcal{A}_1$ and $\mathcal{A}_2$, the empirical goodness-of-fit process $\mathcal{C}_n$ converges weakly to a centered limit $\mathcal{C} = \mathcal{C}_\theta - \Theta \dot{C}_\theta$ having covariance function $\Gamma_{\mathcal{C}}(u,v,u',v') = \operatorname{cov}\{\mathcal{C}(u,v), \mathcal{C}(u',v')\}$ whose expression is explicit but cumbersome. Thanks to this asymptotic result, it is then justified to base a goodness-of-fit test on some continuous functional computed from $\mathcal{C}_n$ in virtue of the continuous mapping theorem (see Billingsley, 1968). An omnibus statistic which has good power properties in general is the Cramer–von Mises distance function

$$\mathcal{V}_n = \int_0^1 \int_0^1 \{\mathcal{C}_n(x,y)\}^2 \,\mathrm{d}x\mathrm{d}y. \tag{3}$$

Note that the use of other functional distances are possible, e.g. the Kolmogorov–Smirnov type statistics, but the latter have been found by Genest *et al.* (2006a) and by Genest *et al.* (2008) to be generally less powerful than the Cramér–von Mises statistic. Since statistic (3) has no explicit form in general, Genest & Rémillard (2008) proposed to rely on the parametric bootstrap version

$$\mathcal{V}_{n,N} = \int_0^1 \int_0^1 \{\mathcal{C}_{n,N}(x,y)\}^2 \, \mathrm{d}x\mathrm{d}y,$$

where $\mathcal{C}_{n,N} = \sqrt{n}(C_n - C_N)$ and C_N is the empirical copula computed via equation (2) from an artificial sample $(X^\star_{1,n}, Y^\star_{1,n}), \ldots, (X^\star_{N,n}, Y^\star_{N,n})$ from $C_{\hat\theta_n}$. These authors show that as $n, N \to \infty$, the process $\mathcal{C}_{n,N}$ converges to the same limit as $\mathcal{C}_n$ and consequently, $\mathcal{V}_{n,N}$ has the same asymptotic distribution as $\mathcal{V}_n$.

2.2. Moment-based goodness-of-fit statistics

Consider two real valued moments m_1 and m_2 of C_θ that are related to θ by one-to-one functions r_1, r_2 defined on $\mathcal{M}$ such that $m_1 = r_1(\theta)$ and $m_2 = r_2(\theta)$. Under the null hypothesis that the unknown copula of a population belongs to $\mathcal{F}$, one has $r_1^{-1}(m_1) = r_2^{-1}(m_2)$. If $\hat m_{1,n}$ and $\hat m_{2,n}$ are consistent for m_1 and m_2 respectively, then $\hat\theta_{1,n} = r_1^{-1}(\hat m_{1,n})$ and $\hat\theta_{2,n} = r_2^{-1}(\hat m_{2,n})$ provide consistent estimations of θ. In most cases of interest, $\sqrt{n}(\hat\theta_{j,n} - \theta)$ is asymptotically normal with mean zero and variance $\sigma_j^2(C_\theta)$ under $\mathcal{H}_0$. A simple, asymptotically normal goodness-of-fit statistic is then given by

$$\mathcal{S}_n = \sqrt{n}\left\{r_1^{-1}(\hat m_{1,n}) - r_2^{-1}(\hat m_{2,n})\right\}. \tag{4}$$

A goodness-of-fit test then consists in rejecting the null hypothesis whenever $|\mathcal{S}_n|/\sigma(C_\theta)$ exceeds $z_{\alpha/2}$, i.e. the $(1-\alpha/2)$-th percentile of a $\mathcal{N}(0,1)$ distribution, where $\sigma^2(C_\theta) = \lim_{n\to\infty} \mathrm{var}(\mathcal{S}_n)$. Note that tests based on $\mathcal{S}_n$ may be inconsistent since it may happen that $r_1^{-1}(m_1) = r_2^{-1}(m_2)$ even if $\mathcal{H}_0$ is false.

The above method can be employed by considering two of the most popular measures of association, namely Spearman's rho and Kendall's tau. The latter are respectively defined, in terms of the underlying copula C_θ of the population, by

$$\rho_{C_\theta}(\theta) = 12\int_0^1\int_0^1 C_\theta(x,y)\mathrm{d}x\mathrm{d}y - 3 \quad \text{and} \quad \tau_{C_\theta}(\theta) = 4\int_0^1\int_0^1 C_\theta(x,y)\mathrm{d}C_\theta(x,y) - 1. \tag{5}$$

Consistent estimators based on inversions of these rank statistics are $\hat\theta_{n,\rho} = \rho_{C_\theta}^{-1}(\rho_n)$ and $\hat\theta_{n,\tau} = \tau_{C_\theta}^{-1}(\tau_n)$, where

$$\rho_n = 1 - \frac{6n}{n^2-1}\sum_{i=1}^n \left(\tilde R_{i,n} - \tilde S_{i,n}\right)^2 \quad \text{and} \quad \tau_n = -1 + \frac{4}{n(n-1)}\sum_{i\neq j} \mathbf{1}\left(X_i \le X_j, Y_i \le Y_j\right)$$

are their sample versions. Another estimator arises from the so-called pseudo maximum-likelihood method, which is similar to the classical likelihood approach but where the normalized ranks are used instead of the observations. The resulting estimator $\hat\theta_{n,PL}$ has been studied by Genest *et al.* (1995), Shih & Louis (1995) and recently by Kim *et al.*

(2006). Based on these three consistent estimators, one can build three goodness-of-fit statistics of the form (4), namely

$$\mathcal{S}_{n1} = \sqrt{n}\left(\hat{\theta}_{n,\rho} - \hat{\theta}_{n,\tau}\right), \; \mathcal{S}_{n2} = \sqrt{n}\left(\hat{\theta}_{n,\rho} - \hat{\theta}_{n,PL}\right) \text{ and } \mathcal{S}_{n3} = \sqrt{n}\left(\hat{\theta}_{n,\tau} - \hat{\theta}_{n,PL}\right). \quad (6)$$

2.3. Shih's goodness-of-fit test for the gamma frailty model

The dependence function associated to the bivariate gamma frailty model, also referred to as Clayton's copula, is given in Equation (13) to be found in Appendix B. Shih (1998) considered unweighed and weighted estimations of the dependence parameter θ via Kendall's tau τ_n and the weighted rank-based statistic

$$\hat{\theta}_{n,W} = \sum_{i<j} \frac{\Delta_{ij}}{W_{ij}} \Big/ \sum_{i<j} \frac{1-\Delta_{ij}}{W_{ij}},$$

where $\Delta_{ij} = \mathbf{1}\{(X_i - X_j)(Y_i - Y_j) > 0\}$ and

$$W_{ij} = \sum_{k=1}^{n} \mathbf{1}\{X_k \le \max(X_i, X_j), \; Y_k \le \max(Y_i, Y_j)\}.$$

Since $\hat{\theta}_{n,\tau} = 2\tau_n/(1-\tau_n)$ and $\hat{\theta}_{n,W}$ are both unbiased for θ under the null hypothesis that C belongs to Clayton's family of copulas, a version of a goodness-of-fit statistic proposed by Shih (1998) is $\mathcal{S}_{n4} = \sqrt{n}(\hat{\theta}_{n,\tau} - \hat{\theta}_{n,W})$. One deduces from arguments to be found in Shih (1998) that $\mathcal{S}_{n4}$ is asymptotically normal under the null hypothesis. Unfortunately, the variance provided by Shih (1998) was found to be wrong by Genest *et al.* (2006b), where a corrected formula is provided. From the work of these authors, one may deduce the asymptotic representation

$$\mathcal{S}_{n4} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{K_\theta(X_i, Y_i) - L_\theta(X_i, Y_i)\} + o_{\mathrm{P}}(1), \quad (7)$$

where

$$K_\theta(x,y) = 2(\theta+2)^2 \left\{ 2\left(x^{-\theta} + y^{-\theta} - 1\right)^{-1/\theta} - x - y + \frac{1}{\theta+2} \right\}$$

and

$$L_\theta(x,y) = (\theta+1)(2\theta+1)\log\left(x^{-\theta} + y^{-\theta} - 1\right)^{-1/\theta} - (\theta+1)^2 \log(xy) + \theta.$$

Genest *et al.* (2006b) then used (7) to compute the asymptotic variance of $\mathcal{S}_{n4}$, whose complicated expression is given by

$$\begin{aligned}\sigma_4^2(C_\theta) &= \frac{136\theta^7 + 1352\theta^6 + 5171\theta^5 + 9449\theta^4 + 8281\theta^3 + 3001\theta^2 + 240\theta + 18}{3\theta^2(\theta+1)^2(\theta+3)^2} \\ &\quad + \frac{8(\theta+2)^4}{\theta^2(\theta+1)^2}\,\mathcal{I}_1(\theta) - \frac{4(\theta+1)^4}{\theta^4} \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1+1/\theta)^2} - \frac{8(\theta+1)(\theta+2)}{\theta^3}\,\mathcal{I}_2(\theta),\end{aligned}$$

where

$$\mathcal{I}_1(\theta) = \sum_{k=0}^{\infty} \frac{\Gamma^2(1/\theta)}{\Gamma(1/\theta)} \frac{k!\Gamma(k+1/\theta)}{\Gamma(k+1+2/\theta)} \quad \text{and} \quad \mathcal{I}_2(\theta) = \sum_{k=0}^{\infty} \frac{\Gamma(2/\theta)k!}{(k+1/\theta)\Gamma(k+1+2/\theta)} .$$

3. Asymptotic behavior under local sequences

In order to derive non-degenerate limiting distributions for a given goodness-of-fit statistic under the sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$ defined in Equation (1), one has to ensure that $\mathcal{Q}_{\delta_n}$ is *close* to $\mathcal{Q}_0 = C_\theta$ in a certain sense. One such criteria is given by van der Vaart & Wellner (1996), where it is supposed that

$$\lim_{n\to\infty} \int_0^1 \int_0^1 \left\{ \sqrt{n} \left(\sqrt{q_{\delta_n}(x,y)} - \sqrt{q_0(x,y)} \right) - \frac{\delta \dot{q}_0(x,y)}{2\sqrt{q_0(x,y)}} \right\}^2 \mathrm{d}x\mathrm{d}y = 0, \tag{8}$$

for q_δ being the density associated to $\mathcal{Q}_\delta$ and $\dot{q}_\delta = \partial q_\delta/\partial\delta$. Note that condition (8) entails that the sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$ is contiguous with respect to $\mathcal{Q}_0$. This is the key requirement that enables to derive the asymptotic local representation of the goodness-of-fit statistics $\mathcal{V}_{n,N}$ and $\mathcal{S}_{n1}, \ldots, \mathcal{S}_{n4}$. This is the subject of the remaining of this section.

3.1. Local behavior of some estimators of the dependence parameter

Many interesting estimators for the unknown parameter of a copula family admit the asymptotic representation

$$\Theta_{n,\Lambda} = \sqrt{n}\left(\hat{\theta}_{n,\Lambda} - \theta\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Lambda_{C_\theta}\left(\tilde{R}_{i,n}, \tilde{S}_{i,n}\right) + o_{\mathrm{P}}(1), \tag{9}$$

where $\Lambda_{C_\theta} : [0,1]^2 \to [0,1]$ is a twice differentiable score function such that for all $\theta \in \mathcal{M}$ and all $(x,y) \in [0,1]^2$, $\mathrm{E}_{C_\theta}\{\Lambda_{C_\theta}(X,Y)\} = 0$ and $|\Lambda''_{C_\theta}(x,y)| \leq g_\theta(x,y)$, where g_θ and $\Lambda^2_{C_\theta}$ are integrable with respect to $c_\theta(x,y) = \partial^2 C_\theta(x,y)/\partial x \partial y$. These conditions ensure that $\Theta_{n,\Lambda}$ converges in law to

$$\Theta_\Lambda = \Theta'_\Lambda + \int_{(0,1)^2} \Lambda_{C_\theta,10}(x,y)\beta_1(x)c_\theta(x,y)\mathrm{d}x\mathrm{d}y + \int_{(0,1)^2} \Lambda_{C_\theta,01}(x,y)\beta_2(y)c_\theta(x,y)\mathrm{d}x\mathrm{d}y,$$

where Θ'_Λ is the limit of $n^{-1/2}\sum_{i=1}^n \Lambda_{C_\theta}(X_i,Y_i)$ and β_1, β_2 are uniform brownian bridges, i.e. gaussian processes with covariance function $\mathrm{cov}\{\beta_j(s),\beta_j(t)\} = \min(s,t) - st$, $j = 1,2$, arising as the limits of $\sqrt{n}\{F_n(x) - x\}$ and $\sqrt{n}\{G_n(y) - y\}$ respectively. Here, $\Lambda_{C_\theta,10}(x,y) = \partial\Lambda_{C_\theta}(x,y)/\partial x$ and $\Lambda_{C_\theta,01}(x,y) = \partial\Lambda_{C_\theta}(x,y)/\partial y$.

Among the estimators that admit representation (9), one has the inversion of Spearman's rho and the pseudo-maximum likelihood estimator explored by Genest *al.* (1995) and Shih & Louis (1995). More details will be given in Example 1 and Example 2. Another popular estimation strategy using a statistic that is not of the form (9) is based on $\hat{\theta}_{n,\tau}$, i.e. on the inversion of Kendall's measure of association.

The next proposition, whose proof is deferred to Appendix A.1, identifies the asymptotic distribution of $\Theta_{n,\Lambda}$ and $\Theta_{n,\tau} = \sqrt{n}(\hat{\theta}_{n,\tau} - \theta)$ under contiguous alternatives of the type (1). This result is a prerequisite in order to compute the local power of moment-based goodness-of-fit statistics described in Section 2.2. It will also enable to characterize the asymptotic behavior of the process $\mathcal{C}_n$, and consequently that of $\mathcal{V}_{n,N}$, under $(\mathcal{Q}_{\delta_n})_{n\geq 1}$ for several strategies that aim to estimate θ.

Proposition 1
Assume that condition (8) holds for the sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$. Then under $(\mathcal{Q}_{\delta_n})_{n\geq 1}$,

(i) $\Theta_{n,\Lambda} \rightsquigarrow \Theta_\Lambda + \delta\mu_\Lambda(C_\theta, D)$, *where* $\mu_\Lambda(C_\theta, D) = \mathrm{E}_D\left\{\Lambda_{C_\theta}(X,Y)\right\} - \mathrm{E}_{C_\theta}\left\{\Lambda_{C_\theta}(X,Y)\right\}$ *and* Θ_Λ *is a normal random variable with mean 0 and variance*

$$\sigma_\Lambda^2(C_\theta) = \mathrm{var}\left\{\Lambda_{C_\theta}(X,Y) + \int_0^1\int_X^1 \Lambda_{C_\theta,10}(x,y)c_\theta(x,y) + \int_Y^1\int_0^1 \Lambda_{C_\theta,01}(x,y)c_\theta(x,y)\right\};$$

(ii) $\Theta_{n,\tau} \rightsquigarrow \Theta_\tau + \delta\mu_\tau(C_\theta, D)$, *where* $\mu_\tau(C_\theta, D) = 4\{\tau'_{C_\theta}(\theta)\}^{-1}\left\{\mathrm{E}_D(C_\theta) - \mathrm{E}_{C_\theta}(C_\theta)\right\}$ *and* Θ_τ *is a normal random variable with mean 0 and variance*

$$\sigma_\tau^2 = \frac{16}{\{\tau'_{C_\theta}(\theta)\}^2}\,\mathrm{var}\left\{2C_\theta(X,Y) - X - Y\right\}.$$

The next two examples are applications of part (i) of Proposition 1 when the estimator is based on an inversion of Spearman's rho and on the pseudo maximum-likelihood estimator.

Example 1. Let $\rho_{C_\theta}(\theta)$ be the population value of Spearman's measure of association for a vector (X,Y) with underlying copula C_θ. Then $\hat{\theta}_{n,\rho} = \rho_{C_\theta}^{-1}(\rho_n)$ is a consistent estimator for θ, where ρ_n is Spearman's rank correlation coefficient. Using a Taylor expansion of order 1, one can show that this estimator can be written in the form (9) with $\Lambda_{C_\theta}(x,y) = \{\rho'_{C_\theta}(\theta)\}^{-1}\{12xy - 3 - \rho_{C_\theta}(\theta)\}$, where $\rho'_{C_\theta}(\theta) = \partial\rho_{C_\theta}(\theta)/\partial\theta$. Thus, under the contiguous sequence (1), $\Theta_{n,\rho} = \sqrt{n}(\hat{\theta}_{n,\rho} - \theta)$ is asymptotically normal with drift parameter $\mu_\rho(C_\theta, D) = \{\rho'_{C_\theta}(\theta)\}^{-1}\{\rho_D - \rho_{C_\theta}(\theta)$ and variance

$$\sigma_\rho^2(C_\theta) = \frac{144}{\{\rho'_{C_\theta}(\theta)\}^2}\,\mathrm{var}\left\{XY + \int_0^1\int_X^1 yc_\theta(x,y)\mathrm{d}x\mathrm{d}y + \int_Y^1\int_0^1 xc_\theta(x,y)\mathrm{d}x\mathrm{d}y\right\}.$$

Example 2. Let $\hat{\theta}_{n,PL}$ be the pseudo likelihood estimator. From the work of Genest *et al.* (1995), one has representation (9) with $\Lambda_{C_\theta}(x,y) = \beta_{C_\theta}^{-1}\ell'_{C_\theta}(x,y)$, where $\ell_{C_\theta}(x,y) = \log c_\theta(x,y)$ and $\beta_{C_\theta} = \mathrm{E}_{C_\theta}[\{\ell'_{C_\theta}(X,Y)\}^2]$, with $\ell'_{C_\theta} = \partial\ell_{C_\theta}/\partial\theta$. An application of Proposition 1 shows that $\Theta_{n,PL} = \sqrt{n}(\hat{\theta}_{n,PL} - \theta)$ converges in law to a normal distribution with variance $\sigma_{PL}^2(C_\theta) = \beta_{C_\theta}^{-2}\mathrm{var}\{\ell'_{C_\theta}(X,Y) - W_{C_\theta,1}(X) - W_{C_\theta,2}(Y)\}$, where

$$W_{C_\theta,1}(u) = \int_u^1\int_0^1 \ell'_{C_\theta}(x,y)\ell'_{C_\theta,1}(x,y)c_\theta(x,y)\mathrm{d}x\mathrm{d}y$$

and

$$W_{C_\theta,2}(u) = \int_0^1 \int_u^1 \ell'_{C_\theta}(x,y)\ell'_{C_\theta,2}(x,y)c_\theta(x,y)\mathrm{d}x\mathrm{d}y,$$

with $\ell'_{C_\theta,1}(x,y) = \partial \ell_{C_\theta}(x,y)/\partial x$ and $\ell'_{C_\theta,2}(x,y) = \partial \ell_{C_\theta}(x,y)/\partial y$. The asymptotic mean is

$$\mu_{PL}(C_\theta, D) = \beta_{C_\theta}^{-1}\mathrm{E}_D\left\{\ell'_{C_\theta}(X,Y)\right\} - \beta_{C_\theta}^{-1}\mathrm{E}_{C_\theta}\left\{\ell'_{C_\theta}(X,Y)\right\} = \beta_{C_\theta}^{-1}\mathrm{E}_D\left\{\ell'_{C_\theta}(X,Y)\right\},$$

since by Lebesgue's dominated convergence theorem,

$$E_{C_\theta}\left\{\ell'_{C_\theta}(X,Y)\right\} = \int_0^1 \int_0^1 \dot{c}_\theta(x,y)\ \mathrm{d}x\ \mathrm{d}y = \frac{\partial}{\partial\theta}\int_0^1 \int_0^1 c_\theta(x,y)\ \mathrm{d}x\ \mathrm{d}y = 0.$$

3.2. Local behavior of the goodness-of-fit statistics

The first theoretical result of this section establishes the large-sample behavior of $\mathcal{C}_n$ under the sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$. It is assumed that the estimator of θ is either of the form (9) or based on the inversion of Kendall's tau.

Proposition 2
Suppose condition (8) *and Assumptions $\mathcal{A}_1$–$\mathcal{A}_2$ hold and assume that $\Theta_n = \sqrt{n}(\hat{\theta}_n - \theta)$ converges in law to $\tilde{\Theta} = \Theta + \delta\mu(C_\theta, D)$ under the sequence* (1)*, where Θ is the limit in law of Θ_n under $\mathcal{H}_0$. Then under $(\mathcal{Q}_{\delta_n})_{n\geq 1}$, the empirical process $\mathcal{C}_n = \sqrt{n}(C_n - C_{\hat{\theta}_n})$ converges weakly to*

$$\tilde{\mathcal{C}} = \mathcal{C} + \delta\left\{D - C_\theta - \mu(C_\theta, D)\dot{C}_\theta\right\},$$

where $\mathcal{C}$ is the weak limit of $\mathcal{C}_n$ under $\mathcal{H}_0$ and $\dot{C}_\theta = \partial C_\theta/\partial\theta$.

Remark. As one may expect, a sequence of the form $\mathcal{Q}_{\delta_n} = C_{\theta+\delta_n}$ yields absolutely no power for statistics based on $\mathcal{C}_n$ since $\mathcal{Q}_{\delta_n} \in \mathcal{F}$ in that case. Indeed, as one can deduce from computations made in the proof of Proposition 2, condition (8) enounced in van der Vaart & Wellner (1996) implies that $\mathcal{C}_{n,\theta}$ converges to $\mathcal{C}_\theta + \delta\dot{C}_\theta$. Moreover, since Θ_n converges to $\Theta+\delta$ in that case, $\sqrt{n}(C_{\hat{\theta}_n} - C_\theta)$ converges to $(\Theta+\delta)\dot{C}_\theta$, so that $\mathcal{C}_n = \mathcal{C}_{n,\theta} - \sqrt{n}(C_{\hat{\theta}_n} - C_\theta)$ converges to $\mathcal{C}$, i.e. to the same limit as under $\mathcal{H}_0$.

The asymptotic local behavior of the moment-based goodness-of-fit statistics (6) can easily be obtained as consequences of Proposition 1. This is the subject of Proposition 3, whose straightforward proof is omitted.

Proposition 3
Suppose condition (8) *holds. Then under $(\mathcal{Q}_{\delta_n})_{n\geq 1}$,*
(i) $\mathcal{S}_{n1} \rightsquigarrow \mathcal{S}_1 + \delta\left\{\mu_\rho(C_\theta, D) - \mu_\tau(C_\theta, D)\right\}$;

(ii) $\mathcal{S}_{n2} \rightsquigarrow \mathcal{S}_2 + \delta\left\{\mu_\rho(C_\theta, D) - \mu_{PL}(C_\theta, D)\right\}$;

(iii) $\mathcal{S}_{n3} \rightsquigarrow \mathcal{S}_3 + \delta\left\{\mu_\tau(C_\theta, D) - \mu_{PL}(C_\theta, D)\right\}$.

This result implies that the limiting distribution of S_{nj}, $j = 1, 2, 3$ under the contiguous sequence is normal with some mean $\delta\mu_j(C_\theta, D)$ and variance $\sigma_j^2(C_\theta)$. As long as $\mu(C_\theta, D) \neq 0$, a goodness-of-fit procedure based on S_{nj} will yield power locally.

3.3. Shih's statistic under contiguity

The asymptotic behavior of $\mathcal{S}_{n4}$ under the contiguous sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$ will follow from an application of Lecam's third lemma and the asymptotic representation (7). The result is summarized in Proposition 4.

Proposition 4
Under the contiguous sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$, the goodness-of-fit statistic $\mathcal{S}_{n4}$ converges in law to a normal distribution with variance $\sigma_4^2(C_\theta)$ and mean $\delta\eta_1(C_\theta, D) - \delta\eta_2(C_\theta, D)$, where

$$\begin{aligned}
\eta_1(C_\theta, D) &= 4(\theta+2)^2 \left\{ \mathrm{E}_D(C_\theta) - \mathrm{E}_{C_\theta}(C_\theta) \right\}, \\
\eta_2(C_\theta, D) &= (\theta+1)(\theta+2) \int_0^1 \int_0^1 \{d(u,v) - c_\theta(u,v)\} \log C_\theta(u,v) \mathrm{d}u\mathrm{d}v \\
&\quad - (\theta+1)^2 \int_0^1 \int_0^1 \{d(u,v) - c_\theta(u,v)\} \log uv \mathrm{d}u\mathrm{d}v.
\end{aligned}$$

4. Local power comparisons

In this section, the asymptotic power of the goodness-of-fit tests based on $\mathcal{V}_{n,N}$ and $\mathcal{S}_{n1}, \ldots, \mathcal{S}_{n4}$ are investigated under alternatives of the form (1). Here, C and D are chosen to be in the same family with different levels of dependence. In other words, local alternatives of the form $\mathcal{Q}_{\delta_n}(x,y) = (1-\delta_n)C_\theta(x,y) + \delta_n C_{\theta'}(x,y)$ are considered, where $\theta < \theta'$. It is assumed that θ is a dependence parameter for the family $\{C_\theta; \theta \in \mathcal{M}\}$, i.e. $C_\theta(x,y) \leq C_{\theta'}(x,y)$ for all $(x,y) \in [0,1]^2$. This requirement is fulfilled for most families of copulas. The above mixture distribution can represent a setting where the data generating process stays in the same family over time but the dependence strength suddenly changes, c.f. regime-shifting models. Structural changes of this kind can occur in mean-reverting processes such as those driving oil and other commodity prices, where the dependence pattern, i.e. the copula family, remains the same over time but the strength of this link becomes significantly stronger or weaker at some moment.

The following analyses will consider local distributions involving mixtures of Clayton, Frank, Gumbel–Barnett and Normal copulas whose analytical expressions are given in equations (13)–(16) to be found in Appendix B.

4.1. Efficiency of the empirical copula process under various estimation strategies

Here, the influence of the estimation strategy on the power of the Cramér–von Mises statistics is investigated under local sequences. Here and in the sequel, $\mathcal{C}_{n,N,\rho}$, $\mathcal{C}_{n,N,\tau}$ and $\mathcal{C}_{n,N,PL}$ refer to the empirical copula goodness-of-fit process with the estimation of θ based respectively on Spearman's rho, Kendall's tau and the pseudo-likelihood approach. Similarly, $\mathcal{V}_{n,N}^{\rho}$, $\mathcal{V}_{n,N}^{\tau}$ and $\mathcal{V}_{n,N}^{PL}$ are the associated Cramér–von Mises functionals.

Table 1: Drift terms for the estimators based on Spearman's rho, the pseudo-maximum likelihood and Kendall's tau under mixtures of Clayton, Frank, Gumbel–Barnett and Normal copulas

τ_{C_θ}	τ_D	Model	μ_ρ	μ_{PL}	μ_τ	Model	μ_ρ	μ_{PL}	μ_τ
0.1	0.2		0.244	0.250	0.030		0.901	0.926	0.111
0.1	0.3		0.475	0.487	0.059		1.789	1.815	0.231
0.1	0.4		0.692	0.697	0.086		2.615	2.704	0.333
0.1	0.5	Clayton	0.889	0.882	0.114	Frank	3.385	3.519	0.435
0.4	0.5		0.527	0.544	0.067		1.319	1.381	0.164
0.4	0.6		0.996	0.995	0.128		2.436	2.619	0.315
0.4	0.7		1.384	1.393	0.183		3.351	3.810	0.452
0.4	0.8		1.679	1.786	0.228		4.021	4.762	0.548
0.1	0.2		0.099	0.101	0.013		0.154	0.154	0.019
0.1	0.3		0.192	0.198	0.025		0.301	0.302	0.037
0.1	0.4		0.281	0.290	0.037		0.440	0.443	0.054
0.1	0.5	Gumbel–	0.485	0.379	0.049	Normal	0.565	0.572	0.071
0.4	0.5	Barnett	0.096	0.101	0.016		0.120	0.123	0.017
0.4	0.6		0.179	0.195	0.029		0.226	0.228	0.032
0.4	0.7		0.250	0.282	0.044		0.312	0.315	0.046
0.4	0.8		0.303	0.797	0.068		0.377	0.387	0.062

According to Proposition 2, the weak limits of the empirical copula goodness-of-fit processes $\mathcal{C}_{n,N,\rho}$, $\mathcal{C}_{n,N,\tau}$ and $\mathcal{C}_{n,N,PL}$ under the contiguous sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$ are

$$\tilde{\mathcal{C}}_\rho = \mathcal{C}_\rho + \delta(g - \mu_\rho \dot{C}_\theta), \quad \tilde{\mathcal{C}}_\tau = \mathcal{C}_\tau + \delta(g - \mu_\tau \dot{C}_\theta) \quad \text{and} \quad \tilde{\mathcal{C}}_{PL} = \mathcal{C}_{PL} + \delta(g - \mu_{PL} \dot{C}_\theta),$$

where $\mathcal{C}_\rho$, $\mathcal{C}_\tau$ and $\mathcal{C}_{PL}$ are the respective limits under $\mathcal{H}_0$ and $g(x,y) = D(x,y) - C_\theta(x,y)$. Computations of μ_ρ, μ_τ and μ_{PL} are detailed in Appendix B for mixtures of Clayton, Frank, Gumbel–Barnett and Normal copulas. The results are reported in Table 1. Generally speaking, these drift terms are higher for $\Theta_{n,\rho}$ and $\Theta_{n,PL}$ than for $\Theta_{n,\tau}$. This indicates that the estimator based on Kendall's tau is more robust under perturbations of $\mathcal{H}_0$ of the type $\mathcal{Q}_{\delta_n}$, which is not necessarily a good property for goodness-of-fit testing where one wants to detect departures from $\mathcal{H}_0$.
There is no hope to obtain explicit representations for the asymptotic distributions of $\mathcal{V}^\rho_{n,N}$, $\mathcal{V}^\tau_{n,N}$ and $\mathcal{V}^{PL}_{n,N}$, and consequently for the associated power curves. A procedure to overcome this difficulty is explained next in order to compute the local power curve of the Cramér–von Mises tests. For simplicity, only the case involving $\mathcal{V}^\rho_{n,N}$ is detailed.

First note that under $(\mathcal{Q}_{\delta_n})_{n\geq 1}$,

$$\mathcal{V}^\rho_{n,N} \rightsquigarrow \tilde{\mathcal{V}}^\rho = \int_0^1 \int_0^1 \left\{\tilde{\mathcal{C}}_\rho(x,y)\right\}^2 \mathrm{d}x\mathrm{d}y = \int_0^1 \int_0^1 \left\{\mathcal{C}_\rho(x,y) + \delta h_\rho(x,y)\right\}^2 \mathrm{d}x\mathrm{d}y,$$

where $h_\rho(x,y) = D(x,y) - C_\theta(x,y) - \mu_\rho(C_\theta, D)\dot{C}_\theta(x,y)$. Hence, for large values of n and

N, an approximation is given by

$$\tilde{\mathcal{V}}^{\rho}_{n,N} = \int_0^1 \int_0^1 \{\mathcal{C}_{n,N,\rho}(x,y) + \delta h_\rho(x,y)\}^2 \mathrm{d}x\mathrm{d}y,$$

where $\mathcal{C}_{n,N,\rho}$ is the empirical copula goodness-of-fit process where θ is estimated through an inversion of Spearman's rho. One can see that $\tilde{\mathcal{V}}^{\rho}_{n,N} = \mathcal{V}^{\rho}_{n,N} + 2\delta V_1 + \delta^2 V_2$, where

$$\begin{aligned} V_1 &= \int_0^1 \int_0^1 h_\rho(x,y)\mathcal{C}_{n,N,\rho}(x,y)\mathrm{d}x\mathrm{d}y \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\tilde{R}_{i,n}}^1 \int_{\tilde{S}_{i,n}}^1 h_\rho(x,y)\mathrm{d}x\mathrm{d}y - \sqrt{n} \int_0^1 \int_0^1 h_\rho(x,y) C_{\hat{\theta}_{n,\rho}}(x,y)\mathrm{d}x\mathrm{d}y \end{aligned}$$

and

$$V_2 = \int_0^1 \int_0^1 \{h_\rho(x,y)\}^2 \,\mathrm{d}x\mathrm{d}y.$$

In Figure 1 and Figure 2, the local power curves of the Cramér–von Mises test statistic computed under the three considered estimation strategies using the above approximations with $n = 2500$ and $N = 2500$ are reported under mixtures of Clayton, Frank, Gumbel–Barnett and Normal copulas. The strength of the dependence of the null copula C and of the perturbation copula D, as measured by Kendall's tau, are $(\tau_C, \tau_D) = (0.1, 0.5)$ in Figure 1 and $(\tau_C, \tau_D) = (0.4, 0.8)$ in Figure 2.
It is first interesting to note that surprisingly, the choice of the estimator has a significant impact on the local power curves in almost all cases considered, except under Normal mixtures. Under Clayton alternatives, the conclusions are the same in Figure 1 and Figure 2, namely that $\mathcal{V}^{\tau}_{n,N}$ has a significantly much larger local power than its two competitors. Overall, $\mathcal{V}^{\rho}_{n,N}$ is the least powerful locally. Probably due to the fact that the drift terms μ_τ associated to the estimation by Kendall's tau are small (see Table 1), $\mathcal{V}^{\tau}_{n,N}$ performs generally very well, especially in the case of small level of dependence, i.e. for $(\tau_C, \tau_D) = (0.1, 0.5)$. For higher degrees of dependence, $\mathcal{V}^{PL}_{n,N}$ is often better than $\mathcal{V}^{\tau}_{n,N}$ and constitutes a good choice under all scenarios, except for Clayton mixtures.

4.2. Comparison of the empirical copula process with the moment-based statistics

In view of Propositions 3 and 4, the asymptotic local power curves $\beta_1, \ldots, \beta_4$ of the goodness-of-fit tests based on $\mathcal{S}_{n1}, \ldots, \mathcal{S}_{n4}$ are of the form

$$\beta_j(\delta, C_\theta, D) = 1 - \Phi\left\{ z_{\alpha/2} - \left| \frac{\delta\mu_j(C_\theta, D)}{\sigma_j(C_\theta)} \right| \right\} + \Phi\left\{ -z_{\alpha/2} - \left| \frac{\delta\mu_j(C_\theta, D)}{\sigma_j(C_\theta)} \right| \right\}, \qquad (10)$$

where $z_{\alpha/2}$ is the $(1-\alpha/2)$-th percentile of a $\mathcal{N}(0,1)$ distribution. Here, $\mu_1 = \mu_\rho - \mu_\tau$, $\mu_2 = \mu_\rho - \mu_{PL}$, $\mu_3 = \mu_\tau - \mu_{PL}$ and $\mu_4 = \eta_1 - \eta_2$. In view of equation (10), the local power of the test based on $\mathcal{S}_{nj}$ only depends on the absolute value of the ratio $\mu_j(C_\theta, D)/\sigma_j(C_\theta)$, i.e. the asymptotic local efficiency. Some values of μ_1, μ_2 and μ_3 are reported in Table 2 under the four choices of mixture distributions. The highest local efficiencies, i.e. the one that yields the most power locally among the three, are identified in bold.

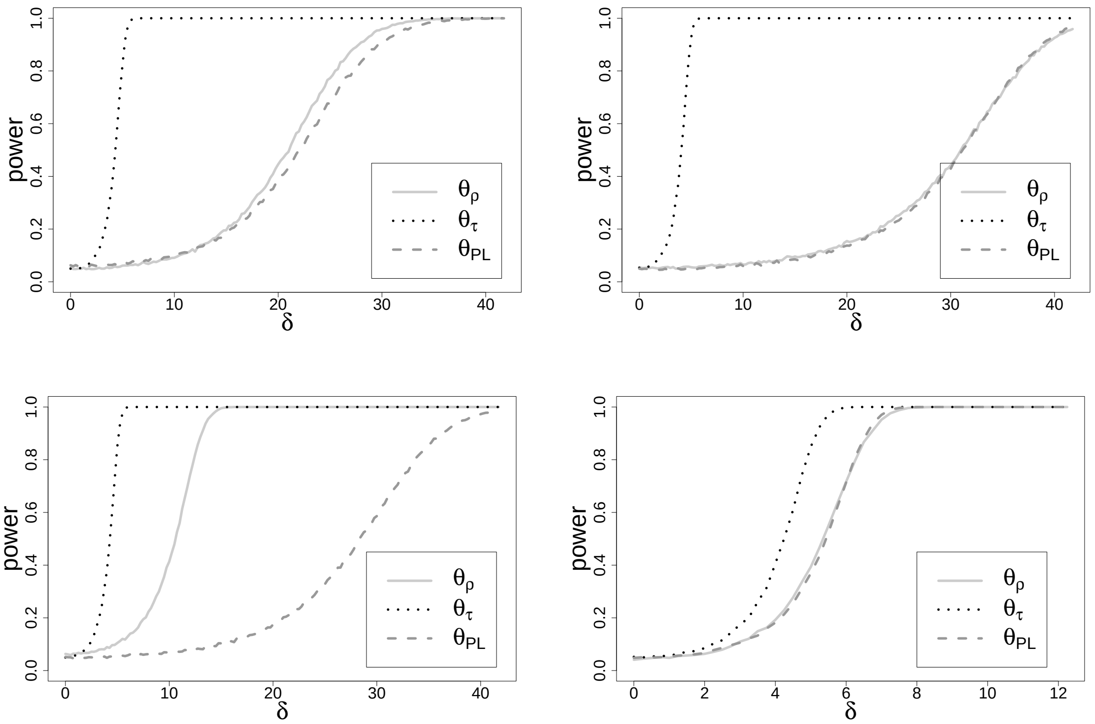

Figure 1: Asymptotic local power curves of the tests based on $\mathcal{V}_{n,N}^{\rho}$, $\mathcal{V}_{n,N}^{\tau}$ and $\mathcal{V}_{n,N}^{PL}$ under mixtures of (a) Clayton, (b) Frank, (c) Gumbel–Barnett and (d) Normal copulas with $\tau_C = 0.1$ and $\tau_D = 0.5$.

Table 2 establishes a clear picture of which statistic is the best under a given scenario of mixture distributions : for Clayton, Gumbel–Barnett and Normal mixtures, $\mathcal{S}_{n1}$ is the most powerful locally, while $\mathcal{S}_{n3}$ is the best for local mixtures of Frank copulas. The test statistic $\mathcal{S}_{n2}$ is very poor in all cases, except when $(\tau_C, \tau_D) = (0.4, 0.8)$ under Gumbel–Barnett alternatives. It is also interesting to note that under Clayton mixtures, $\mathcal{S}_{n1}$ performs better than Shih's statistic $\mathcal{S}_{n4}$, even if the latter is specifically conceived for this particular case. To come to this conclusion, note that $|\mu_4|/\sigma_4 = 0.655$ when $(\tau_C, \tau_D) = (0.1, 0.5)$ and $|\mu_4|/\sigma_4 = 0.347$ when $(\tau_C, \tau_D) = (0.4, 0.8)$.

Figure 3 compares the local power curves of $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$ and $\mathcal{S}_{n3}$ to the best statistic among $\mathcal{V}_{n,N,\rho}$, $\mathcal{V}_{n,N,\tau}$ and $\mathcal{V}_{n,N,PL}$ according to the results of subsection 4.1. Only the case $(\tau_C, \tau_D) = (0.4, 0.8)$ is considered. For the mixture of Clayton copulas, the goodness-of-fit statistic of Shih, suitable only for this family, is also investigated.

The test statistic $\mathcal{S}_{n1}$ exhibit high power locally in all cases, while $\mathcal{S}_{n3}$ also performs very well. The most surprising discovery here is the rather poor performance of the Cramér–von Mises statistics compared to the very simple, asymptotically normal moment-based statistics. These conclusions must however be treated with care since the nature of the alternative distributions considered could have favored the moment-based statistics. Nevertheless, the latter deserve further investigations under other types of alternatives.

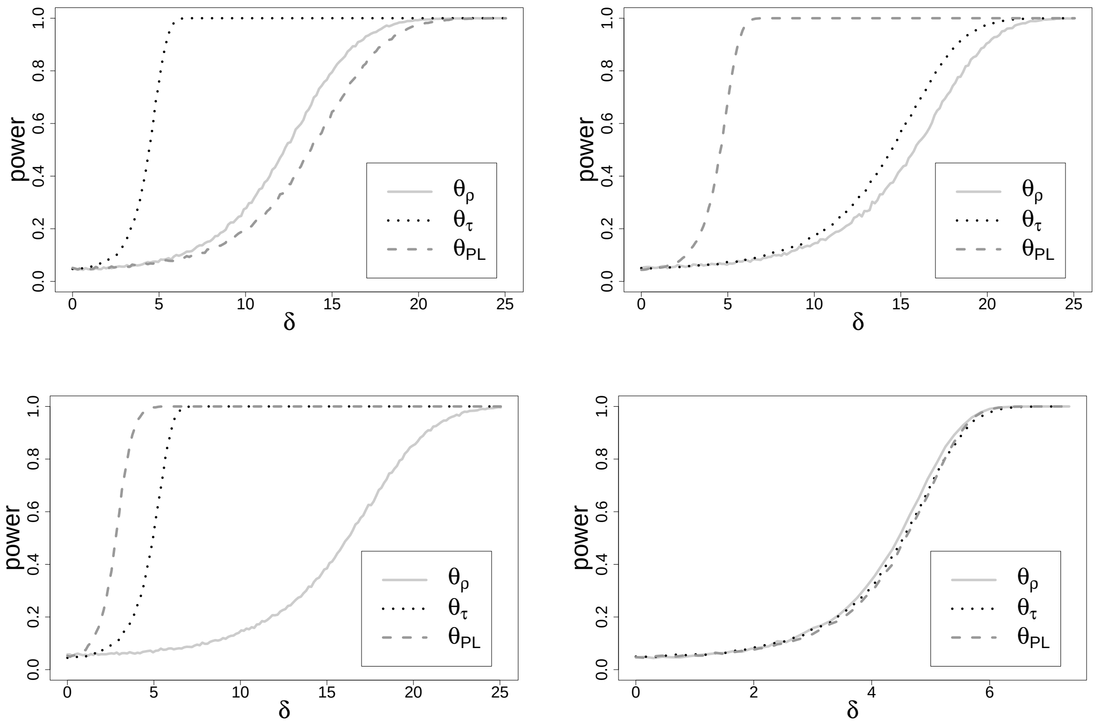

Figure 2: Asymptotic local power curves of the tests based on $\mathcal{V}^{\rho}_{n,N}$, $\mathcal{V}^{\tau}_{n,N}$ and $\mathcal{V}^{PL}_{n,N}$ under mixtures of (a) Clayton, (b) Frank, (c) Gumbel–Barnett and (d) Normal copulas with $\tau_C = 0.4$ and $\tau_D = 0.8$.

Also, multivariate extensions of $\mathcal{S}_{n1}, \ldots, \mathcal{S}_{n4}$ could be considered as serious competitors to $\mathcal{V}^{\rho}_{n,N}$, $\mathcal{V}^{\tau}_{n,N}$ and $\mathcal{V}^{PL}_{n,N}$, the latter being very costly in terms of computing time.

In some cases, e.g. in panel (b) of Figure 3, it is difficult to decide whether $\mathcal{S}_{n2}$ performs better than $\mathcal{V}^{PL}_{n,N}$, locally. A way to circumvent this problem consists in computing some measure of asymptotic relative efficiency. This idea is developed in the next section.

5. Asymptotic relative efficiencies

5.1. A new ARE measure

For a goodness-of-fit statistic whose limiting distribution is normal with mean $\delta\mu(C_\theta, D)$ and variance $\sigma^2(C_\theta)$, the associated local power curve $\beta(\delta, C_\theta, D)$ is an increasing function of $\mu(C_\theta, D)/\sigma(C_\theta, D)$ for all fixed values of $\delta > 0$. It thus seems natural to compare the efficiency of two such statistics $\mathcal{S}_{nj}$ and $\mathcal{S}_{nk}$ via Pitman's measure of asymptotic relative efficiency (ARE), namely

$$\mathcal{ARE}_{\text{Pitman}}(\mathcal{S}_{nj}, \mathcal{S}_{nk}) = \left\{ \frac{\mu_j(C_\theta, D)/\sigma_j(C_\theta)}{\mu_k(C_\theta, D)/\sigma_k(C_\theta)} \right\}^2 .$$

Table 2: Asymptotic local efficiency terms for the test statistics $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$ and $\mathcal{S}_{n3}$ under mixtures of Clayton, Frank, Gumbel–Barnett and Normal copulas

τ_{C_θ}	τ_D	Mixture model	$\mathcal{S}_{n1}$ $\lvert\mu_1\rvert/\sigma_1$	$\mathcal{S}_{n2}$ $\lvert\mu_2\rvert/\sigma_2$	$\mathcal{S}_{n3}$ $\lvert\mu_3\rvert/\sigma_3$	Mixture model	$\mathcal{S}_{n1}$ $\lvert\mu_1\rvert/\sigma_1$	$\mathcal{S}_{n2}$ $\lvert\mu_2\rvert/\sigma_2$	$\mathcal{S}_{n3}$ $\lvert\mu_3\rvert/\sigma_3$
0.1	0.2		**1.627**	0.006	0.227		3.329	0.065	**4.269**
0.1	0.3		**3.163**	0.013	0.442		6.566	0.067	**8.298**
0.1	0.4		**4.608**	0.009	0.632		9.617	0.230	**12.420**
0.1	0.5	Clayton	**5.894**	0.005	0.794	Frank	12.432	0.346	**16.155**
0.4	0.5		**0.762**	0.007	0.234		1.162	0.039	**1.873**
0.4	0.6		**1.438**	0.000	0.426		2.134	0.115	**3.547**
0.4	0.7		**1.989**	0.004	0.594		2.916	0.290	**5.169**
0.4	0.8		**2.403**	0.046	0.765		3.494	0.468	**6.487**
0.1	0.2		**1.920**	0.006	0.289		**3.971**	0.000	0.444
0.1	0.3		**3.728**	0.019	0.568		**7.765**	0.003	0.871
0.1	0.4		**5.446**	0.029	0.831		**11.353**	0.010	1.278
0.1	0.5	Gumbel–	**9.732**	0.338	1.084	Normal	**14.529**	0.023	1.646
0.4	0.5	Barnett	**0.795**	0.017	0.361		**1.459**	0.011	0.446
0.4	0.6		**1.491**	0.056	0.706		**2.748**	0.007	0.824
0.4	0.7		**2.048**	0.112	1.012		**3.768**	0.011	1.131
0.4	0.8		2.336	1.725	**3.099**		**4.462**	0.036	1.367

In bold, the most powerful statistic locally among $\mathcal{S}_{n1}$*,* $\mathcal{S}_{n2}$ *and* $\mathcal{S}_{n3}$

However, it is not entirely clear how to extend this measure in the case when the limiting distribution of a test statistic is no longer normal, which is the case with many of the goodness-of-fit statistics. A generalization of Pitman's measure proposed by Genest *et al.* (2006c) and Genest *et al.* (2007) is

$$\widetilde{\mathcal{ARE}}(\mathcal{S}_{nj}, \mathcal{S}_{nk}) = \lim_{\delta\to 0} \frac{\beta_{\mathcal{S}_{nj}}(\delta) - \beta_{\mathcal{S}_{nj}}(0)}{\beta_{\mathcal{S}_{nk}}(\delta) - \beta_{\mathcal{S}_{nk}}(0)}$$

in terms of the local power functions $\beta_{\mathcal{S}_{nj}}$, $\beta_{\mathcal{S}_{nk}}$ of two tests $\mathcal{S}_{nj}$ and $\mathcal{S}_{nk}$. For most cases of interest, however, this measure requires the derivatives of the power curves in a neighborhood of $\delta = 0$. Since the asymptotic local power functions of the tests based on $\mathcal{V}_{n,N,\rho}$, $\mathcal{V}_{n,N,\tau}$ and $\mathcal{V}_{n,N,PL}$ admit no explicit representations, this causes a serious problem when trying to apply the latter definition.

Here, another generalization of $\mathcal{ARE}_{\text{Pitman}}$ is proposed :

$$\mathcal{ARE}(\mathcal{S}_{nj}, \mathcal{S}_{nk}) = \left\{ \lim_{M\to\infty} \frac{\int_0^M \{1 - \beta_{\mathcal{S}_{nk}}(\delta)\}\, \mathrm{d}\delta}{\int_0^M \{1 - \beta_{\mathcal{S}_{nj}}(\delta)\}\, \mathrm{d}\delta} \right\}^2. \tag{11}$$

The first motivation for such a definition is the possibility to estimate $\int_0^M \{1 - \beta_{\mathcal{S}_{nj}}(\delta)\}\, \mathrm{d}\delta$ and $\int_0^M \{1 - \beta_{\mathcal{S}_{nk}}(\delta)\}\, \mathrm{d}\delta$ when accurate approximations $\hat\beta_{\mathcal{S}_{nj}}$ and $\hat\beta_{\mathcal{S}_{nk}}$ are available. This is the case for the power curves of the tests based on $\mathcal{V}_{n,N}^{\rho}$, $\mathcal{V}_{n,N}^{\tau}$ and $\mathcal{V}_{n,N}^{PL}$. To be specific, suppose $\hat\beta(\delta)$ is available at the points iM/N, $i = 1, \ldots, N$, for sufficiently large

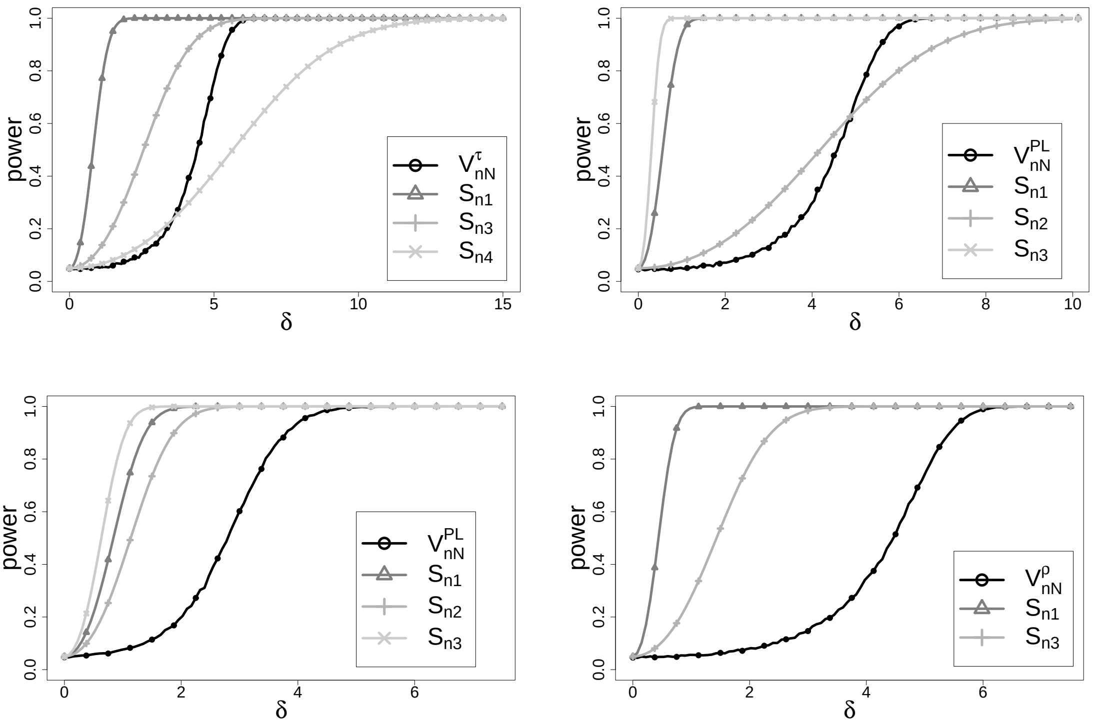

Figure 3: Asymptotic local power of the Cramér–von Mises tests and of $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$, $\mathcal{S}_{n3}$ and $\mathcal{S}_{n4}$ under (a) Clayton, (b) Frank, (c) Gumbel–Barnett and (d) Normal mixtures with $\tau_C = 0.4$ and $\tau_D = 0.8$.

N in order to achieve some numerical accuracy. Upper and lower approximations of $\int_0^M \left\{1 - \beta_{\mathcal{S}_{nj}}(\delta)\right\} \mathrm{d}\delta$ are

$$I_1 = \frac{M}{N} \sum_{i=1}^{N} \left\{1 - \hat{\beta}\left(\frac{iM}{N}\right)\right\} \quad \text{and} \quad I_2 = \frac{M}{N} \sum_{i=0}^{N-1} \left\{1 - \hat{\beta}\left(\frac{iM}{N}\right)\right\},$$

and the chosen approximation, provided M is selected such that $\hat{\beta}(M) = 1$, is

$$\frac{I_1 + I_2}{2} = \frac{M}{N} \sum_{i=1}^{N-1} \left\{1 - \hat{\beta}\left(\frac{iM}{N}\right)\right\} + \frac{M}{N}\left(\frac{1-\alpha}{2}\right).$$

Another interesting feature of $\mathcal{ARE}(\mathcal{S}_{nj}, \mathcal{S}_{nk})$ is the fact that it generalizes Pitman's notion of asymptotic relative efficiency. To see this, let $\beta(\delta) = 1 - \Phi(z_{\alpha/2} - \delta\mu) + \Phi(-z_{\alpha/2} - \delta\mu)$ and compute

$$\begin{aligned} \int_0^\infty \{1 - \beta(\delta)\} \, \mathrm{d}\delta &= \int_0^\infty \Phi\left(z_{\alpha/2} - \delta\mu\right) \mathrm{d}\delta - \int_0^\infty \Phi\left(-z_{\alpha/2} - \delta\mu\right) \mathrm{d}\delta \\ &= \frac{1}{\mu}\left\{\int_{-\infty}^{z_{\alpha/2}} \Phi(x)\mathrm{d}x - \int_{-\infty}^{-z_{\alpha/2}} \Phi(x)\mathrm{d}x\right\} = \frac{1}{\mu}\int_{-z_{\alpha/2}}^{z_{\alpha/2}} \Phi(x)\mathrm{d}x = \frac{z_{\alpha/2}}{\mu}. \end{aligned}$$

As a consequence, one has

$$\int_0^\infty \{1 - \beta_j(\delta, C_\theta, D)\}\, \mathrm{d}\delta = z_{\alpha/2} \left\{ \frac{\mu_j(C_\theta, D)}{\sigma_j(C_\theta)} \right\}^{-1} \tag{12}$$

for local power functions of the form (10). Computations of $\mathcal{ARE}$ for some of the goodness-of-fit statistics encountered in this paper are provided in the next subsection.

5.2. Local efficiency comparisons

In all situations considered in subsection 4.2, the best moment-based statistic locally outperform the best Cramér–von Mises statistic. Hence, it seems useless to compare the latter in terms of their asymptotic relative efficiency. However, since the power curves of $\mathcal{V}^\rho_{n,N}$, $\mathcal{V}^\tau_{n,N}$ and $\mathcal{V}^{PL}_{n,N}$ are often very close to each other, such computations could be very interesting. They are presented in Table 3.

Table 3: Estimated values of $\lim_{M\to\infty} \int_0^M \{1-\beta(\delta)\}\mathrm{d}\delta$ for the goodness-of-fit statistics $\mathcal{V}^\rho_{n,N}$, $\mathcal{V}^\tau_{n,N}$ and $\mathcal{V}^{PL}_{n,N}$ and asymptotic relative efficiencies under mixtures of Clayton, Frank, Gumbel–Barnett and Normal copulas.

Mixture model	τ_{C_θ}	τ_D	$\lim_{M\to\infty} \int_0^M \{1-\beta(\delta)\}\mathrm{d}\delta$ $\mathcal{V}^\rho_{n,N}$	$\mathcal{V}^\tau_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	Asymptotic relative efficiency $(\mathcal{V}^\rho_{n,N}, \mathcal{V}^\tau_{n,N})$	$(\mathcal{V}^\rho_{n,N}, \mathcal{V}^{PL}_{n,N})$	$(\mathcal{V}^\tau_{n,N}, \mathcal{V}^{PL}_{n,N})$
Clayton	0.1	0.5	12.018	2.540	12.618	0.211	1.050	4.968
	0.4	0.8	23.469	8.349	26.091	0.356	1.112	3.125
Frank	0.1	0.5	17.464	2.381	17.594	0.136	1.007	7.389
	0.4	0.8	29.483	27.079	8.670	0.918	0.294	0.320
Gumbel–Barnett	0.1	0.5	5.954	2.506	16.143	0.421	2.711	6.442
	0.4	0.8	30.369	9.282	5.475	0.306	0.180	0.590
Normal	0.1	0.5	3.142	2.491	3.150	0.793	1.003	1.265
	0.4	0.8	8.390	8.527	8.609	1.016	1.026	1.010

These computations show, among other things, that $\mathcal{V}^\tau_{n,N}$ is generally more powerful than $\mathcal{V}^{PL}_{n,N}$ for low dependence alternatives, i.e. close to independence. An opposite conclusion arises for mixture of high dependence copulas, namely when $(\tau_C, \tau_D) = (0.4, 0.8)$. The performance of $V^{PL}_{n,N}$ and $V^\rho_{n,N}$ are quite similar for low dependence, except under Gumbel–Barnett mixtures. Overall, $V^\tau_{n,N}$ seems the best choice close to the independence copula, while $V^{PL}_{n,N}$ performs well under high levels of dependence.

Looking back at panel (b) of Figure 3, it is difficult to decide whether $\mathcal{S}_{n2}$ performs better than $\mathcal{V}^{PL}_{n,N}$. Even though the local power curve of $\mathcal{V}^{PL}_{n,N}$ reaches 1 more quickly, the asymptotic relative efficiency is given by $\mathcal{ARE}(\mathcal{V}_{n,N}^{PL}, \mathcal{S}_{n2}) = 0.950$, which supports the choice of $\mathcal{S}_{n2}$ if a mixture of Frank distributions is suspected as a possible alternative.

6. Sensitivity in small samples

This section is devoted to the sensitivity in small samples and under fixed alternatives of the test statistics encountered in this paper, namely $\mathcal{V}_{n,N}^{\rho}$, $\mathcal{V}_{n,N,}^{\tau}$, $\mathcal{V}_{n,N}^{PL}$, $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$, $\mathcal{S}_{n3}$ and $\mathcal{S}_{n4}$. The main goal is to relate the asymptotic local efficiency results of Section 4 and Section 5 with empirical situations. In subsection 6.1, the specific influence of the estimators on the power of the Cramér–von Mises statistics is investigated. In subsection 6.2, comparisons with the moment-based statistics are made. These results will be paralleled with those presented in subsections 4.1 and 4.2 under contiguous sequences.

6.1. Influence of the estimators on the power of the Cramér–von Mises statistics

It was seen in subsection 4.1 that the asymptotic local powers of the goodness-of-fit tests based on the empirical copula process are sensitive to the choice of the estimator of the dependence parameter, at least under the mixture distributions considered. In this section, the ability of $\mathcal{V}_{n,N}^{\rho}$, $\mathcal{V}_{n,N}^{\tau}$ and $\mathcal{V}_{n,N}^{PL}$ to reject the null hypothesis is first examined under fixed alternatives and many sample sizes. The results can be found in Tables 4–7. First note that all 5% nominal levels are maintained, keeping in mind a margin of error of the magnitude of $\pm 1\%$ when estimating proportions from 10 000 replicates.
When Clayton's family of copulas is in the null hypothesis, one can see from Table 4 that $\mathcal{V}_{n,N}^{\rho}$ performs very well against all alternatives, especially in small samples, while $\mathcal{V}_{n,N}^{\tau}$ is almost as powerful. The latter are significantly superior to $\mathcal{V}_{n,N}^{PL}$ under Gumbel–Barnett alternatives, especially in small samples. The performance of $\mathcal{V}_{n,N}^{PL}$ however surpasses that of $\mathcal{V}_{n,N}^{\rho}$ and $\mathcal{V}_{n,N}^{\tau}$ under Frank and Normal alternatives, and this advantage is particularly marked for higher degrees of dependence.

Things are much simpler in Table 5 when testing the membership to Frank's family, where the three considered estimation strategies yield almost the same power for the Cramér–von Mises statistics. For the null hypothesis of belonging to Gumbel–Barnett's class, the statistic $\mathcal{V}_{n,N}^{PL}$ is remarkably better than its two competitors under Frank and Normal alternatives, especially for large sample sizes, as one can notice from the entries in Table 6. An opposite conclusion must however be made under Clayton alternatives, where $\mathcal{V}_{n,N}^{\rho}$ and $\mathcal{V}_{n,N}^{\tau}$ are slightly better.

Finally, the most powerful statistics for testing the Normal hypothesis are $\mathcal{V}_{n,N}^{\rho}$ and $\mathcal{V}_{n,N}^{\tau}$ under Clayton alternatives, while $\mathcal{V}_{n,N}^{PL}$ is the best choice under observations that come from the Frank copula. Here again, the performance of the latter increases as the sample size becomes larger.

In a second series of analyses, the power of the Cramér–von Mises statistics under mixture distributions of the type $\mathcal{Q}_{\delta_n} = (1 - \delta_n)C_\theta + \delta_n C_{\theta'}$ have been considered for samples of size $n = 500$. The corresponding empirical power curves are presented in Figure 4. In this setting, $100 \times \delta/\sqrt{500}$ % of the observations come from the distribution $C_{\theta'}$, so the power increases with δ. However, from a certain threshold, the observed powers suddenly decreases toward the nominal level. This occurs because $C_{\theta'}$ also belongs to the

Table 4: Estimated percentage of rejection of the null hypothesis of belonging to Clayton's family for the goodness-of-fit tests based on $\mathcal{V}^{\rho}_{n,N}$, $\mathcal{V}^{\tau}_{n,N}$, $\mathcal{V}^{PL}_{n,N}$, $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$ and $\mathcal{S}_{n3}$ under fixed copula alternatives.

		$\mathcal{H}_1$: Clayton						$\mathcal{H}_1$: Gumbel–Barnett					
n	τ	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$
100	0.10	5.0	4.5	4.2	3.2	3.4	6.2	22.4	20.2	12.0	2.7	11.2	6.0
	0.15	5.7	5.0	5.2	3.1	4.2	7.2	38.7	36.6	24.0	2.3	18.6	11.0
	0.20	6.0	5.5	5.6	2.8	4.5	6.8	55.4	53.3	38.6	1.4	27.6	19.1
250	0.10	5.1	4.6	5.1	3.9	4.9	6.1	37.7	36.1	26.6	2.4	33.8	26.6
	0.15	5.3	5.1	5.0	4.0	4.9	5.9	65.4	64.6	53.2	1.7	58.2	49.9
	0.20	5.1	5.3	5.1	3.1	5.1	5.9	86.2	85.5	77.7	1.1	78.0	74.5
500	0.10	5.0	5.0	4.3	3.5	5.1	5.3	57.1	54.9	46.6	1.3	64.8	59.8
	0.15	5.6	5.1	4.8	4.8	4.9	5.7	86.6	86.2	79.9	1.4	90.1	88.2
	0.20	5.0	5.2	5.3	3.6	5.1	5.8	97.5	97.3	95.6	1.1	98.4	98.2
1000	0.10	5.1	5.1	4.7	3.0	5.2	5.4	73.3	73.8	69.7	0.6	90.5	89.9
	0.15	4.8	5.3	5.3	5.0	5.5	5.7	97.4	97.5	96.0	0.8	99.7	99.7
	0.20	5.1	5.3	5.3	4.8	4.9	5.2	99.9	100	99.9	2.2	100	100
2500	0.10	4.7	4.7	6.2	4.2	4.8	5.2	90.1	89.8	90.5	0.4	99.9	99.9
	0.15	4.8	4.8	5.5	4.7	4.3	4.6	99.9	99.9	99.9	0.3	100	100
	0.20	4.5	5.3	5.5	5.8	5.9	5.2	100	100	100	13.0	100	100

		$\mathcal{H}_1$: Frank						$\mathcal{H}_1$: Normal					
n	τ	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$
100	0.10	12.8	11.6	8.1	3.9	7.5	3.8	12.3	11.2	6.7	3.4	5.6	3.1
	0.15	20.8	19.2	13.9	4.3	11.0	4.8	20.2	18.4	11.5	3.6	8.0	3.9
	0.20	31.3	29.5	23.9	5.1	17.6	9.3	29.4	26.9	17.9	3.2	11.5	6.1
250	0.10	18.8	18.3	17.8	4.7	24.0	16.4	18.6	17.3	12.7	4.0	17.4	11.4
	0.15	36.0	34.8	36.5	7.3	44.0	34.2	33.0	32.3	25.5	5.3	29.4	21.1
	0.20	55.7	54.6	58.3	9.3	64.7	56.5	49.7	47.8	41.2	5.0	44.3	36.1
500	0.10	28.7	27.1	32.0	5.3	49.9	42.8	25.8	24.6	20.8	3.9	35.8	29.1
	0.15	54.5	52.9	61.3	10.2	81.2	75.4	48.6	46.9	43.6	7.0	61.4	53.7
	0.20	77.1	76.1	84.7	14.1	95.3	93.6	69.7	68.4	66.7	7.9	81.8	78.0
1000	0.10	37.3	37.8	50.5	5.5	81.4	78.2	33.8	33.0	35.5	3.9	63.6	59.5
	0.15	72.3	72.3	83.5	15.7	98.7	98.1	66.7	65.8	65.8	10.2	92.5	89.4
	0.20	92.4	92.8	97.6	25.0	100	100	88.2	87.3	89.4	15.2	99.0	98.8
2500	0.10	50.2	48.5	73.3	9.0	99.6	99.6	43.4	42.4	52.4	5.8	96.1	95.9
	0.15	88.9	88.6	96.5	20.5	100	100	83.5	82.2	88.0	14.0	100	100
	0.20	99.1	99.2	99.9	40.4	100	100	97.5	97.8	98.6	28.7	100	100

Table 5: Estimated percentage of rejection of the null hypothesis of belonging to Frank's family for the goodness-of-fit tests based on $\mathcal{V}^{\rho}_{n,N}$, $\mathcal{V}^{\tau}_{n,N}$, $\mathcal{V}^{PL}_{n,N}$, $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$ and $\mathcal{S}_{n3}$ under fixed copula alternatives.

		$\mathcal{H}_1$: Clayton						$\mathcal{H}_1$: Gumbel–Barnett					
n	τ	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$
100	0.10	7.1	6.4	6.4	4.0	1.1	6.5	10.5	9.6	10.0	3.0	1.0	6.5
	0.15	11.3	10.2	10.4	3.6	1.3	8.3	15.2	14.2	14.6	2.4	1.0	8.5
	0.20	16.3	14.8	15.5	2.6	1.1	10.1	17.3	17.0	17.3	1.8	1.3	11.0
250	0.10	12.7	12.0	12.9	3.2	1.6	9.4	12.6	12.8	12.9	2.1	0.8	9.6
	0.15	24.8	24.8	26.0	2.3	1.4	15.1	19.7	19.6	20.3	0.9	1.6	15.4
	0.20	43.4	43.5	43.7	1.4	2.3	18.9	28.3	29.6	29.6	0.5	3.7	20.3
500	0.10	22.6	22.1	21.7	2.8	1.1	15.9	16.5	16.2	16.0	1.0	1.0	16.1
	0.15	47.3	47.1	47.0	1.8	2.4	25.9	28.3	28.6	28.8	0.5	5.1	27.1
	0.20	73.2	74.2	73.7	1.0	6.5	31.8	42.8	45.7	45.1	1.0	14.4	34.1
1000	0.10	36.4	39.3	38.5	2.2	1.5	26.8	21.4	22.9	22.7	0.4	3.6	29.0
	0.15	72.5	73.1	72.1	1.3	8.4	41.7	41.1	42.8	41.9	0.9	19.2	45.8
	0.20	92.8	92.9	93.2	1.3	18.0	47.9	60.6	62.1	63.9	8.0	37.1	54.7
2500	0.10	53.2	52.2	51.9	1.4	8.5	51.1	26.0	26.2	26.5	1.0	26.3	59.5
	0.15	90.6	91.1	91.5	1.1	32.3	74.8	53.6	56.0	56.9	19.8	66.0	82.4
	0.20	99.6	99.5	99.4	10.9	54.1	79.9	79.9	79.7	81.2	65.9	85.6	88.7

		$\mathcal{H}_1$: Frank						$\mathcal{H}_1$: Normal					
n	τ	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$
100	0.10	6.0	5.6	5.5	4.6	2.3	4.3	6.0	5.1	5.4	3.8	2.1	4.7
	0.15	6.0	5.7	5.6	4.5	2.8	4.5	6.4	5.7	5.7	3.6	2.5	5.0
	0.20	5.4	5.4	5.2	4.5	3.3	4.7	6.7	6.4	6.7	3.3	2.8	5.9
250	0.10	4.8	4.9	4.9	4.2	3.1	4.5	6.2	5.6	6.0	3.3	3.2	5.3
	0.15	4.8	4.7	4.7	3.9	3.4	4.6	6.6	6.0	6.3	2.8	3.5	6.6
	0.20	4.5	5.1	4.8	4.2	3.6	4.7	8.3	7.7	7.9	2.0	3.0	7.8
500	0.10	4.6	4.6	4.5	3.9	4.1	4.6	6.2	5.6	5.4	2.7	3.7	6.6
	0.15	4.7	4.9	4.6	4.5	4.2	5.2	8.0	7.5	7.6	2.2	4.3	8.1
	0.20	5.0	5.3	5.1	4.7	4.4	5.1	10.8	11.4	10.1	1.6	4.1	8.6
1000	0.10	4.3	5.9	5.1	4.7	4.7	5.3	7.3	7.6	6.8	2.5	4.1	8.2
	0.15	5.4	5.7	5.1	5.0	5.0	5.1	9.8	10.3	9.2	1.6	3.9	8.9
	0.20	4.8	5.1	5.1	4.7	4.9	4.7	14.7	14.3	13.8	1.0	3.5	8.8
2500	0.10	5.3	5.1	4.3	4.6	4.6	4.4	7.6	7.4	6.8	1.7	4.5	9.0
	0.15	5.0	5.5	5.0	4.8	5.1	5.7	10.9	11.3	11.2	1.2	4.8	11.2
	0.20	5.4	5.0	5.3	4.2	4.7	4.5	17.2	16.2	17.3	0.6	4.0	9.5

Table 6: Estimated percentage of rejection of the null hypothesis of belonging to Gumbel–Barnett's family for the goodness-of-fit tests based on $\mathcal{V}^{\rho}_{n,N}$, $\mathcal{V}^{\tau}_{n,N}$, $\mathcal{V}^{PL}_{n,N}$, $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$ and $\mathcal{S}_{n3}$ under fixed copula alternatives.

		$\mathcal{H}_1$: Clayton						$\mathcal{H}_1$: Gumbel–Barnett					
n	τ	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$
100	0.10	9.3	7.9	5.4	6.2	16.5	5.8	3.6	3.7	4.9	5.7	5.9	7.0
	0.15	18.8	16.8	12.1	7.6	22.4	7.2	4.7	4.3	5.2	5.6	5.1	6.3
	0.20	31.5	29.9	22.2	7.8	26.7	9.6	4.9	4.7	5.3	5.3	4.4	4.8
250	0.10	26.8	25.1	20.7	8.4	36.3	24.2	4.6	4.5	5.4	5.5	5.8	5.9
	0.15	53.9	52.4	45.7	9.4	53.1	38.4	4.8	4.9	4.9	4.9	5.0	4.8
	0.20	78.9	77.5	70.6	10.4	67.9	55.1	5.2	5.0	5.0	5.1	4.7	4.6
500	0.10	48.8	48.0	41.7	9.9	59.8	51.8	4.7	5.1	4.8	4.8	4.9	4.8
	0.15	83.2	83.0	78.6	12.4	83.8	78.1	5.1	5.0	5.6	4.1	5.0	4.7
	0.20	96.9	96.8	95.6	13.0	94.2	92.3	4.9	4.4	4.7	4.4	5.1	4.9
1000	0.10	73.5	72.4	69.1	13.6	88.8	85.9	5.0	4.8	5.6	4.8	4.9	4.8
	0.15	97.5	97.0	96.7	19.3	98.9	98.7	5.1	4.9	5.3	4.5	4.8	5.1
	0.20	100	100	99.9	22.5	99.9	99.9	5.3	4.9	5.3	4.6	5.0	4.8
2500	0.10	92.6	91.6	90.4	16.3	99.9	99.9	5.8	4.9	5.0	3.5	5.6	5.4
	0.15	99.9	99.9	99.9	35.0	100	100	5.6	5.2	5.6	4.4	5.2	4.9
	0.20	100	100	100	41.7	100	100	5.5	5.0	5.3	5.2	5.1	5.5

		$\mathcal{H}_1$: Frank						$\mathcal{H}_1$: Normal					
n	τ	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$
100	0.10	4.2	3.8	3.2	7.7	13.2	4.7	4.0	3.6	3.0	7.2	10.3	4.1
	0.15	5.4	5.2	5.1	9.7	16.4	4.6	4.9	4.3	3.8	8.5	11.9	3.7
	0.20	7.3	7.3	7.1	12.9	20.1	4.8	5.6	5.0	4.4	9.5	12.2	3.0
250	0.10	7.2	7.4	10.1	10.2	27.7	16.5	6.4	6.3	6.6	9.5	18.4	10.7
	0.15	12.8	13.0	18.1	15.3	42.0	26.0	10.4	9.8	9.8	11.9	25.3	13.4
	0.20	18.2	18.4	27.2	22.1	55.8	38.8	15.2	13.7	13.6	16.1	32.1	17.9
500	0.10	12.4	12.9	21.4	14.4	48.7	38.8	10.3	10.2	12.0	12.6	31.4	22.4
	0.15	22.2	23.1	37.5	24.1	73.3	63.6	18.3	16.8	19.4	18.5	47.3	35.2
	0.20	36.2	35.4	56.1	36.5	89.9	84.4	26.4	23.9	28.6	27.7	62.7	49.7
1000	0.10	18.7	19.2	36.8	19.5	79.6	74.5	14.9	14.2	20.7	16.9	57.4	50.1
	0.15	37.0	36.1	62.3	38.8	97.0	95.7	28.1	25.6	35.5	32.3	79.8	73.5
	0.20	56.9	57.1	81.9	58.3	99.7	99.6	41.0	40.7	51.6	50.3	91.1	86.3
2500	0.10	28.4	25.5	54.8	25.8	99.6	99.5	23.1	20.3	31.8	21.6	94.3	92.5
	0.15	54.7	55.1	81.4	65.9	100	100	39.8	40.6	54.8	62.1	99.4	99.2
	0.20	79.8	79.5	95.2	88.0	100	100	62.6	60.5	72.8	85.1	100	100

Table 7: Estimated percentage of rejection of the null hypothesis of belonging to the Normal family for the goodness-of-fit tests based on $\mathcal{V}^{\rho}_{n,N}$, $\mathcal{V}^{\tau}_{n,N}$, $\mathcal{V}^{PL}_{n,N}$, $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$ and $\mathcal{S}_{n3}$ under fixed copula alternatives.

		$\mathcal{H}_1$: Clayton						$\mathcal{H}_1$: Gumbel–Barnett					
n	τ	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$
100	0.10	4.8	4.4	4.5	4.9	5.0	6.2	7.1	6.8	7.0	4.1	5.4	6.7
	0.15	7.7	7.5	7.7	4.3	3.9	6.0	10.5	9.6	9.6	2.9	5.1	6.9
	0.20	12.3	12.2	11.4	4.3	4.2	5.9	13.9	13.5	13.5	2.4	5.1	6.9
250	0.10	10.5	9.6	9.5	5.0	4.2	6.8	10.3	9.1	9.7	2.8	5.4	8.3
	0.15	21.7	21.4	19.6	4.5	4.5	7.3	15.5	15.6	14.6	2.0	5.5	8.2
	0.20	36.6	37.4	32.9	3.2	3.9	6.5	21.5	21.4	19.7	1.2	5.3	8.0
500	0.10	19.7	20.5	16.9	4.8	5.5	8.3	14.1	13.7	12.8	2.1	6.6	9.6
	0.15	41.3	42.3	36.7	3.5	5.8	8.7	22.6	22.8	19.8	0.8	7.2	9.5
	0.20	65.1	65.4	58.9	2.6	5.3	7.2	32.9	32.8	28.8	1.6	6.6	8.3
1000	0.10	33.7	31.8	28.2	4.6	8.0	10.5	18.6	16.8	15.9	1.1	9.5	11.7
	0.15	64.3	64.2	59.1	2.7	10.4	12.7	31.3	30.2	27.7	1.7	11.5	12.1
	0.20	87.8	88.2	84.2	2.6	9.8	11.1	46.0	48.4	43.4	10.2	10.9	10.3
2500	0.10	43.2	45.1	41.9	3.0	15.3	17.8	18.1	19.7	18.7	1.6	18.4	18.4
	0.15	83.1	85.1	82.1	3.3	21.9	23.2	37.9	41.6	37.8	27.0	21.8	18.3
	0.20	98.6	98.5	98.0	11.2	25.8	24.9	61.9	62.2	59.0	67.0	26.7	18.4

		$\mathcal{H}_1$: Frank						$\mathcal{H}_1$: Normal					
n	τ	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$	$\mathcal{V}^{\rho}_{n,N}$	$\mathcal{V}^{\tau}_{n,N}$	$\mathcal{V}^{PL}_{n,N}$	$\mathcal{S}_{n1}$	$\mathcal{S}_{n2}$	$\mathcal{S}_{n3}$
100	0.10	4.1	4.0	4.3	5.2	11.7	7.4	4.5	4.1	4.1	4.9	6.0	5.1
	0.15	4.5	4.4	5.1	5.9	15.6	10.0	4.8	4.7	4.4	4.9	5.6	4.7
	0.20	5.9	6.1	7.3	7.3	20.9	14.7	4.7	4.7	5.0	5.3	6.3	5.4
250	0.10	5.0	5.0	5.8	6.1	15.9	11.2	4.6	4.1	4.2	5.2	5.2	5.2
	0.15	6.3	6.7	9.7	7.2	25.6	19.6	4.8	4.9	5.1	5.1	5.3	5.3
	0.20	7.3	7.8	12.7	9.4	35.3	30.2	4.5	4.8	4.6	4.9	5.0	5.4
500	0.10	6.1	6.6	8.1	6.8	22.7	18.2	5.0	5.0	4.5	4.9	4.7	5.4
	0.15	7.1	8.2	13.2	8.8	36.9	31.5	5.2	5.4	4.7	5.1	5.2	5.6
	0.20	9.5	11.1	20.1	11.1	55.2	51.2	4.9	4.9	4.6	5.2	4.9	4.9
1000	0.10	7.5	7.1	11.2	8.4	34.9	30.0	6.2	5.2	5.0	5.5	5.3	5.1
	0.15	8.7	9.5	19.5	11.0	59.8	55.3	5.5	5.1	4.9	5.4	4.8	4.6
	0.20	13.2	14.7	31.5	13.8	82.2	80.4	5.0	5.1	5.4	4.5	5.1	5.1
2500	0.10	6.4	7.2	14.1	9.8	63.7	60.7	4.3	5.0	4.9	4.5	4.8	4.8
	0.15	10.0	11.7	26.7	14.5	91.7	90.5	4.4	5.2	5.1	5.4	4.8	4.7
	0.20	17.7	18.1	45.1	17.0	99.3	99.4	5.2	5.1	5.3	5.0	5.3	5.7

family of copulas under $\mathcal{H}_0$. One may have expected, however, that the powers would start to decrease at the middle point, i.e. when $\delta = \sqrt{500}/2 \approx 11.2$. The observed asymmetry in all four cases is probably an indication that the goodness-of-fit tests are better to detect discrepancies from $\mathcal{H}_0$ when the data come from a copula with a high level of dependence. The fact that $\theta' > \theta$ probably explained that the middle point is skewed to the right.

As expected, the differences in power between $\mathcal{V}^{\rho}_{n,N}$, $\mathcal{V}^{\tau}_{n,N}$ and $\mathcal{V}^{PL}_{n,N}$ are less apparent in small sample sizes than it was asymptotically (see Figure 2 to compare). Nevertheless, the conclusions here are very similar to the asymptotic situation, except that the performance of $\mathcal{V}^{\rho}_{n,N}$ is not as bad as for $n \to \infty$ under Clayton and Gumbel-Barnett mixtures. Briefly, the choice of the estimator doesn't seem to have a significant influence under Gumbel–Barnett and Normal mixtures, while for Clayton mixtures, the pseudo-likelihood estimator is not recommended. The latter is however the best choice under Frank mixtures.

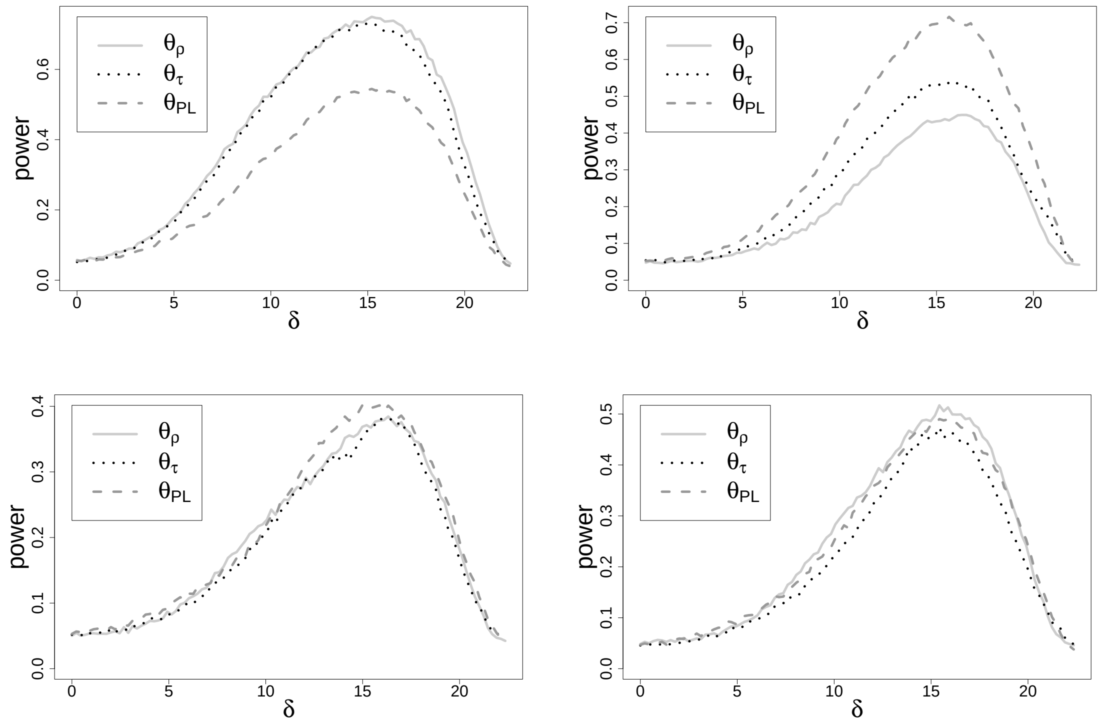

Figure 4: Power curves for the tests based on $\mathcal{V}^{\rho}_{n,N}$, $\mathcal{V}^{\tau}_{n,N}$ and $\mathcal{V}^{PL}_{n,N}$ under (a) Clayton, (b) Frank, (c) Gumbel–Barnett and (d) Normal mixtures with $(\tau_C, \tau_d) = (0.4, 0.8)$, $n = 500$ and $N = 2500$.

6.2. Power of the Cramér–von Mises statistics compared to the moment-based statistics

It was seen in subsection 6.1 that the test statistic $\mathcal{V}^{\rho}_{n,N}$ was a good choice for small sample sizes when testing the goodness-of-fit under the hypothesis of belonging to the Clayton family. The ability to reject $\mathcal{H}_0$ in that case is almost as good for tests based on $\mathcal{S}_{n2}$ and $\mathcal{S}_{n3}$, with a slight advantage to $\mathcal{S}_{n2}$. The power of the latter even becomes larger than that of $\mathcal{V}^{\rho}_{n,N}$ when $n \geq 500$ and is often better than the best Cramér–von Mises statistic in large samples, namely $\mathcal{V}^{PL}_{n,N}$. Note the poor performance of $\mathcal{S}_{n1}$ in all cases considered.

When testing the hypothesis of belonging to the Frank family, $\mathcal{S}_{n1}$ and $\mathcal{S}_{n2}$ are bad choices. However, $\mathcal{S}_{n3}$ is sometimes comparable with the Cramér–von Mises statistics when the sample size is large, especially under Gumbel–Barnett alternatives.

The null hypothesis of a Gumbel–Barnett family provides an example of a very powerful moment-based statistic. Here, $\mathcal{S}_{n2}$ is more powerful than the best Cramér–von Mises statistic, namely $\mathcal{V}^{\rho}_{n,N}$ under Clayton and $\mathcal{V}^{PL}_{n,N}$ under Frank and Normal copulas. Another example is given when testing the hypothesis of belonging to the Normal family against Frank alternatives, where $\mathcal{S}_{n2}$ and $\mathcal{S}_{n3}$ are clearly the most powerful. The latter are unfortunately inefficient to detect Clayton and Gumbel–Barnett dependence structures.

A final analysis have been made to compare the power of the tests under $\mathcal{Q}_{\delta_n} = (1-\delta_n)C_\theta + \delta_n C_{\theta'}$. The results are to be found in Figure 5. Here, the ordering in the power curves are often quite different to the ones encountered in Figure 3 in the asymptotic situation. An explanation probably lies in the fact that the moment-based statistics are especially good in very large samples, and the result is that the latter outclass the Cramér–von Mises statistics when $n \to \infty$. This domination is weaker in moderate sample sizes. This is particularly evident under Clayton mixtures where the best Cramér–von Mises statistic outperforms all moment-based statistics. Note here the very poor performance of $\mathcal{S}_{n1}$, in contrast to the extremely good performance of the same statistic when $n \to \infty$. Under Frank mixtures the moment-based statistics perform very well even for moderate sample sizes, where they outperform the best Cramér–von Mises statistic. Under Gumbel–Barnett mixtures, $\mathcal{S}_{n1}$ is clearly the best statistic while under Normal mixtures, $\mathcal{S}_{n3}$ is the best and $\mathcal{S}_{n1}$ provides a very poor performance.

7. Discussion

In this paper, the local power curves of tests based on Cramér–von Mises distances of the empirical copula goddness-of-fit process have been investigated and compared to that of moment-based statistics involving Spearman's rho, Kendall's tau and the pseudo-maximum likelihood estimator. Many discoveries have been made, in particular that the estimation strategy can have a significant impact on the power of the Cramér–von Mises statistics, and that some of the moment-based statistics provide very powerful tests under many distributional scenarios. Also, it seems that the ability of the Cramér–von Mises statistics to detect departures from $\mathcal{H}_0$ is better under fixed alternatives rather than under mixtures, while an opposite conclusion can be expressed for the moment-based statistics.

In future works, these kind of investigations could also be accomplished for other popular goodness-of-fit tests like those proposed by Scaillet (2006), Huard *et al.* (2006) and Genest *et al.* (2006a). The latter authors based their tests on Kendall's process $\mathcal{K}_n(t) = \sqrt{n}\{K_n(t) - K_{\hat\theta_n}(t)\}$, where $K_\theta(t) = \mathrm{P}\{C_\theta(X,Y) \leq t\}$, with $(X,Y) \sim C_\theta$, is the bivariate probability integral transformation of C_θ and K_n is a fully nonparametric estimator of K_θ. Suitable adaptations of the arguments to be found in Ghoudi & Rémillard (1998) should enable to establish that $\mathcal{K}_n \rightsquigarrow \mathcal{K} + \delta(\dot L_0 - \mu \dot K_\theta)$ under alternatives of the type $\mathcal{Q}_{\delta_n}$, where $\mathcal{K}$ is the weak limit of $\mathcal{K}_n$ under $\mathcal{H}_0$, L_δ is the probability integral transformation of $\mathcal{Q}_\delta$ and μ is the drift term associated to the limit of $\Theta_n = \sqrt{n}(\hat\theta_n - \theta)$ identified in Proposition 1.

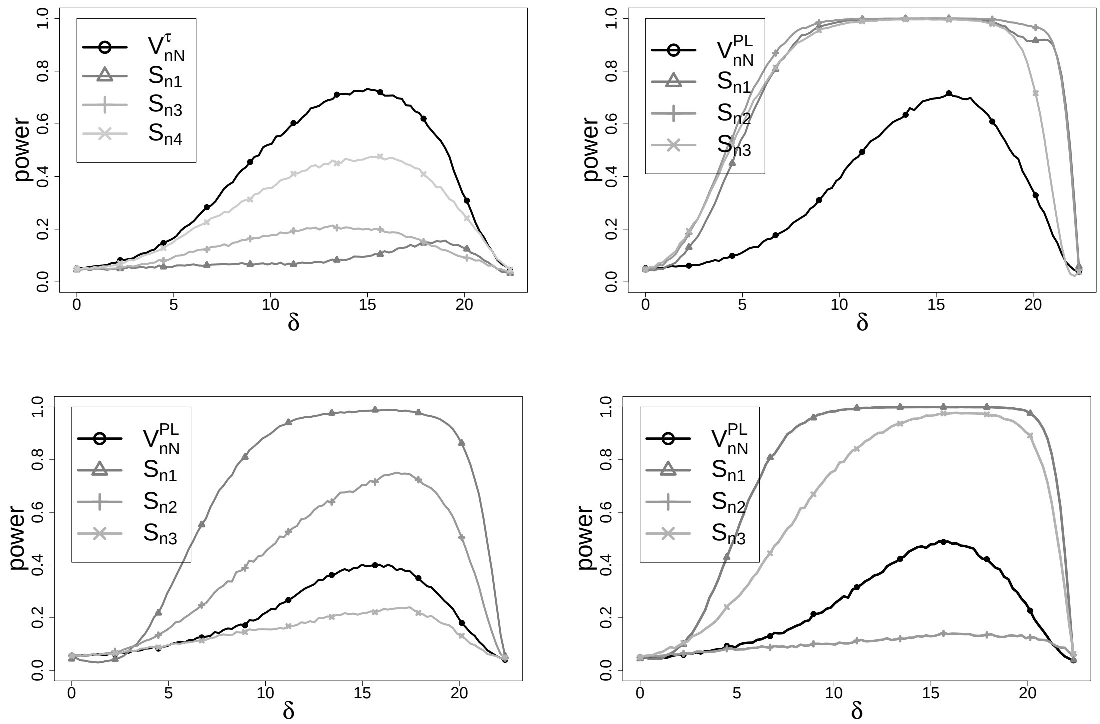

Figure 5: Power of the tests based on $\mathcal{V}_{n,N}$, $\mathcal{S}_{n1}$, $\mathcal{S}_{n2}$, $\mathcal{S}_{n3}$ and $\mathcal{S}_{n4}$ when $n = 500$ under (a) Clayton, (b) Frank, (c) Gumbel–Barnett and (d) Normal mixtures with $\tau_C = 0.4$ and $\tau_D = 0.8$

It could also be interesting to exploit the idea of moment-based statistics to test the fit to families of multivariate copulas. For example, possible estimators of a univariate parameter θ are those based on inversions of the multivariate extensions of Spearman's rho described by Schmid & Schmidt (2007), namely

$$\rho_{n,\star} = \xi(d)\left\{2^d \int_{(0,1)^d} C_n(u)\mathrm{d}u - 1\right\} \quad \text{and} \quad \rho_{n,\star\star} = \xi(d)\left\{2^d \int_{(0,1)^d} \bar{C}_n(u)\mathrm{d}u - 1\right\},$$

where $\xi(d) = (d+1)(2^d - d - 1)^{-1}$, C_n is the multivariate empirical copula and $\bar{C}_n$ is the survival version of C_n. Then, the local behavior of the goodness-of-fit statistic

$$\mathcal{S}_n = \sqrt{n}\left\{\rho_\star^{-1}\left(\rho_{n,\star}\right) - \rho_{\star\star}^{-1}\left(\rho_{n,\star\star}\right)\right\},$$

where $\rho_\star$ and $\rho_{\star\star}$ are the population versions of $\rho_{n,\star}$ and $\rho_{n,\star\star}$, will be a consequence of that of $\mathcal{C}_{n,\theta}$ that can be deduced from the proof of Proposition 2.

It may be noted that the form of the alternative hypothesis (1) is not the only one under which asymptotic power curves could be derived. Another possibility is given by

$$\mathcal{Q}^\star_\delta(x,y) = \psi_\delta^{-1}\left[C\left\{\psi_\delta(x), \psi_\delta(y)\right\}\right],$$

where ψ_δ must satisfy some conditions to ensure that $\mathcal{Q}_\delta^\star$ is a copula and the perturbation function ψ_δ is chosen such that $\psi_0(t) = t$. Then, by arguments similar to that in the proof of Proposition 2, it would be possible to establish that $\mathcal{C}_{n,\theta} \rightsquigarrow \mathcal{C}_\theta + \delta \dot{\mathcal{Q}}_0^\star$, where

$$\dot{\mathcal{Q}}_0^\star(x,y) = C_{10}(x,y)\dot{\psi}_0(x) + C_{01}(x,y)\dot{\psi}_0(y) - \dot{\psi}_0\left\{C(x,y)\right\}.$$

Acknowledgements

Partial funding in support of this work was provided by the Norwegian Research Council (grant number 154079/420) and by the Natural Sciences and Engineering Research Council of Canada.

References

Berg, D. (2007). Copula goodness-of-fit testing : An overview and power comparison. Working paper, University of Oslo & Norwegian computing center.

Billingsley, P. (1999). *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, New York.

Chen, X. & Fan, Y. (2005). Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. *Canad. J. Statist.* **33**, 389–414.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendance. *Acad. Royal Bel., Bull. Class. Sci.* **65**, 274–292.

Fermanian, J.-D. (2005). Goodness-of-fit tests for copulas. *J. Multivariate Anal.* **95**, 119–152.

Fermanian, J.-D., Radulović, D. & Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10**, 847–860.

Frees, E. W. & Valdez, E. A. (1998). Understanding relationships using copulas. *North Am. Act. J.* **2**, 1–25.

Gänssler, P. & Stute, W. (1987). *Seminar on empirical processes.* DMV Seminar **9**, Birkhäuser, Basel.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

Genest, C., Quessy, J.-F. & Rémillard, B. (2006a). Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scand. J. Statist.* **33**, 337–366.

Genest, C., Quessy, J.-F. & Rémillard, B. (2006b). On the joint asymptotic behavior of two rank-based estimators of the association parameter in the gamma frailty model. *Statist. Probab. Lett.* **76**, 10–18.

Genest, C., Quessy, J.-F. & Rémillard, B. (2006c). Local efficiency of a Cramér–von Mises test of independence. *J. Multivariate Anal.* **97**, 274–294.

Genest, C., Quessy, J.-F. & Rémillard, B. (2007). Asymptotic local efficiency of Cramér–von Mises tests for multivariate independence. *Ann. Statist.* **35**, 166–191.

Genest, C. & Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann. Henri Poincaré* **44**, in press.

Genest, C., Rémillard, B. & Beaudoin, D. (2008). Goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.* **42**, in press.

Ghoudi, K. & Rémillard, B. (1998). Empirical processes based on pseudo-observations. In *Asymptotic Methods in Probability and Statistics (Ottawa, Ontario, 1997).* North-Holland, Amsterdam, 171–197.

Hájek, J. & Sidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

Huard, D., Évin, G. & Favre, A.-C. (2006). Bayesian copula selection. *Comput. Statist. Data Anal.* **51**, 809–822.

Kim, G., Silvapulle, M. J. & Silvapulle, P. (2006). Comparison of semiparametric and parametric methods for estimating copulas. *Comput. Statist. Data Anal.* **51**, 2836–2850.

Quessy, J.-F. (2005). Théorie et application des copules : tests d'adéquation, tests d'indépendance et bornes pour la valeur-à-risque. *Doctoral dissertation, Université Laval.*

Scaillet, O. (2006). Kernel based goodness-of-fit tests for copulas with fixed smoothing parameters. *J. Multivariate Anal.* **98**, 533–543.

Schmid, F. & Schmidt, R. (2007). Multivariate extensions of Spearman's rho and related statistics. *Statist. Probab. Lett.* **77**, 407–416.

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

Shih, J. H. & Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51**, 1384–1399.

Shih, J. H. (1998). A goodness-of-fit test for association in a bivariate survival model. *Biometrika* **85**, 189–200.

Shorack, G. R. & Wellner, J. A. (1986). *Empirical processes with applications to statistics.* Wiley, New York.

Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist.* **33**, 357–375.

van der Vaart, A. W. & Wellner, J. A. (1996). *Weak convergence and empirical processes. With applications to statistics.* Springer Series in Statistics. Springer-Verlag, New York.

Corresponding author: Jean-François Quessy, Département de mathématiques et informatique, Université du Québec à Trois-Rivières, P.B. 500, Trois-Rivières, Canada, G9A 5H7. E-mail: jean-francois.quessy@uqtr.ca

Appendix A : Proofs

A.1. Proof of Proposition 1. Assumption (8) enables to deduce, from Lemma 3.10.11 of Van der Vaart & Wellner (1996), that the log-likelihood ratio of $\mathcal{Q}_{\delta_n}$ with respect to $\mathcal{Q}_0$ has the asymptotic representation

$$L_n = \frac{\delta}{\sqrt{n}} \sum_{i=1}^n \left\{ \frac{d(X_i, Y_i) - c_\theta(X_i, Y_i)}{c_\theta(X_i, Y_i)} \right\} - \frac{\delta^2}{2n} \sum_{i=1}^n \left\{ \frac{d(X_i, Y_i) - c_\theta(X_i, Y_i)}{c_\theta(X_i, Y_i)} \right\}^2 + o_{\mathrm{P}}(1).$$

The proofs for (i) and (ii) are achieved in separate steps.
(i) From the asymptotic representation (9), it follows that

$$\Theta_{n,\Lambda} = \Theta'_{n,\Lambda} + \frac{1}{n} \sum_{i=1}^n \Lambda_{C_\theta,10}(X_i, Y_i) \beta_{n1}(X_i) + \frac{1}{n} \sum_{i=1}^n \Lambda_{C_\theta,01}(X_i, Y_i) \beta_{n2}(y_i) + o_{\mathrm{P}}(1),$$

where $\beta_{n1}(x) = \sqrt{n}\{F_n(x) - x\}$ and $\beta_{n2}(y) = \sqrt{n}\{G_n(y) - y\}$. From Slutsky's lemma, the bivariate central limit theorem and arguments that one can find in Ghoudi & Rémillard (1998), the vector $(\Theta_{n,\Lambda}, L_n)$ converges to a bivariate normal distribution with mean vector and covariance matrix

$$\mu = \left(0, \frac{-\delta^2\sigma^2(L)}{2}\right) \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_\Lambda^2(C_\theta) & \delta\mu_\Lambda(C_\theta, D) \\ \delta\mu_\Lambda(C_\theta, D) & \delta^2\sigma_\Lambda^2(\mathcal{Q}_\delta) \end{pmatrix},$$

where $\sigma^2(L) = \mathrm{var}_{C_\theta}\{d(X,Y)/c_\theta(X,Y)\}$. One may then conclude, in view of Lecam's third lemma, that $\Theta_{n,\Lambda}$ is asymptotically normal with mean $\delta\mu_\Lambda(C_\theta, D)$ and variance $\sigma_\Lambda^2(C_\theta)$ under the contiguous sequence $(\mathcal{Q}_{\delta_n})_{n\geq 1}$.
(ii) From Hájek's projection method (Hájek & Sidák, 1967), one deduces the large-sample representation

$$\Theta_{n,\tau} = \frac{4}{\tau'_{C_\theta}(\theta)} \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{2C_\theta(X_i, Y_i) - X_i - Y_i + \frac{1 - \tau_{C_\theta}(\theta)}{2}\right\} + o_{\mathrm{P}}(1).$$

Hence, the vector (Θ_n, L_n) converges to a bivariate normal distribution with mean vector and covariance matrix

$$\mu = \left(0, \frac{-\delta^2\sigma^2(L)}{2}\right) \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_\tau^2 & \delta\mu_\tau(\theta) \\ \delta\mu_\tau(\theta) & \delta^2\sigma_\tau^2(\mathcal{Q}_\delta) \end{pmatrix},$$

from which it follows that $\Theta_{n,\tau}$ is asymptotically $\mathcal{N}(\delta\mu_\tau(\theta), \sigma_\tau^2)$ under $(\mathcal{Q}_{\delta_n})_{n\geq 1}$.

A.2. Proof of Proposition 2. Let $(X_1^{(n)}, Y_1^{(n)}), \ldots, (X_n^{(n)}, Y_n^{(n)})$ be a random sample from $\mathcal{Q}_{\delta_n}$. Write $\mathcal{C}_n^{(n)} = \mathcal{C}_{n,\theta}^{(n)} - \mathcal{B}_n^{(n)}$, where $\mathcal{C}_{n,\theta}^{(n)} = \sqrt{n}(C_n^{(n)} - C_\theta)$ and $\mathcal{B}_n^{(n)} = \sqrt{n}(C_{\hat{\theta}_n^{(n)}} - C_\theta)$. Here, $\hat{\theta}_n^{(n)}$ is the estimator based on the sample from $\mathcal{Q}_{\delta_n}$ and

$$C_n^{(n)}(x,y) = H_n^{(n)}\left\{\left(F_n^{(n)}\right)^{-1}(x), \left(G_n^{(n)}\right)^{-1}(y)\right\},$$

where

$$H_n^{(n)}(s,t) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\left(X_i^{(n)} \leq s, Y_i^{(n)} \leq t\right),$$

$F_n^{(n)}(s) = H_n^{(n)}(s,1)$ and $G_n^{(n)}(t) = H_n^{(n)}(1,t)$. From van der Vaart & Wellner (1996), condition (8) implies that $\mathbb{H}_n^{(n)}(s,t) = \sqrt{n}(H_n^{(n)} - C_\theta) \rightsquigarrow \mathbb{H} + \delta(D - C_\theta)$. In particular,

$$\beta_{1,n}^{(n)}(x) = \sqrt{n}\{F_n^{(n)}(x) - x\} = \mathbb{H}_n^{(n)}(x,1) \rightsquigarrow \mathbb{H}(x,1)$$

and

$$\beta_{2,n}^{(n)}(y) = \sqrt{n}\{G_n^{(n)}(y) - y\} = \mathbb{H}_n^{(n)}(1,y) \rightsquigarrow \mathbb{H}(1,y)$$

since $D(x,1) - C_\theta(x,1) = D(1,y) - C_\theta(1,y) = 0$. From Chapter 3 in Shorack & Wellner (1986), both

$$\sup_{0\leq x\leq 1}\left|F_n^{(n)}(x) - x\right| = \sup_{0\leq x\leq 1}\left|\left(F_n^{(n)}\right)^{-1}(x) - x\right|$$

and

$$\sup_{0\le y\le 1}\left|G_n^{(n)}(y)-y\right| = \sup_{0\le y\le 1}\left|\left(G_n^{(n)}\right)^{-1}(y)-y\right|$$

converge in probability to zero, so that

$$\sqrt{n}\left\{\left(F_n^{(n)}\right)^{-1}-I\right\} \rightsquigarrow -I\!\!H(\cdot,1) \quad\text{and}\quad \sqrt{n}\left\{\left(G_n^{(n)}\right)^{-1}-I\right\} \rightsquigarrow -I\!\!H(1,\cdot).$$

Hence, since one can write

$$\begin{aligned}
\mathcal{C}_{n,\theta}^{(n)}(x,y) &= I\!\!H_n^{(n)}\left\{\left(F_n^{(n)}\right)^{-1}(x),\left(G_n^{(n)}\right)^{-1}(y)\right\} \\
&\quad + \sqrt{n}\left\{C_\theta\left(\left(F_n^{(n)}\right)^{-1}(x),\left(G_n^{(n)}\right)^{-1}(y)\right)-C_\theta(x,y)\right\} \\
&= I\!\!H_n^{(n)}\left\{\left(F_n^{(n)}\right)^{-1}(x),\left(G_n^{(n)}\right)^{-1}(y)\right\} + C_{\theta,10}(x,y)\sqrt{n}\{(F_n^{(n)})^{-1}(x)-x\} \\
&\quad + C_{\theta,01}(x,y)\sqrt{n}\{(G_n^{(n)})^{-1}(y)-y\} + o_{\mathrm{P}}(1),
\end{aligned}$$

one deduces that $\mathcal{C}_{n,\theta}^{(n)}$ converges weakly to $\mathcal{C}_\theta+\delta(D-C_\theta)$, where $\mathcal{C}_\theta = I\!\!H - C_{\theta,10}I\!\!H(\cdot,1) - C_{\theta,01}I\!\!H(1,\cdot)$ is the limit identified, e.g. by Gänssler & Stute (1987) and Tsukahara (2005) under the null hypothesis. The second part of Assumption $\mathcal{A}_2$ and the mean-value theorem enable to establish that $\mathcal{B}_n^{(n)}$ converges to $\tilde{\Theta}\dot{C}_\theta = \Theta\dot{C}_\theta + \mu(C_\theta,D)\dot{C}_\theta$, while the joint consistency of $(\mathcal{C}_{n,\theta}^{(n)},\mathcal{B}_n^{(n)})$ to $(\mathcal{C}_\theta+\delta(D-C_\theta),\Theta\dot{C}_\theta+\mu(C_\theta,D)\dot{C}_\theta)$ rises from Assumption $\mathcal{A}_1$.

Appendix B : Computation of the drift terms

In the case of Clayton, Frank and Gumbel–Barnett copulas, the value of Spearman's rho cannot be expressed explicitly in terms of the dependence parameter, and hence the population value of formula (5) must be estimated through numerical methods. Such is also the case for

$$\rho'_{C_\theta}(\theta) = 12\int_0^1\int_0^1 \dot{C}_\theta(x,y)\mathrm{d}x\mathrm{d}y, \quad \mathrm{E}_D\left\{C_\theta(X,Y)\right\} = \int_0^1\int_0^1 C_\theta(x,y)d(x,y)\mathrm{d}x\mathrm{d}y,$$

$$\beta_{C_\theta} = \int_0^1\int_0^1 \frac{\{\dot{c}_\theta(x,y)\}^2}{c_\theta(x,y)}\,\mathrm{d}x\mathrm{d}y \quad\text{and}\quad \mathrm{E}_D\left\{\ell'_{C_\theta}(X,Y)\right\} = \int_0^1\int_0^1 \frac{\dot{c}_\theta(x,y)}{c_\theta(x,y)}\,d(x,y)\mathrm{d}x\mathrm{d}y,$$

where $c_\theta(x,y)=\partial^2 C_\theta(x,y)/\partial x\partial y$, $\dot{c}_\theta(x,y)=\partial c_\theta(x,y)/\partial\theta$ and $\dot{C}_\theta(x,y)=\partial C_\theta(x,y)/\partial\theta$. Note that for Archimedean copulas, i.e. dependence models of the form $C_\theta(x,y) = \phi_\theta^{-1}\{\phi_\theta(x)+\phi_\theta(y)\}$, one can show that

$$\dot{C}_\theta(x,y) = \frac{\dot{\phi}_\theta(x)+\dot{\phi}_\theta(y)-\dot{\phi}_\theta\left\{C(x,y)\right\}}{\phi'_\theta\left\{C_\theta(x,y)\right\}},$$

where $\dot{\phi}_\theta(x)=\partial\phi_\theta(x)/\partial\theta$ and $\phi'_\theta(x)=\partial\phi_\theta(x)/\partial x$. The Clayton, Frank and Gumbel–Barnett copulas are member of this important class of models.

B.1. The Clayton family. The copulas in this class and their associated densities are

$$C_\theta^{\mathrm{CL}}(x,y) = \left(x^{-\theta} + y^{-\theta} - 1\right)^{-1/\theta} \text{ and } c_\theta^{\mathrm{CL}}(x,y) = (\theta+1)\,(xy)^{-\theta-1}\left(x^{-\theta} + y^{-\theta} - 1\right)^{-1/\theta-2}, \tag{13}$$

where $\theta > -1$. The associated value of Kendall's tau is $\tau_{C_\theta^{\mathrm{CL}}}(\theta) = \theta/(\theta+2)$, from which one deduces easily that $\mathrm{E}_{C_\theta}(C_\theta) = (\theta+1)/2$ and $\tau'_{C_\theta^{\mathrm{CL}}}(\theta) = 2/(\theta+2)^2$. Further,

$$\dot{C}_\theta(x,y) = \frac{C_\theta(x,y)}{\theta}\left\{\frac{x^{-\theta}\log x + y^{-\theta}\log y}{x^{-\theta} + y^{-\theta} - 1} - \log C_\theta(x,y)\right\}.$$

B.2. The Frank family. Frank's copula is given by

$$C_\theta^{\mathrm{F}}(x,y) = -\frac{1}{\theta}\ln\left\{1 - \frac{\left(1 - e^{-\theta x}\right)\left(1 - e^{-\theta y}\right)}{1 - e^{-\theta}}\right\}, \tag{14}$$

where $\theta \in \mathbb{R}\setminus\{0\}$. As reported in Frees & Valdez (1999), Spearman's rho and Kendall's tau in this family are expressed by

$$\rho_{C_\theta^{\mathrm{F}}}(\theta) = 1 + \frac{12}{\theta^2}\int_0^\theta \frac{t(2t-\theta)}{e^t - 1}\mathrm{d}t \quad \text{and} \quad \tau_{C_\theta^{\mathrm{F}}}(\theta) = 1 - \frac{4}{\theta} + \frac{4}{\theta^2}\int_0^\theta \frac{t}{e^t - 1}\mathrm{d}t.$$

Hence, one deduces

$$\rho'_{C_\theta^{\mathrm{F}}}(\theta) = \frac{12}{\theta\left(e^\theta - 1\right)} - \frac{24}{\theta^4}\int_0^\theta \frac{t(3t-\theta)}{e^t - 1}\mathrm{d}t$$

and

$$\tau'_{C_\theta^{\mathrm{F}}}(\theta) = \frac{4}{\theta^2} + \frac{4}{\theta\left(e^\theta - 1\right)} - \frac{8}{\theta^3}\int_0^\theta \frac{t}{e^t - 1}\mathrm{d}t.$$

The other necessary computations, however, must be accomplished numerically.

B.3. The Gumbel–Barnett family. The analytical form of this extreme-value copula (see Ghoudi *et al.*, 1998) is

$$C_\theta^{\mathrm{GB}}(x,y) = \exp\left\{-\left(|\log x|^{1/(1-\theta)} + |\log y|^{1/(1-\theta)}\right)^{1-\theta}\right\}, \tag{15}$$

where $0 \le \theta \le 1$. Computations of the drift terms in this class of models are difficult and must be done numerically. The only explicit expressions are for Kendall's tau and its derivative, namely $\tau_{C_\theta^{\mathrm{GB}}}(\theta) = \theta$ and $\tau'_{C_\theta^{\mathrm{GB}}}(\theta) = 1$.

B.4. The Normal family. The Normal copula, which arises as the dependence function associated to the classical normal model, is defined by

$$C_\theta^{\mathrm{N}}(x,y) = \int_{-\infty}^{\Phi^{-1}(x)}\int_{-\infty}^{\Phi^{-1}(y)} h_\theta(s,t)\mathrm{d}s\mathrm{d}t, \tag{16}$$

where

$$h_\theta(s,t) = \frac{\left(1-\theta^2\right)^{-1/2}}{2\pi} \exp\left\{ -\frac{1}{2(1-\theta^2)}\left(s^2+t^2-2\theta st\right)\right\}$$

is the standard bivariate normal density with correlation coefficient θ. Despite the implicit form of C_θ^{N} involving the percentile function of a standard univariate normal distribution, there exists explicit relationships between the dependence parameter θ and Kendall and Spearman measures of association. Explicitly,

$$\tau_{C_\theta}(\theta) = \frac{2}{\pi}\sin^{-1}(\theta) \quad \text{and} \quad \rho_{C_\theta}(\theta) = \frac{6}{\pi}\sin^{-1}\left(\frac{\theta}{2}\right),$$

from which it follows easily that

$$\mathrm{E}_{C_\theta}(C_\theta) = \frac{2\sin^{-1}(\theta)+\pi}{4\pi}, \quad \tau'_{C_\theta}(\theta) = \frac{2}{\pi\sqrt{1-\theta^2}} \quad \text{and} \quad \rho'_{C_\theta}(\theta) = \frac{6}{\pi\sqrt{4-\theta^2}}.$$

Hence, if $D \equiv C_{\theta_D}^{\mathrm{N}}$, i.e. if one considers a mixture of Normal copulas, then

$$\mu_\rho(C_\theta, D) = \frac{\sin^{-1}(\theta_D/2) - \sin^{-1}(\theta/2)}{\sqrt{4-\theta^2}}.$$

Also, the density associated to C_θ^{N} is

$$c_\theta^{\mathrm{N}}(x,y) = h_\theta\left\{\Phi^{-1}(x), \Phi^{-1}(y)\right\}\left(\Phi^{-1}\right)'(x)\left(\Phi^{-1}\right)'(y),$$

and it is possible to establish that

$$\ell'_{C_\theta^{\mathrm{N}}}(x,y) = \frac{\dot{c}_\theta^{\mathrm{N}}(x,y)}{c_\theta^{\mathrm{N}}(x,y)} = \left.\frac{\theta(1-\theta^2)-\theta(s^2+t^2)+(\theta^2+1)st}{(1-\theta^2)^2}\right|_{s=\Phi^{-1}(x),\, t=\Phi^{-1}(y)}.$$

This enables to compute

$$\begin{aligned}
\mathrm{E}_D\left\{\ell'_{C_\theta^{\mathrm{N}}}(X,Y)\right\} &= \int_0^1\int_0^1 \ell'_{C_\theta^{\mathrm{N}}}(x,y) c_{\theta_D}^{\mathrm{N}}(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{\frac{\theta(1-\theta^2)-\theta(s^2+t^2)+(\theta^2+1)st}{(1-\theta^2)^2}\right\} h_{\theta_D}(s,t)\mathrm{d}s\mathrm{d}t \\
&= \mathrm{E}_{\theta_D}\left\{\frac{\theta(1-\theta^2)-\theta(S^2+T^2)+(\theta^2+1)ST}{(1-\theta^2)^2}\right\},
\end{aligned}$$

where (S,T) follows a bivariate normal distribution with means 0, variances 1 and correlation coefficient θ_D. Here, E_{θ_D} denotes expectation with respect to h_{θ_D}. Thus, noting that $\mathrm{E}_{\theta_D}(S^2) = \mathrm{E}_{\theta_D}(T^2) = 1$ and $\mathrm{E}_{\theta_D}(ST) = \theta_D$, straightforward computations yield

$$\mathrm{E}_D\left\{\ell'_{C_\theta^{\mathrm{N}}}(X,Y)\right\} = \frac{(\theta^2+1)(\theta_D-\theta)}{(1-\theta^2)^2}.$$

Long but similar computations enable to obtain $\beta_{C_\theta} = \theta^2+1$ and hence

$$\mu_{PL}(C_\theta, D) = \frac{\theta_D - \theta}{(1-\theta^2)^2}.$$