Kybernetika

Salim Bouzebda; Amor Keziou

New estimates and tests of independence in semiparametric copula models

Kybernetika, Vol. 46 (2010), No. 1, 178--201

Persistent URL: http://dml.cz/dmlcz/140046

KYBERNETIKA — VOLUME *46* (2010), NUMBER 1, PAGES 178–201

NEW ESTIMATES AND TESTS OF INDEPENDENCE IN SEMIPARAMETRIC COPULA MODELS

Salim Bouzebda and Amor Keziou

We introduce new estimates and tests of independence in copula models with unknown margins using ϕ-divergences and the duality technique. The asymptotic laws of the estimates and the test statistics are established both when the parameter is an interior or a boundary value of the parameter space. Simulation results show that the choice of χ^2-divergence has good properties in terms of efficiency-robustness.

Keywords: dependence function, multivariate rank statistics, semiparametric inference, copulas, boundary, divergences, duality

Classification: 62F03, 62F10, 62F12, 62H12, 62H15

1. INTRODUCTION AND MOTIVATIONS

Copulas are a useful tool to model dependent data as they allow to separate the dependence properties of the data from their marginal properties and to construct multivariate models with marginal distributions of arbitrary form. In particular, parametric models for copulas with unknown margins have been intensively investigated during the last decades. In the monographs by [29] and [19] the reader finds detailed accounts of the theory as well as surveys of commonly used copulas.

It is known that some commonly used dependence measures such as Pearson's correlation coefficient, Kendall's tau and Spearman's rho cannot completely capture the dependence structure among variables. Copulas have become popular in applied statistics, because of the fact that they constitute a flexible and robust way to model dependence between the margins of random vectors.

In this framework, semiparametric inference methods, based on *pseudo-likelihood*, have been applied to copulas by a number of authors (see, e. g., [38, 42, 43] and the references therein). Throughout the available literature, investigations on the asymptotic properties of parametric estimators, as well as the relevant test statistics, have privileged the case where the parameter is an interior point of the admissible domain. However, for most parametric copula models of interest, the boundaries of the admissible parameter spaces include some important parameter values, typically among which, that corresponding to the independence of margins. We find in [19] many examples of parametric copulas, for which marginal independence is verified for some specific values of the parameter θ, on the boundary $\partial\boldsymbol{\Theta}$ of the admissible parameter set $\boldsymbol{\Theta} \subseteq \mathbb{R}^p$, $p \geq 1$.

This paper concentrates on this specific problem. We aim, namely, to investigate parametric inference procedures, in the case where the parameter belongs to the boundary of the admissible domain. In particular, it will become clear, that the usual limit laws both for parametric copula estimators and test statistics become invalid under these limiting cases, and, in particular, under marginal independence. Motivated by this observation, we will introduce a new semiparametric inference procedure based on *ϕ-divergences* and the *duality* technique extending the paper by [4] to the general context of ϕ-divergences for multivariate copulas with multivariate parameter. The proposed method extends the pseudo-maximum likelihood procedure introduced by [15]. It will be seen that the last method corresponds to the particular choice of the KL_m-divergence. We obtain a class of estimates and test statistics depending upon the divergence. We are interested by comparing the proposed estimates (including the pseudo-maximum likelihood one) in terms of efficiency and robustness according to the choice of the divergence. We will show that the proposed estimators, under suitable conditions, remain asymptotically normal, even under the marginal independence assumption for appropriate choice of the divergence. This will allow us to introduce test statistics of independence, whose study will be made, both under the null and the alternative hypotheses. Let

$$F(x_1, \ldots, x_d) := \mathrm{P}\{X_1 \leq x_1, \ldots, X_d \leq x_d\}$$

be a d-dimensional distribution function, and $F_i(x_i) := \mathrm{P}(X_i \leq x_i),\ i = 1, \ldots, d$, the marginal distributions of $F(\cdot)$. It is well known since the work of [40] that there exists a distribution function $\mathbf{C}(\cdot)$ on $[0,1]^d$ with uniform marginals such that

$$\mathbf{C}(\mathbf{u}) := \mathbf{C}(u_1, \ldots, u_d) := \mathrm{P}\left\{F_1(X_1) \leq u_1, \ldots, F_d(X_d) \leq u_d\right\}. \tag{1.1}$$

See also [7, 8, 9, 10, 28, 33] and [36]. We can refer to [39], where the author sketches the proof of (1.1), develops some of its consequences, and surveys some of the work on copulas. Formally, copulas can be defined in the common way as follows.

Definition 1.1. An d-dimensional copula is a function $\mathbf{C} : [0,1]^d \to [0,1]$ with the following properties

1. $\mathbf{C}(\cdot)$ is grounded, i. e., for every $\mathbf{u} = (u_1, \ldots, u_d)$, $\mathbf{C}(\mathbf{u}) = 0$ if at least one coordinate $u_i = 0,\ i = 1, \ldots, d$;

2. $\mathbf{C}(\cdot)$, is d-increasing, i. e., for every $\mathbf{u} \in [0,1]^d$ and $\mathbf{v} \in [0,1]^d$ such that $\mathbf{u} \leq \mathbf{v}$, the $\mathbf{C}$-volume $V_{\mathbf{C}}[\mathbf{u}, \mathbf{v}]$ of the box $[\mathbf{u}, \mathbf{v}]$ is non negative (see [29]);

3. $\mathbf{C}(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for all $u_i \in [0,1]^d$, $u_i = 0,\ i = 1, \ldots, d$.

Many useful multivariate models for dependence between $X_1, \ldots, X_d$ turn out to be generated by *parametric* families of copulas of the form $\{\mathbf{C}_\theta;\ \theta \in \boldsymbol{\Theta}\}$, typically indexed by a vector valued parameter $\theta \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$ (see, e. g., [22, 23, 29], and [18] among others). In the sequel, we assume that $\mathbf{C}_\theta(\cdot)$ admits a density $\mathbf{c}_\theta(\cdot)$ with respect to the Lebesgue measure λ on $\mathbb{R}^d$, i. e., $\mathbf{c}_\theta(\cdot) = \frac{\partial^d}{\partial u_1 \ldots \partial u_d}\mathbf{C}_\theta(\cdot)$. The nonparametric approach to copula estimation has been initiated by [7], who introduced and investigated the *empirical copula process*. In addition, [8, 9, 10] described the limiting behavior of this empirical process see, also [12] and the references therein. The empirical copula process has been studied in full generality in [13] and [41].

In the present paper, we consider the estimation and test problems for semi-parametric copula models with unknown general margins. Let $(X_{1k},\ldots,X_{dk})^\top$, for $k=1,\ldots,n$, be a d-variate sample with distribution function $F_{\theta_{\mathbf{T}},F_1,\ldots,F_d}(\cdot,\ldots,\cdot) = \mathbf{C}_{\theta_{\mathbf{T}}}(F_1(\cdot),\ldots,F_d(\cdot))$ where $\theta_{\mathbf{T}} \in \boldsymbol{\Theta}$ is used to denote the true unknown value of the parameter. In order to estimate $\theta_{\mathbf{T}}$, some semiparametric estimation procedures, based on the maximization, on the parameter space $\boldsymbol{\Theta}$, of properly chosen *pseudo-likelihood* criterion, have been proposed by [15, 24, 30, 38, 43] and [42] among others. In each of these papers, some asymptotic normality properties are established for

$$\sqrt{n}(\tilde{\theta} - \theta_{\mathbf{T}})$$

where $\tilde{\theta} = \tilde{\theta}_n$ denotes a properly chosen estimator of $\theta_{\mathbf{T}}$. This is achieved, provided that $\theta_{\mathbf{T}}$ lies in the *interior*, denoted by $\mathring{\boldsymbol{\Theta}}$, of the parameter space $\boldsymbol{\Theta} \subseteq \mathbb{R}^p$. On the other hand, the case where $\theta_{\mathbf{T}} \in \partial\boldsymbol{\Theta} := \overline{\boldsymbol{\Theta}} - \mathring{\boldsymbol{\Theta}}$ is a *boundary value* of $\boldsymbol{\Theta}$, has not been studied in a systematical way until present. Moreover, it turns out that, for the above-mentioned estimators, the asymptotic normality of $\sqrt{n}(\tilde{\theta} - \theta_{\mathbf{T}})$, may fail to hold for $\theta_{\mathbf{T}} \in \partial\boldsymbol{\Theta}$; indeed, under some regularity conditions, when θ is univariate, we can prove that the limit law is the distribution of $Z\mathbb{1}_{(Z\geq 0)}$ where Z is a centered normal variable, and that the limit law of the generalized pseudo-likelihood ratio statistic is a mixture of chi-square laws with one degree of freedom and Dirac measure at zero; see [2]. Furthermore, when the parameter is multivariate, the derivation of the limit distributions under the null hypothesis of independence, becomes much more complex; see [37]. Also, the limit distributions are not standard which yields formidable numerical difficulties to calculate the critical value of the test.

We cite below some examples of parametric copulas, for which marginal independence is verified for some specific values of the parameter θ, on the boundary $\partial\boldsymbol{\Theta}$ of the admissible parameter set $\boldsymbol{\Theta}$. We start with examples for which θ varies within subsets of $\mathbb{R}$. Such is the case for the extreme value copulas, namely

$$\mathbf{C}_A(u_1,u_2) := \exp\left\{\log u_1u_2 A\left(\frac{\log u_1}{\log u_1u_2}\right)\right\}, \tag{1.2}$$

where $A(\cdot)$ is a convex function on $[0,1]$, satisfying

$$A : [0,1] \mapsto [1/2,1] \text{ such that } \max(t,1-t) \leq A(t) \leq 1 \text{ for all } 0 \leq t \leq 1.$$

For

$$A(t) := A_\theta(t) = (t^\theta + (1-t)^\theta)^{1/\theta}; \quad \theta \in [1,\infty[\tag{1.3}$$

we have [16] family of copulas, which is one of the most popular model used to model bivariate extreme values. For

$$A_\theta(t) = 1 - (t^{-\theta} + (1-t)^{-\theta})^{-1/\theta}; \quad \theta \in [0,\infty[\tag{1.4}$$

we obtain [14] family of copulas. Finally for

$$A_\theta(t) = t\Phi\left(\theta^{-1} + \frac{1}{2}\theta\log\left(\frac{t}{1-t}\right)\right) + (1-t)\Phi\left(\theta^{-1} - \frac{1}{2}\theta\log\left(\frac{t}{1-t}\right)\right), \quad (1.5)$$

where $\theta \in [0, \infty[$ and $\Phi(\cdot)$ denoting the standard normal $N(0,1)$ distribution function, we obtain the [17] family of copulas. A useful family of copulas, due to [18], is given, for $0 < u_1, u_2 < 1$, by

$$\mathbf{C}_\theta(u_1, u_2) := 1 - \left[(1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta\right]^{1/\theta}; \quad \theta \in [1, \infty[. \quad (1.6)$$

The Gumbel–Barnett copulas are given, for $0 < u_1, u_2 < 1$, by

$$\mathbf{C}_\theta(u_1, u_2) := u_1 u_2 \exp\{-(1-\theta)(\log u_1)(\log u_2)\}; \quad \theta \in [0, 1]. \quad (1.7)$$

The Clayton copulas of positive dependence are such that, for $0 < u_1, u_2 < 1$,

$$\mathbf{C}_\theta(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}; \quad \theta \in]0, \infty[. \quad (1.8)$$

Parametric families of copulas with parameter θ varying in $\mathbb{R}^p$, for some $p \geq 2$, include the following classical examples. Below, we set $\theta = \left(\theta_1, \theta_2\right)^\top \in \mathbb{R}^2$.

$$\mathbf{C}_\theta(u_1, u_2) := \left\{1 + \left[(u_1^{-\theta_1} - 1)^{\theta_2} + (u_2^{-\theta_1} - 1)^{\theta_2}\right]^{1/\theta_2}\right\}^{-1/\theta_1}, \ \theta \in]0, \infty[\times[1, \infty[; \quad (1.9)$$

$$\mathbf{C}_\theta(u_1, u_2) := \exp\Big\{ - \Big[\theta_2^{-1}\log\Big(\exp\left(-\theta_2(\log u_1)^{\theta_1}\right) \quad (1.10)$$
$$+ \exp\left(-\theta_2(\log u_2)^{\theta_1}\right) - 1\Big)\Big]^{1/\theta_1}\Big\}, \ \theta \in [1, \infty[\times]0, \infty[.$$

For other examples of the kind, we refer to [19].

For each of the above examples, the independence case $\mathbf{C}_{\theta_\mathbf{T}}(u_1, u_2) = u_1 u_2$ (or $A(t) = 1$) occurs at the boundary of the parameter space $\boldsymbol{\Theta}$, i. e., when $\theta_\mathbf{T} = 1$ for the models (1.3), (1.6) and (1.7), $\theta_\mathbf{T} = 0$ for the models (1.4), (1.5) and (1.8), $\theta_\mathbf{T} = (0, 1)^\top$ for the bivariate parameter model (1.9), and $\theta_\mathbf{T} = (1, 0)^\top$ for the bivariate parameter model (1.10). In the sequel, we will denote by θ_0 the value of the parameter (when it exists), corresponding to the independence of the marginals, i. e., the value of the parameter for which we have

$$\mathbf{C}_{\theta_0}(\mathbf{u}) := \prod_{i=1}^d u_i, \quad \text{for all } \mathbf{u} \in (0,1)^d.$$

Hence, $\theta_0 = 1$ for the models (1.3), (1.6) and (1.7), $\theta_0 = 0$ for the models (1.4), (1.5) and (1.8), $\theta_0 = (0, 1)^\top$ for the model (1.9), and $\theta_0 = (1, 0)^\top$ for the model (1.10). Note that for the models (1.4), (1.5), (1.8), (1.9) and (1.10), $\mathbf{C}_{\theta_0}(u_1, u_2) = u_1 u_2$ is naturally defined to be the limit of $\mathbf{C}_\theta(\cdot)$ when θ tends to θ_0 with values in $\boldsymbol{\Theta}$.

Recall that $\mathbf{c}_\theta(\cdot) := \frac{\partial^d}{\partial u_1 \ldots \partial u_d} \mathbf{C}_\theta(\cdot)$ is the density of $\mathbf{C}_\theta(\cdot)$ and we define $\mathbf{c}_{\theta_0}(\cdot)$ to be the limit of $\mathbf{c}_\theta(\cdot)$ when θ tends to θ_0 with values in $\boldsymbol{\Theta}$. Hence, we can show that for all the above models $\mathbf{c}_{\theta_0}(u_1, u_2) = 1$ for all $0 < u_1, u_2 < 1$.

In contrast with the preceding examples, where $\theta_0 \in \partial\boldsymbol{\Theta}$ is a boundary value of $\boldsymbol{\Theta}$, the case where θ_0 is an interior point of $\boldsymbol{\Theta}$ may, at times, occur, but is more seldom. An example where $\theta_0 \in \mathring{\boldsymbol{\Theta}}$ is given by the Farlie–Gumbel–Morgenstern (FGM) copula, defined by

$$\mathbf{C}_\theta(u_1, u_2) := u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2), \ \theta \in \boldsymbol{\Theta} := [-1, 1], \tag{1.11}$$

and for which $\theta_0 = 0 \in \mathring{\boldsymbol{\Theta}} =]-1, 1[$.

In the present article, we will treat parametric estimation of $\theta_{\mathbf{T}}$, and tests of the independence assumption $\theta_{\mathbf{T}} = \theta_0$. We consider both the case where $\theta_0 \in \mathring{\boldsymbol{\Theta}}$ is an interior point of $\boldsymbol{\Theta}$, and the case where $\theta_0 \in \partial\boldsymbol{\Theta}$ is a boundary value of $\boldsymbol{\Theta}$. Our approach is novel in this setting and it will become clear later on from our results, that the asymptotic normality of the estimate based on ϕ-divergences holds, even under the independence assumption, when, either, θ_0 is an interior, or a boundary point of $\boldsymbol{\Theta}$, independently of the dimension of the parameter space. The proposed test statistics of independence using ϕ-divergences are also studied, under the null hypothesis $\mathscr{H}_0$ of independence, as well as under the alternative hypothesis. The asymptotic distributions of the test statistics under the alternative hypothesis are used to derive an approximation to the power functions. An application of the forthcoming results will allow us to evaluate the sample size necessary to guarantee a pre-assigned power level, with respect to a specified alternative. To establish our results, we use similar arguments as those developed by [42] in connection with the instrumental statements on rank statistics established by [35] and [32] among others, combined with a new technique, (based on the law of iterated logarithm given in Lemma A.1 below) to show both existence and consistency of our estimates and test statistics. In Section 5, we investigate the finite-sample performance of the newly proposed estimators. To avoid interrupting the flow of the presentation, all mathematical developments are relegated to the appendix.

2. A NEW INFERENCE PROCEDURE

Recall that the ϕ-divergence between a bounded signed measure $\mathbf{Q}$, and a probability P on $\mathscr{D}$, when $\mathbf{Q}$ is absolutely continuous with respect to $\mathbf{P}$, is defined by

$$D_\phi(\mathbf{Q}, \mathbf{P}) := \int_{\mathscr{D}} \phi\left(\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}(\mathbf{x})\right) \mathrm{d}\mathbf{P}(x),$$

where ϕ is a convex function from $]-\infty, \infty[$ to $[0, \infty]$ with $\phi(1) = 0$. We will consider only ϕ-divergences for which the function ϕ is strictly convex and satisfies: the domain of ϕ, $\operatorname{dom}\phi := \{x \in \mathbb{R} : \phi(x) < \infty\}$ is an interval with end points $a_\phi < 1 < b_\phi$, $\phi(a_\phi) = \lim_{x \downarrow a_\phi} \phi(x)$ and $\phi(a_\phi) = \lim_{x \uparrow b_\phi} \phi(x)$. The Kullback–Leibler, modified Kullback–Leibler, χ^2, modified χ^2 and Hellinger divergences are examples of ϕ-divergences; they are obtained respectively for $\phi(x) = x \log x - x + 1$,

$\phi(x) = -\log x + x - 1$, $\phi(x) = \frac{1}{2}(x-1)^2$, $\phi(x) = \frac{1}{2}\frac{(x-1)^2}{x}$ and $\phi(x) = 2(\sqrt{x}-1)^2$. We extend the definition of these divergences on the whole space of all bounded signed measures via the extension of the definition of the corresponding ϕ functions on the whole real space $\mathbb{R}$ as follows: when ϕ is not well defined on $\mathbb{R}_-$ or well defined but not convex on $\mathbb{R}$, we set $\phi(x) = +\infty$ for all $x < 0$. Observe for the χ^2-divergence, the corresponding ϕ function is defined on whole $\mathbb{R}$ and strictly convex. We refer to [25] for a systematic theory of divergences. We denote by ϕ^* the Fenchel–Legendre transform of the convex function ϕ, i. e., the function defined by

$$t \in \mathbb{R} \mapsto \phi^*(t) := \sup_{x \in \mathbb{R}} \{tx - \phi(x)\}.$$

From [31], Section 26, we can prove that it is strictly convex, its domain is an interval (a^*_ϕ, b^*_ϕ) with

$$a^*_\phi < 0 < b^*_\phi, \quad a^*_\phi = \lim_{x \to -\infty} \frac{\phi(x)}{x}, \quad b^*_\phi = \lim_{x \to +\infty} \frac{\phi(x)}{x},$$

and it satisfies $\phi^*(0) = 0$,

$$\phi^*(a^*_\phi) = \lim_{t \downarrow a^*_\phi} \phi^*(t) \quad \text{and} \quad \phi^*(b^*_\phi) = \lim_{t \uparrow b^*_\phi} \phi^*(t).$$

Furthermore, it holds that ϕ is the Fenchel–Legendre transform of ϕ^*. In the sequel, for all θ, we denote by $D_\phi(\theta, \theta_\mathbf{T})$ the ϕ-divergences between $\mathbf{C}_\theta(\cdot)$ and $\mathbf{C}_{\theta_\mathbf{T}}(\cdot)$, i. e.,

$$D_\phi(\theta, \theta_\mathbf{T}) := \int_\mathbf{I} \phi\left(\frac{\mathrm{d}\mathbf{C}_\theta}{\mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}}(\mathbf{u})\right) \mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}(\mathbf{u}) = \int_\mathbf{I} \phi\left(\frac{\mathbf{c}_\theta(\mathbf{u})}{\mathbf{c}_{\theta_\mathbf{T}}(\mathbf{u})}\right) \mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}(\mathbf{u}), \tag{2.1}$$

where $\mathbf{I} = (0,1)^d$. Denote $\mathbf{C}_n(\cdot)$ the empirical copula associated to the data, i. e.,

$$\mathbf{C}_n(\mathbf{u}) := \frac{1}{n} \sum_{k=1}^{n} \prod_{i=1}^{d} \mathbb{1}_{\{F_{in}(X_{ik}) \le u_1\}}, \ \mathbf{u} \in \mathbf{I}, \tag{2.2}$$

and

$$F_{in}(t) := \left\{\frac{n}{n+1}\right\} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{]-\infty,t]}(X_{ik}) = \frac{1}{n+1} \sum_{k=1}^{n} \mathbb{1}_{]-\infty,t]}(X_{ik}), \ i = 1, \dots, d,$$

where $\mathbb{1}_A$ stands for the indicator function of the event A. The rescaling by the factor $n/(n+1)$, avoids difficulties arising from potential unboundedness of $\mathbf{c}_\theta(\mathbf{u})$ when one of u_i's tends to 1. Observe that the plug-in estimate

$$\int_\mathbf{I} \phi\left(\frac{\mathrm{d}\mathbf{C}_\theta}{\mathrm{d}\mathbf{C}_n}(\mathbf{u})\right) \mathrm{d}\mathbf{C}_n(\mathbf{u})$$

of $D_\phi(\theta, \theta_\mathbf{T})$ is not well defined since $\mathbf{C}_\theta(\cdot)$ is not absolutely continuous with respect to $\mathbf{C}_n(\cdot)$. In order to avoid this difficulty, and to estimate the divergences $D_\phi(\theta, \theta_\mathbf{T})$ for a given $\theta \in \mathbf{\Theta}$ in particular for $\theta = \theta_0$, we will make use of the dual representation of ϕ-divergences obtained by [1] Theorem 4.4 and [20] Theorem 2.3. By this, when ϕ is differentiable, we readily obtain that $D_\phi(\theta_0, \theta_\mathbf{T})$ can be rewritten into

$$D_\phi(\theta_0, \theta_\mathbf{T}) := \sup_{f \in \mathscr{F}} \left\{\int_\mathbf{I} f \, \mathrm{d}\mathbf{C}_{\theta_0} - \int_\mathbf{I} \phi^*(f) \, \mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}\right\}, \tag{2.3}$$

where $\mathscr{F}$ is an arbitrary class of measurable functions fulfilling the following two conditions:

$$\forall f \in \mathscr{F}, \quad \int |f| \, \mathrm{d}\mathbf{C}_{\theta_0} < \infty$$

and

$$\phi'(\mathrm{d}\mathbf{C}_{\theta_0}/\mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}) = \phi'(\mathbf{c}_{\theta_0}/\mathbf{c}_{\theta_\mathbf{T}}) \in \mathscr{F}.$$

Furthermore, the sup in the above display is unique and is achieved at $f = \phi'(\mathbf{c}_{\theta_0}/\mathbf{c}_{\theta_\mathbf{T}})$. Note that for the specific value θ_0, corresponding to the independence, we have $\mathbf{c}_{\theta_0}(\mathbf{u}) = 1, \ \forall \mathbf{u} \in \mathbf{I}$. So, by the above statement, taking the class of functions

$$\mathscr{F} = \left\{\mathbf{u} \in \mathbf{I} \mapsto \phi'\left(1/\mathbf{c}_\theta(\mathbf{u})\right); \ \theta \in \mathbf{\Theta}\right\},$$

we obtain the formula

$$\begin{aligned} & D_\phi(\theta_0, \theta_\mathbf{T}) \\ &= \sup_{\theta \in \mathbf{\Theta}} \left\{ \int_\mathbf{I} \phi'\left(\frac{\mathbf{c}_{\theta_0}}{\mathbf{c}_\theta}\right) \mathrm{d}\mathbf{C}_{\theta_0}(\mathbf{u}) - \int_\mathbf{I} \left[\frac{\mathbf{c}_{\theta_0}}{\mathbf{c}_\theta} \phi'\left(\frac{\mathbf{c}_{\theta_0}}{\mathbf{c}_\theta}\right) - \phi\left(\frac{\mathbf{c}_{\theta_0}}{\mathbf{c}_\theta}\right)\right] \mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}(\mathbf{u}) \right\} \\ &= \sup_{\theta \in \mathbf{\Theta}} \left\{ \int_\mathbf{I} \phi'\left(\frac{1}{\mathbf{c}_\theta}\right) \mathrm{d}u_1 \dots \mathrm{d}u_d - \int_\mathbf{I} \left[\frac{1}{\mathbf{c}_\theta} \phi'\left(\frac{1}{\mathbf{c}_\theta}\right) - \phi\left(\frac{1}{\mathbf{c}_\theta}\right)\right] \mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}(\mathbf{u}) \right\}, \quad (2.4) \end{aligned}$$

whenever

$$\int_\mathbf{I} \left|\phi'\left(1/\mathbf{c}_\theta\right)\right| \mathrm{d}u_1 \dots \mathrm{d}u_d < \infty \quad \text{for all} \quad \theta \in \mathbf{\Theta}.$$

Furthermore, the sup is unique and reached at $\theta = \theta_\mathbf{T}$. Hence, the divergence $D_\phi(\theta_0, \theta_\mathbf{T})$ and the parameter $\theta_\mathbf{T}$ can be estimated respectively by

$$\sup_{\theta \in \mathbf{\Theta}} \left\{ \int_\mathbf{I} \phi'\left(\frac{1}{\mathbf{c}_\theta}\right) \mathrm{d}u_1 \dots du_d - \int_\mathbf{I} \left[\frac{1}{\mathbf{c}_\theta} \phi'\left(\frac{1}{\mathbf{c}_\theta}\right) - \phi\left(\frac{1}{\mathbf{c}_\theta}\right)\right] \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\} \qquad (2.5)$$

and

$$\arg\sup_{\theta \in \mathbf{\Theta}} \left\{ \int_\mathbf{I} \phi'\left(\frac{1}{\mathbf{c}_\theta}\right) \mathrm{d}u_1 \dots \mathrm{d}u_d - \int_\mathbf{I} \left[\frac{1}{\mathbf{c}_\theta} \phi'\left(\frac{1}{\mathbf{c}_\theta}\right) - \phi\left(\frac{1}{\mathbf{c}_\theta}\right)\right] \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\}, \qquad (2.6)$$

in which $\mathbf{C}_{\theta_\mathbf{T}}(\cdot)$ is replaced by $\mathbf{C}_n(\cdot)$. Note that this class of estimates contains the maximum pseudo-likelihood (MPL) estimator proposed by [30]; it is obtained for the KL_m-divergence taking $\phi(x) = -\log(x) + x - 1$. Under some regularity conditions, we can prove that these estimates are consistent and asymptotically normal in the same way as the MPL estimate when the parameter $\theta_\mathbf{T}$ is an interior point of the parameter space $\mathbf{\Theta}$. The interest of divergence remains in the fact that a properly choice of the divergence may ameliorate the MPL estimator in terms of efficiency-robustness. The results in [2] show that, for $\mathbf{\Theta} = [\theta_0, \infty)$, and when the true value $\theta_\mathbf{T}$ of the parameter is equal to θ_0 (corresponding to the independence assumption), the classical asymptotic normality property of the MPL estimate is no more satisfied. To overcome this difficulty, in what follows, we enlarge the parameter space $\mathbf{\Theta}$ into a wider space $\mathbf{\Theta}_e \supset \mathbf{\Theta}$. This is tailored to let θ_0 become an interior point of $\mathbf{\Theta}_e$. More precisely, set

$$\mathbf{\Theta}_e := \left\{ \theta \in \mathbb{R}^p \text{ such that } \int_\mathbf{I} \left|\phi'(1/\mathbf{c}_\theta(\mathbf{u}))\right| \mathrm{d}u_1 \dots \mathrm{d}u_d < \infty \right\}. \qquad (2.7)$$

So, applying (2.3), with the class of functions

$$\mathscr{F} := \{\mathbf{u} \in \mathbf{I} \mapsto \phi'(1/\mathbf{c}_\theta(\mathbf{u}));\ \theta \in \mathbf{\Theta}_e\},$$

we obtain

$$D_\phi(\theta_0, \theta_\mathbf{T}) = \sup_{\theta \in \mathbf{\Theta}_e} \left\{ \int_\mathbf{I} \phi'\left(\frac{1}{\mathbf{c}_\theta}\right) \mathrm{d}\mathbf{u} - \int_\mathbf{I} \left[\frac{1}{\mathbf{c}_\theta}\phi'\left(\frac{1}{\mathbf{c}_\theta}\right) - \phi\left(\frac{1}{\mathbf{c}_\theta}\right)\right] \mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}(\mathbf{u}) \right\}. \quad (2.8)$$

Furthermore, the sup in this display is unique and reached in $\theta = \theta_\mathbf{T}$. Hence, we propose to estimate $D_\phi(\theta_0, \theta_\mathbf{T})$ by

$$\widehat{D_\phi}(\theta_0, \theta_\mathbf{T}) := \sup_{\theta \in \mathbf{\Theta}_e} \int_\mathbf{I} \mathbf{m}(\theta, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}), \quad (2.9)$$

and to estimate the parameter $\theta_\mathbf{T}$ by

$$\widehat{\theta}_n := \arg \sup_{\theta \in \mathbf{\Theta}_e} \left\{ \int_\mathbf{I} \mathbf{m}(\theta, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\}, \quad (2.10)$$

where

$$\mathbf{m}(\theta, \mathbf{u}) \quad := \quad \int_\mathbf{I} \phi'\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right) \mathrm{d}\mathbf{u} - \left\{ \phi'\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right) \frac{1}{\mathbf{c}_\theta(\mathbf{u})} - \phi\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right) \right\}.$$

In the sequel we denote by $\frac{\partial}{\partial\theta}\mathbf{m}(\theta, \mathbf{u})$ the p-dimensional vector with entries $\frac{\partial}{\partial\theta_i}\mathbf{m}(\theta, \mathbf{u})$ and $\frac{\partial^2}{\partial\theta^2}\mathbf{m}(\theta, \mathbf{u})$ the $p \times p$-matrix with entries $\frac{\partial^2}{\partial\theta_i\partial\theta_j}\mathbf{m}(\theta, \mathbf{u})$. In what follows, we give some examples of divergences and the associated estimates.

2.1. Examples

- Our first example is the common used modified Kullback–Leibler divergence

$$\begin{aligned} \phi(x) &= -\log x + x - 1 \\ \phi'(x) &= -\frac{1}{x} + 1 \\ x\phi'(x) - \phi(x) &= \log x. \end{aligned}$$

The estimate of $D_{\mathrm{KL_m}}(\theta_0, \theta_\mathbf{T})$ is given by

$$\begin{aligned} \widehat{D}_{\mathrm{KL_m}}(\theta_0, \theta_\mathbf{T}) &= \sup_{\theta \in \mathbf{\Theta}_e} \left\{ - \int_\mathbf{I} \log\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right) \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\} \\ &= \sup_{\theta \in \mathbf{\Theta}_e} \left\{ \int_\mathbf{I} \log\left(\mathbf{c}_\theta(\mathbf{u})\right) \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\} \end{aligned}$$

and the estimate of the parameter $\theta_\mathbf{T}$ is given by

$$\widehat{\theta}_n := \arg \sup_{\theta \in \mathbf{\Theta}_e} \left\{ \int_\mathbf{I} \log\left(\mathbf{c}_\theta(\mathbf{u})\right) \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\},$$

which coincides with the MPL one.

- The second one is the Kullback–Leibler divergence

$$\begin{aligned}\phi(x) &= x\log x - x + 1\\ \phi'(x) &= \log x\\ x\phi'(x) - \phi(x) &= x - 1.\end{aligned}$$

The estimate of $D_{\mathrm{KL}}(\theta_0, \theta_{\mathbf{T}})$ is given by

$$\widehat{D}_{\mathrm{KL}}(\theta_0, \theta_{\mathbf{T}}) = \sup_{\theta\in\mathbf{\Theta}_e}\left\{\int_{\mathbf{I}} \log\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right) \mathrm{d}\mathbf{u} - \int_{\mathbf{I}}\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})} - 1\right) \mathrm{d}\mathbf{C}_n(\mathbf{u})\right\}$$

and the estimate of the parameter $\theta_{\mathbf{T}}$ is defined as follows

$$\widehat{\theta}_n := \arg\sup_{\theta\in\mathbf{\Theta}_e}\left\{\int_{\mathbf{I}} \log\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right) \mathrm{d}\mathbf{u} - \int_{\mathbf{I}}\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})} - 1\right) \mathrm{d}\mathbf{C}_n(\mathbf{u})\right\}.$$

- The third one is the χ^2-divergence

$$\begin{aligned}\phi(x) &= \frac{1}{2}(x-1)^2\\ \phi'(x) &= x - 1\\ x\phi'(x) - \phi(x) &= \frac{1}{2}x^2 - \frac{1}{2}.\end{aligned}$$

The estimate of $D_{\chi^2}(\theta_0, \theta_{\mathbf{T}})$ is given by

$$\begin{aligned}\widehat{D}_{\chi^2}(\theta_0, \theta_{\mathbf{T}}) &= \sup_{\theta\in\mathbf{\Theta}_e}\left\{\int_{\mathbf{I}}\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})} - 1\right) \mathrm{d}\mathbf{u}\right.\\ &\quad - \left.\int_{\mathbf{I}}\frac{1}{2}\left(\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)^2 - 1\right) \mathrm{d}\mathbf{C}_n(\mathbf{u})\right\}\end{aligned}$$

and the estimate of the parameter $\theta_{\mathbf{T}}$ is defined by

$$\widehat{\theta}_n := \arg\sup_{\theta\in\mathbf{\Theta}_e}\left\{\int_{\mathbf{I}}\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})} - 1\right) \mathrm{d}\mathbf{u} - \int_{\mathbf{I}}\frac{1}{2}\left(\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)^2 - 1\right) \mathrm{d}\mathbf{C}_n(\mathbf{u})\right\}.$$

- The last example is the Hellinger divergence

$$\begin{aligned}\phi(x) &= 2(\sqrt{x} - 1)^2\\ \phi'(x) &= 2 - \frac{1}{\sqrt{x}}\\ x\phi'(x) - \phi(x) &= 2\sqrt{x} - 2.\end{aligned}$$

The estimate of $D_{\mathrm{H}}(\theta_0, \theta_{\mathbf{T}})$ is given by

$$\widehat{D}_{\mathrm{H}}(\theta_0, \theta_{\mathbf{T}}) = \sup_{\theta\in\mathbf{\Theta}_e}\left\{\int_{\mathbf{I}}\left(2 - 2\sqrt{\mathbf{c}_\theta(\mathbf{u})}\right) \mathrm{d}\mathbf{u} - \int_{\mathbf{I}} 2\left(\frac{1}{\sqrt{\mathbf{c}_\theta(\mathbf{u})}} - 1\right) \mathrm{d}\mathbf{C}_n(\mathbf{u})\right\}$$

and the estimate of the parameter $\theta_{\mathbf{T}}$ is defined by

$$\widehat{\theta}_n := \arg\sup_{\theta\in\boldsymbol{\Theta}_e} \left\{ \int_{\mathbf{I}} \left(2 - 2\sqrt{\mathbf{c}_\theta(\mathbf{u})}\right) \mathrm{d}\mathbf{u} - \int_{\mathbf{I}} 2\left(\frac{1}{\sqrt{\mathbf{c}_\theta(\mathbf{u})}} - 1\right) \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\}.$$

All the above examples are particular cases of the so-called "*power divergences*", introduced by [5] (see also [25] Chapter 2), which are defined through the class of convex real valued functions

$$x \in \mathbb{R}_+^* \to \varphi_\gamma(x) := \frac{x^\gamma - \gamma x + \gamma - 1}{\gamma(\gamma - 1)}$$

for γ in $\mathbb{R}\setminus\{0,1\}$. The estimate of $D_\gamma(\theta_0, \theta_{\mathbf{T}})$ is given by

$$\begin{aligned} \widehat{D}_\gamma(\theta_0, \theta_{\mathbf{T}}) &= \sup_{\theta\in\boldsymbol{\Theta}_e} \left\{ \int_{\mathbf{I}} \frac{1}{\gamma - 1} \left(\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)^{\gamma-1} - 1 \right) \mathrm{d}\mathbf{u} \right. \\ &\quad - \left. \int_{\mathbf{I}} \frac{1}{\gamma} \left(\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)^{\gamma} - 1 \right) \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\} \end{aligned}$$

and the parameter estimate is defined by

$$\begin{aligned} \widehat{\theta}_n &:= \arg\sup_{\theta\in\boldsymbol{\Theta}_e} \left\{ \int_{\mathbf{I}} \frac{1}{\gamma - 1} \left(\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)^{\gamma-1} - 1 \right) \mathrm{d}\mathbf{u} \right. \\ &\quad - \left. \int_{\mathbf{I}} \frac{1}{\gamma} \left(\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)^{\gamma} - 1 \right) \mathrm{d}\mathbf{C}_n(\mathbf{u}) \right\}. \end{aligned}$$

Remark 2.1. Divergences measures have been intensively used in estimation and test in the framework of the discrete parametric models with independent identically distributed data; the estimates of the divergences and the parameter are obtained by the plug-in method; see [25] including the references therein. For continuous parametric models the plug-in procedure does not lead to well defined estimates; [3, 20, 26] introduce new estimates and tests, using the dual representation of divergences, extending the maximum likelihood procedure.

Remark 2.2. We give an example of copulas for which the likelihood-based procedure fails. We consider the Gumbel copulas $\mathbf{C}_\theta(\cdot)$ given in (1.3), it's corresponding density copula is defined by

$$\mathbf{c}_\theta(u_1, u_2) := \mathbf{C}_\theta(u_1, u_2)(u_1 u_2)^{-1} \frac{(\widetilde{u_1}\widetilde{u_2})^{(\theta-1)}}{(\widetilde{u_1}^\theta + \widetilde{u_2}^\theta)^{(2-1/\theta)}} \left[(\widetilde{u_1}^\theta + \widetilde{u_2}^\theta)^{(1/\theta)} + \theta - 1 \right], \quad (2.11)$$

where $\widetilde{x} = -\log x$. We can show that $\mathbf{c}_\theta(\cdot)$ may takes negative values for some $\theta \in \boldsymbol{\Theta}_e$. In fact $\mathbf{c}_{0.7}(u_1, u_2)$ is negative for $(u_1, u_2) \in [0.9, 1]^2$, hence the likelihood function is not well defined. The choice of the χ^2-divergence is particularly well adapted to this situation for example.

3. THE ASYMPTOTIC BEHAVIOR OF THE ESTIMATES

In this section, we provide the consistency of the estimates (2.10). We also state their asymptotic normality and evaluate their limiting variance. Statistics of the form

$$\mathbf{\Psi}_n := \int_{\mathbf{I}} \psi(\mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}),$$

belong to the general class of *multivariate rank statistics*. Their asymptotic properties have been investigated at length by a number of authors, among whom we may cite [34, 35] and [32]. In particular, the previous authors have provided regularity conditions, imposed on $\psi(\cdot)$, which imply the asymptotic normality of $\mathbf{\Psi}_n$. The corresponding arguments have been modified by [15], as to establish almost sure convergence of the estimators that they consider (see, e. g., [15] Proposition A.1). In the same spirit, the limiting behavior, as n tends to the infinity, of the estimators and test statistics which we will introduce later on, will make an instrumental use of the general theory of multivariate rank statistics, and rely, in particular, on Proposition A.1 in [15]. The existence and consistency of our estimators will be established through an application of the law of the iterated logarithm for empirical copula processes, in combination with general arguments from multivariate rank statistics theory (we refer to [6, 12] and references therein). We will use the following notations

$$\mathbb{K}_1(\theta, \mathbf{u}) := \phi'\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)$$

and

$$\mathbb{K}_2(\theta, \mathbf{u}) := \left\{\phi'\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)\frac{1}{\mathbf{c}_\theta(\mathbf{u})} - \phi\left(\frac{1}{\mathbf{c}_\theta(\mathbf{u})}\right)\right\}.$$

Definition 3.1. (i) Let $\mathscr{Q}$ be the set of continuous functions q on $[0,1]$ which are positive on $(0,1)$, symmetric about $1/2$, increasing on $[0,1/2]$ and satisfy $\int_0^1 \{q(t)\}^{-2}\,\mathrm{d}t < \infty$.

(ii) A function $r : (0,1) \to (0,\infty)$ is called u-shaped if it is symmetric about $1/2$ and increasing on $(0,1/2]$.

(iii) For $0 < \beta < 1$ and u-shaped function r, we define

$$r_\beta(t) = \begin{cases} r(\beta t) & \text{if} \quad 0 < t \leq 1/2; \\ r\{1-\beta(1-t)\} & \text{if} \quad 1/2 < t \leq 1/2. \end{cases}$$

If for $\beta > 0$ in a neighborhood of 0, there exists a constant M_β, such that $r_\beta \leq M_\beta r$ on $(0,1)$, then r is called a reproducing u-shaped function. We denote by $\mathscr{R}$ the set of reproducing u-shaped functions.

Typical examples of elements in $\mathscr{Q}$ and $\mathscr{R}$ are given by

$$q(t) = [t(1-t)]^\zeta,\ 0 < \zeta < 1/2, \qquad r(t) = \varrho\,[t(1-t)]^{-\varsigma},\ \varsigma \geq 0,\ \varrho \geq 0.$$

We make use of the following conditions.

(C.1) There exists a neighborhood $N(\theta_{\mathbf{T}}) \subset \boldsymbol{\Theta}_e$ of $\theta_{\mathbf{T}}$ such that the first and the second partial derivatives with respect to θ of $\mathbb{K}_1(\theta, \mathbf{u})$ are dominated on $N(\theta_{\mathbf{T}})$ by some λ-integrable functions;

(C.2) There exists a neighborhood $N(\theta_{\mathbf{T}})$ of $\theta_{\mathbf{T}}$, such that for all $\theta \in N(\theta_{\mathbf{T}})$, the functions $\frac{\partial}{\partial \theta_i}\mathbf{m}(\theta, \mathbf{u}) : (0,1)^d \to \mathbb{R}$ are continuously differentiable, and there exist functions $r_i \in \mathscr{R}$, $\widetilde{r}_i \in \mathscr{R}$ and $q \in \mathscr{Q}$ $(i, j = 1, \ldots, d, \quad i \neq j$ and $\ell, \ell', \ell'' = 1, \ldots, p)$ with

(i) $$\left|\frac{\partial}{\partial \theta_\ell}\mathbf{m}(\theta, \mathbf{u})\right| \leq \prod_{i=1}^{d} r_i(u_i), \quad \left|\frac{\partial^2}{\partial \theta_\ell \partial u_i}\mathbf{m}(\theta, \mathbf{u})\right| \leq \widetilde{r}_i(u_i) \prod_{i \neq j}^{d} r_j(u_j);$$

(ii) $$\left|\frac{\partial^3}{\partial \theta_\ell \partial \theta_{\ell'} \partial \theta_{\ell''}}\mathbb{K}_2(\theta, \mathbf{u})\right| \leq \prod_{i=1}^{d} r_i(u_i);$$

(iii) $$\left|\mathbf{m}(\theta, \mathbf{u})\right| \leq \prod_{i=1}^{d} r(u_i), \quad \left|\frac{\partial}{\partial u_i}\mathbf{m}(\theta, \mathbf{u})\right| \leq \widetilde{r}_i(u_i) \prod_{i \neq j}^{d} r_j(u_j);$$

(iv) $$\left|\frac{\partial}{\partial \theta_\ell}\mathbf{m}(\theta, \mathbf{u})\right|^2 \leq \prod_{i=1}^{d} r_i(u_i), \quad \left|\frac{\partial^2}{\partial \theta_\ell \partial \theta_{\ell'}}\mathbf{m}(\theta, \mathbf{u})\right| \leq \prod_{i=1}^{d} r_i(u_i)$$

and

$$\int_{\mathbf{I}} \left\{\prod_{i=1}^{d} r_i(u_i)\right\}^2 \mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u}) < \infty,$$

$$\int_{\mathbf{I}} \left\{q_i(u_i)\widetilde{r}_i(u_i) \prod_{i \neq j}^{d} r_j(u_j)\right\} \mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u}) < \infty, \quad \text{for} \quad i = 1, \ldots, d;$$

(C.3) The matrix $\int_{\mathbf{I}} (\partial^2/\partial^2\theta)\mathbf{m}(\theta, \mathbf{u})\, \mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u})$ is non singular;

(C.4) The function $\mathbf{u} \in \mathbf{I} \mapsto \frac{\partial}{\partial \theta}\mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})$ is of bounded variation on $\mathbf{I}$.

The main result to be proved here may now be stated precisely as follows.

Theorem 3.2. Assume that conditions C.1 − C.4 hold.

1. Let $B(\theta_T, n^{-1/3}) := \left\{\theta \in \boldsymbol{\Theta}_e, \|\theta - \theta_T\| \leq n^{-1/3}\right\}$, then as n tends to infinity, with probability one, the function $\theta \mapsto \int_{\mathbf{I}} \mathbf{m}(\theta, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u})$ attains its maximum value at some point $\widehat{\theta}_n$ in the interior of $B(\theta_T, n^{-1/3})$, which implies that the estimate $\widehat{\theta}_n$ is consistent and satisfies

$$\int_{\mathbf{I}} \frac{\partial}{\partial \theta}\mathbf{m}(\widehat{\theta}_n, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) = 0.$$

2. $\sqrt{n}(\widehat{\theta}_n - \theta)$ converges in distribution to a centered multivariate normal random variable with covariance matrix

$$\mathbf{\Xi}_\phi = \mathbf{S}^{-1}\mathbf{M}\mathbf{S}^{-1}, \tag{3.1}$$

with

$$\mathbf{S} := -\int_{\mathbf{I}} \frac{\partial^2}{\partial \theta^2} \mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u}) \, \mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u}), \tag{3.2}$$

and

$$\mathbf{M} := \mathrm{Var}\left[\frac{\partial}{\partial \theta}\mathbf{m}(\theta_{\mathbf{T}}, F_1(X_1), \ldots, F_d(X_d)) + \sum_{i=1}^{d} \mathbb{W}_i(\theta_{\mathbf{T}}, X_i)\right], \tag{3.3}$$

where

$$\mathbb{W}_i(\theta_{\mathbf{T}}, X_i) := \int_{\mathbf{I}} \left\{ \mathbb{1}_{\{F_i(X_i) \le u_i\}} - u_i \right\} \tfrac{\partial^2}{\partial \theta \partial u_i} \mathbf{m}\left(\theta_{\mathbf{T}}, \mathbf{u}\right) \, \mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u}), \; i = 1, \ldots, d.$$

The proof of Theorem 3.2 is postponed to the Appendix.

Remark 3.3. The aim of Theorem 3.2 part (a) is not to establish the optimal rate of the estimate but merely the existence and the consistency (a.s.) of the estimate. We have considered $n^{1/3}$ because it works well, indeed, in Taylor expansion (A.11), in the proof, the third term of the RHS is $O(1)$ only for this rate, which is the major key of the demonstration.

4. NEW TESTS OF INDEPENDENCE

One of our motivation is to build a statistical test of independence, based on ϕ-divergence. In the framework of the parametric copula model, the null hypothesis, i. e., the independence case

$$\mathbf{C}_{\theta_0}(u_1, \ldots, u_d) = \prod_{i=1}^{d} u_i$$

corresponds to

$$\mathscr{H}_0 \; : \; \theta_{\mathbf{T}} = \theta_0.$$

We consider the composite alternative hypothesis

$$\mathscr{H}_1 \; : \; \theta_{\mathbf{T}} \neq \theta_0.$$

Since, θ_0 is a boundary value of the parameter space $\mathbf{\Theta}$, we can see that the convergence in distribution of the corresponding pseudo-likelihood ratio statistic to a χ^2 random variable does not hold; see [2]. We give now a solution to this problem. We propose the following statistics

$$\mathbf{T}_n := \frac{2n}{\phi''(1)} \widehat{D_\phi}(\theta_0, \theta_{\mathbf{T}}). \tag{4.1}$$

We will see that the proposed statistic converges in distribution, under the null hypothesis $\mathscr{H}_0$, to a χ^2 random variable with p degrees of freedom, which permits to build a test of $\mathscr{H}_0$ against $\mathscr{H}_1$ asymptotically of level α. The limit law of $\mathbf{T}_n$ is given also under the alternative hypothesis $\mathscr{H}_1$. We will use the following additional conditions.

(C.5) We have

$$\lim_{\theta\to\theta_0} \frac{\partial^2}{\partial\theta_\ell \partial u_i} \mathbf{m}(\theta, \mathbf{u}) = 0,$$

and there exists a neighborhood $N(\theta_0)$ of θ_0 and there exist functions $r_i \in \mathscr{R}$, $\widetilde{r}_i \in \mathscr{R}$ and $q_i \in \mathscr{Q}$ $(i = 1, \ldots, d$ and $\ell = 1, \ldots, p)$, such that for all $\theta \in N(\theta_0)$,

$$\left| \frac{\partial^2}{\partial\theta_\ell \partial u_i} \mathbf{m}(\theta, \mathbf{u}) \right| < \widetilde{r}_i(u_i) \prod_{j\neq j}^{d} r(u_j)$$

and

$$\int_{\mathbf{I}} \left\{ q_i(u_i) \widetilde{r}_i(u_i) \prod_{i\neq j}^{d} r_j(u_j) \right\} \mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}(\mathbf{u}) < \infty.$$

Remark 4.1. When $\theta_\mathbf{T} = \theta_0$, under the conditions (C.1) and (C.5) we can see that $\mathbf{S}$ and $\mathbf{M}$ can be written as

$$\mathbf{S} = \mathbf{M} = \int_{\mathbf{I}} \left[\frac{\partial}{\partial\theta} \mathbf{m}(\theta_\mathbf{T}, \mathbf{u}) \right] \left[\frac{\partial}{\partial\theta} \mathbf{m}(\theta_\mathbf{T}, \mathbf{u}) \right]^\top \mathrm{d}\mathbf{C}_{\theta_\mathbf{T}}(\mathbf{u}).$$

The following theorem gives the limiting law of the statistics $\mathbf{T}_n$ under the both hypothesis $\mathscr{H}_0$ and $\mathscr{H}_1$.

Theorem 4.2. (1) Assume that conditions C.1 − C.5 hold. If $\theta_\mathbf{T} = \theta_0$, then the statistic T_n converges in distribution to a χ^2 variable with p degrees of freedom.

(2) Assume that conditions C.1 − C.5 hold. If $\theta_\mathbf{T} \neq \theta_0$, then

$$\sqrt{n} \left(\widehat{D_\phi}(\theta_0, \theta_T) - D_\phi(\theta_0, \theta_T) \right)$$

converges in distribution to a centered normal variable with variance

$$\sigma_\phi^2(\theta_0, \theta_\mathbf{T}) := \mathrm{Var} \left[\mathbf{m}(\theta_\mathbf{T}, F_1(X_1), \ldots, F_d(X_d)) + \sum_{i=1}^{d} \mathbb{Y}_i(\theta_\mathbf{T}, X_i) \right], \tag{4.2}$$

where

$$\mathbb{Y}_i(\theta_\mathbf{T}, X_i) := \int_{\mathbf{I}} \left\{ \mathbb{1}_{\{F_i(X_i)\leq u_i\}} - u_i \right\} \tfrac{\partial}{\partial u_i} \mathbf{m}\left(\theta_\mathbf{T}, \mathbf{u}\right) \mathbf{c}_{\theta_\mathbf{T}}(\mathbf{u})\, \mathrm{d}u_1 \ldots \mathrm{d}u_d,\ i = 1, \ldots, d.$$

The proof of Theorem 4.2 is postponed to the Appendix.

Remark 4.3. An application of Theorem 4.2, leads to reject the null hypothesis $\mathscr{H}_0 : \theta_\mathbf{T} = \theta_0$, whenever the value of the statistic $\mathbf{T}_n$ exceeds $q_{1-\alpha}$, namely, the $(1-\alpha)$-quantile of the χ^2 law with p degrees of freedom. The corresponding test is then, asymptotically of level α, when $n \to \infty$. The critical region is, accordingly, given by

$$CR := \{\mathbf{T}_n > q_{1-\alpha}\}.$$

The fact that this test is consistent follows from Theorem 4.2. Further, this theorem can be used to give an approximation to the power function $\theta_{\mathbf{T}} \in \mathbf{\Theta} \mapsto \beta(\theta_{\mathbf{T}}) := P_{\theta_{\mathbf{T}}}\{CR\}$ in a similar way to [21]. We so obtain that

$$\beta(\theta_{\mathbf{T}}) \approx 1 - \Phi\left(\frac{\sqrt{n}}{\sigma_\phi(\theta_0, \theta_{\mathbf{T}})}\left(\frac{q_{1-\alpha}}{2n} - D_\phi(\theta_0, \theta_T)\right)\right), \tag{4.3}$$

where Φ denotes, as usual, the cumulative distribution function of a $\mathcal{N}(0,1)$ standard normal random variable. A useful consequence of (4.3) is the possibility of computing an approximate value of the sample size ensuring a specified power $\beta(\theta_{\mathbf{T}})$, with respect to some pre-assigned alternative $\theta_{\mathbf{T}} \neq \theta_0$. Let n_0 be the positive root of the equation

$$\beta = 1 - \Phi\left(\frac{\sqrt{n}}{\sigma_\phi(\theta_0, \theta_{\mathbf{T}})}\left(\frac{q_{1-\alpha}}{2n} - D_\phi(\theta_0, \theta_T)\right)\right),$$

which can be rewritten into

$$n_0 = \frac{(a+b) - \sqrt{a(a+2b)}}{2D_\phi(\theta_0, \theta_T)^2},$$

where $a := \sigma_\phi(\theta_0, \theta_{\mathbf{T}})\left(\Phi^{-1}(1-\beta)\right)^2$ and $b := q_{1-\alpha} D_\phi(\theta_0, \theta_T)$. The sought-after approximate value of the sample size is then given

$$n^* := \lfloor n_0 \rfloor + 1,$$

where $\lfloor u \rfloor$ denote the integer part of u.

Remark 4.4. For point estimation, the estimator based on ϕ-divergence when we extend the parameter space, may not have a meaningful interpretation and most probably has a larger mean square error. However, from Theorem 3.2 and 4.2, it is clear that an asymptotic $1-\alpha$ confidence interval or region, $\mathscr{R}_\alpha$ about θ can be easily constructed using the intersection method as described in [11].

Remark 4.5. The above regularity conditions are satisfied for a large number of single-parameter families of bivariate copulas including the standard bivariate normal, the Farlie–Gumbel–Morgenstern system, and copulas of the Archimedean variety such as those of Ali–Mikhail–Haq and Frank; see, e. g., [15] and [42]. Note that the score functions for some copulas are unbounded near the origin or the point $(1,1)$, so we need to know the above regularity conditions, at least theoretically as be mentioned in [42].

Remark 4.6. The parameters (3.2) and (3.3) may be consistently estimated respectively by the sample mean of

$$\frac{\partial^2}{\partial\theta^2}\mathbf{m}(\widehat{\theta}_n, F_{1n}(X_{1,k}), \ldots, F_{dn}(X_{d,k})), \quad k = 1, \ldots, n, \tag{4.4}$$

and the sample variance of

$$\tfrac{\partial}{\partial\theta}\,\mathbf{m}\left(\widehat{\theta}_n, F_{1n}(X_{1,k}), \ldots, F_{dn}(X_{d,k})\right) + \sum_{i=1}^{d} \mathbb{W}_i(\widehat{\theta}_n, X_{i,k}), \quad k = 1, \ldots, n, \tag{4.5}$$

as was done in [15]. The asymptotic variance (4.2) can be consistently estimated in the same way.

Remark 4.7. The set $\mathbf{\Theta}_e$ defined in (2.7) is generally with non empty interior $\mathring{\mathbf{\Theta}}_e$. In particular, we may check that θ_0 (the value corresponding to independence) belongs to $\mathring{\mathbf{\Theta}}_e$, since the integral in (2.7) is finite; it is equal to zero when $\theta = \theta_0$, for any copula density $\mathbf{c}_\theta(\cdot)$. However, it is hard to determine the whole set $\mathbf{\Theta}_e$ for some copulas, but in order to test the independence, we need only to prove the existence of a neighborhood $N(\theta_0)$ of θ_0 for which the integral in (2.7) is finite since we calculate the estimate $\widehat{\theta}_n$ in (2.10) by Newton–Raphson algorithm using θ_0 as initial point. The explicit calculation of the integral in (2.7) may be complicated for some copulas, in such cases we use the Monte Carlo method to compute this integral.

5. SIMULATIONS

In this section, we report the results from simulation experiments carried out to assess the performance of the proposed estimators. To this end, we have considered the FGM copula. For the experiment considered here, we compute the MPL, KL-divergence, χ^2-divergence, Hellinger divergence and some power divergence estimates, and report their variance value, bias and mean-squared error. In order to compare the robustness of the proposed estimates we consider several scenarios of contamination. To be more precise, we considered ϵ-contaminated models, where a proportion ϵ of observations were replaced by atypical ones generated from a contaminating distribution $F^*(\cdot,\cdot)$. We set ϵ equal to 0 %, 5 %, and 10 %, and $F^*(\cdot,\cdot)$ as the bivariate normal distribution with correlation coefficient $\rho = 0$ and very small variances, acting as a point mass contamination as in [27]. The sample size is $n = 500$ and the estimates are obtained from 1000 independent runs.

(i) Under no contamination: All procedures showed reasonable accuracy. The Hellinger and χ^2-divergence estimates seem to be as good as the MPL estimator. This is more evident as the sample size gets larger; see Table 1.

(ii) Under 5 % and 10 % contamination: MPL estimator is recommended when there is no contamination but its performance may deteriorate rapidly if the sample is pooled, see Tables 2 and 3. In the contaminated case the power divergence estimator with $\gamma = 2.5$ is superior with respect to the others. It seems that the χ^2-divergence estimate behaves well also for contaminated data.

From the three Tables 1, 2 and 3, we can see that the χ^2-divergence estimates is a good trad-off between efficiency and robustness.

In future work, it would be interesting to provide a complete investigation of robustness of semiparametric copula estimator which requires nontrivial mathematics, this would go well beyond the scope of the present paper.

Table 1. No contamination: $\epsilon = 0.00$.

$\theta_T = 0.01,\ n = 500,\ rep = 1000$				
Divergence	Estimate	Variance	Bias	MSE
$\gamma = -0.5$	0.0924	0.0177	0.0076	0.0177
$\gamma = 0$ (MPL)	0.0920	0.0176	0.0080	0.0176
$\gamma = 0.5$ (H)	0.0916	0.0175	0.0084	0.0176
$\gamma = 1$ (KL)	0.0913	0.0175	0.0087	0.0176
$\gamma = 1.5$	0.0911	0.0176	0.0089	0.0177
$\gamma = 2$ (χ^2)	0.0911	0.0178	0.0089	0.0179
$\gamma = 2.5$	0.0912	0.0181	0.0088	0.0181

Table 2. Contamination: $\epsilon = 0.05$.

$\theta_T = 0.01,\ n = 500,\ rep = 1000$				
Divergence	Estimate	Variance	Bias	MSE
$\gamma = -0.5$	0.0949	0.0191	0.0051	0.0191
$\gamma = 0$ (MPL)	0.0921	0.0181	0.0079	0.0181
$\gamma = 0.5$ (H)	0.0895	0.0172	0.0105	0.0173
$\gamma = 1$ (KL)	0.0873	0.0164	0.0127	0.0166
$\gamma = 1.5$	0.0853	0.0158	0.0147	0.0160
$\gamma = 2$ (χ^2)	0.0835	0.0153	0.0165	0.0155
$\gamma = 2.5$	0.0820	0.0149	0.0180	0.0152

Table 3. Contamination: $\epsilon = 0.10$.

$\theta_T = 0.01,\ n = 500,\ rep = 1000$				
Divergence	Estimate	Variance	Bias	MSE
$\gamma = -0.5$	0.0916	0.0191	0.0084	0.0191
$\gamma = 0$ (MPL)	0.0867	0.0171	0.0133	0.0173
$\gamma = 0.5$ (H)	0.0825	0.0155	0.0175	0.0158
$\gamma = 1$ (KL)	0.0787	0.0141	0.0213	0.0146
$\gamma = 1.5$	0.0754	0.0130	0.0246	0.0136
$\gamma = 2$ (χ^2)	0.0724	0.0120	0.0276	0.0128
$\gamma = 2.5$	0.0698	0.0112	0.0302	0.0121

6. CONCLUDING REMARKS

We have introduced a new estimation and test procedure in parametric copula models with unknown margins. The method is based on divergences between copulas and the duality technique. It generalizes the maximum pseudo-likelihood one, and applies both when the parameter is an interior or a boundary value, in particular for testing the null hypothesis of independence. Simulation results show that the χ^2-divergence estimate is a good trade-off between efficiency and robustness. It will be interesting to investigate theoretically the problem of the choice of the divergence which leads to an "*optimal*" (in some sense) estimate or test in terms of efficiency and robustness.

APPENDIX

First we give a technical Lemma which we will use to prove our results.

Lemma A.1. Let $F_{\theta_{\mathbf{T}},F_1,\ldots,F_d}(\cdot)$ have a continuous margins and let $\mathbf{C}_{\theta_{\mathbf{T}}}(\cdot)$ have continuous partial derivatives. Assume that $\xi(\cdot)$ is a continuous function, with bounded variation. Then

$$\int_{\mathbf{I}} \xi(\mathbf{u})\,\mathrm{d}\left(\mathbf{C}_n(\mathbf{u}) - \mathbf{C}(\mathbf{u})\right) = O\left(n^{-1/2}(\log\log n)^{1/2}\right) \quad \text{(a.s.)}. \tag{A.1}$$

P r o o f o f L e m m a A.1. Recall that the *modified empirical copula* $\mathbf{C}_n(\cdot)$, is slightly different from the *empirical copula* $\mathbb{C}_n(\cdot)$, introduced by [6], and defined by

$$\mathbb{C}_n(\mathbf{u}) = F_n\left(F_{1n}^{-1}(u_1), \ldots, F_{dn}^{-1}(u_d)\right) \quad \text{for} \quad \mathbf{u} \in (0,1)^d, \tag{A.2}$$

where $F_{in}^{-1}(\cdot)$ for $i = 1, \ldots, d$ denote the empirical quantile functions, associated with $F_{in}(\cdot)$ for $i = 1, \ldots, d$, respectively, and defined by

$$F_{in}^{-1}(t) := \inf\{x \in \mathbb{R} \mid F_{in}(x) \geq t\}, \quad i = 1, \ldots, d.$$

Note that the subtle difference lies in the fact that $\mathbb{C}_n(\cdot)$ is left-continuous with right-hand limits, whereas $\mathbf{C}_n(\cdot)$ on the other hand is right continuous with left-hand limits. The difference between $\mathbb{C}_n(\cdot)$ and $\mathbf{C}_n(\cdot)$, however, is small

$$\sup_{\mathbf{u}\in I} |\mathbb{C}_n(\mathbf{u}) - \mathbf{C}_n(\mathbf{u})| = \frac{1}{n}. \tag{A.3}$$

As in the proof of Lemma 5.1 in [2], using integration by parts, we can prove that there exists a constant $\kappa > 0$, depending upon d only, such that

$$\left|\sqrt{n}\int_{\mathbf{I}} \xi(\mathbf{u})\ \mathrm{d}(\mathbf{C}_n - \mathbf{C})(\mathbf{u})\right| \leq \kappa\sqrt{n}\sup_{\mathbf{u}\in I} |(\mathbf{C}_n - \mathbf{C})(\mathbf{u})| \int_{\mathbf{I}} \mathrm{d}\,|\xi(\mathbf{u})|\,.$$

In fact, by Fubini's Theorem, we can write

$$\begin{aligned}
&\left|\sqrt{n}\int_{\mathbf{I}}\xi(u_1,\ldots,u_d)\,\mathrm{d}(\mathbf{C}_n-\mathbf{C})(u_1,\ldots,u_d)\right| \\
&= \left|\sqrt{n}\int_{\mathbf{I}}\left\{\int_0^{u_1}\cdots\int_0^{u_d} d\xi(s_1,\ldots,s_d)\right\}\mathrm{d}(\mathbf{C}_n-\mathbf{C})(u_1,\ldots,u_d)\right| \\
&= \left|\sqrt{n}\int_{\mathbf{I}}\left\{\int_{\mathbf{I}}\mathbb{1}_{\{s_1\le u_1\}}\ldots\mathbb{1}_{\{s_d\le u_d\}}\mathrm{d}\xi(s_1,\ldots,s_d)\right\}\mathrm{d}(\mathbf{C}_n-\mathbf{C})(u_1,\ldots,u_d)\right| \\
&= \left|\sqrt{n}\int_{\mathbf{I}}\left\{\int_{\mathbf{I}}\mathbb{1}_{\{s_1\le u_1\}}\ldots\mathbb{1}_{\{s_d\le u_d\}}\mathrm{d}(\mathbf{C}_n-\mathbf{C})(u_1,\ldots,u_d)\right\}\mathrm{d}\xi(s_1,\ldots,s_d)\right| \\
&= \left|\sqrt{n}\int_{\mathbf{I}}\left\{\Delta_{\mathbf{s}}^{\mathbf{1}}(\mathbf{C}_n-\mathbf{C})(u_1,\ldots,u_d)\right\}\ \mathrm{d}\xi(s_1,\ldots,s_d)\right| \\
&\le (2^d-1)\sqrt{n}\sup_{\mathbf{u}\in I}|(\mathbf{C}_n-\mathbf{C})(\mathbf{u})|\int_{\mathbf{I}}\mathrm{d}\,|\xi(\mathbf{u})|\,,
\end{aligned}$$

where for $\mathbf{a}$ and $\mathbf{b}$ in $\mathbf{I}$

$$\Delta_{\mathbf{a}}^{\mathbf{b}}(\mathbf{C}_n-\mathbf{C})(\mathbf{u}) := \Delta_{a_d}^{b_d}\Delta_{a_{d-1}}^{b_{d-1}}\cdots\Delta_{a_2}^{b_2}\Delta_{a_1}^{b_1}(\mathbf{C}_n-\mathbf{C})(\mathbf{u})$$

and for $j=1,\ldots,d$

$$\begin{aligned}
\Delta_{a_j}^{b_j}(\mathbf{C}_n-\mathbf{C})(\mathbf{u}) \quad := \quad & (\mathbf{C}_n-\mathbf{C})(u_1,\ldots,u_{j-1},b_j,u_{j-1},\ldots,u_d) \\
& -(\mathbf{C}_n-\mathbf{C})(u_1,\ldots,u_{j-1},a_j,u_{j-1},\ldots,u_d).
\end{aligned}$$

One may check (see Theorem 3.1 in [6]) that there exists a constant γ (depending upon $\mathbf{C}(\cdot)$ only) such that, with probability 1,

$$\limsup_{n\to\infty}\left\{\frac{n}{\log\log n}\right\}^{1/2}\sup_{\mathbf{u}\in I}|\mathbb{C}_n(\mathbf{u})-\mathbf{C}(\mathbf{u})| = \gamma < \infty. \tag{A.4}$$

From this and (A.3), applying (A.4), we obtain

$$\int_{\mathbf{I}}\xi(\mathbf{u})\,\mathrm{d}(\mathbf{C}_n-\mathbf{C})(\mathbf{u}) = O\left(n^{-1/2}(\log\log n)^{1/2}\right) \quad \text{(a.s.)}.$$

□

Proof of Theorem 3.2. (1) Under the Assumptions (C.1) and (C.2.ii), a simple calculation gives

$$\int_{\mathbf{I}}\frac{\partial}{\partial\theta}\mathbf{m}(\theta,\mathbf{u})\,\mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u}) = 0, \tag{A.5}$$

and

$$\int_{\mathbf{I}}\frac{\partial^2}{\partial\theta^2}\mathbf{m}(\theta,\mathbf{u})\,\mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u}) = -\int_{\mathbf{I}}\phi''\left(\frac{1}{\mathbf{c}_{\theta_{\mathbf{T}}}}\right)\frac{\dot{\mathbf{c}}_{\theta_{\mathbf{T}}}\dot{\mathbf{c}}_{\theta_{\mathbf{T}}}^{\top}}{\mathbf{c}_{\theta_{\mathbf{T}}}^3}\,\mathrm{d}\lambda = -\mathbf{S}. \tag{A.6}$$

We see that the matrix $\mathbf{S}$ is symmetric and positive using the fact that the second derivative $\phi''(\cdot)$ is nonnegative by the assumption that the function $\phi(\cdot)$ is convex. Hence, $\mathbf{S}$ is positive definite by (C.3). Introduce the statistic $\mathbf{\Phi}_n(\theta_{\mathbf{T}})$ defined by

$$\mathbf{\Phi}_n(\theta_{\mathbf{T}}) := \int_{\mathbf{I}}\frac{\partial}{\partial\theta}\mathbf{m}(\theta,\mathbf{u})\,\mathrm{d}\mathbf{C}_n(\mathbf{u}), \tag{A.7}$$

and combine (A.5) and condition (C2)(i) with Theorem 2.1 in [35] to show that, as $n \to \infty$

$$\sqrt{n}\mathbf{\Phi}_n(\theta_{\mathbf{T}}) \xrightarrow{d} \mathcal{N}(0, \mathbf{M}), \tag{A.8}$$

where $\mathbf{M}$ is defined in (3.3). We can refer also to the Proposition 3 in [42] for the same result in (A.8). Denote

$$\mathbf{\Upsilon}_n(\theta_{\mathbf{T}}) := \int_{\mathbf{I}} \frac{\partial^2}{\partial\theta^2}\mathbf{m}(\theta, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}), \tag{A.9}$$

we make use of (A.6) and (C.2.iv) in connection with Proposition A.1 of [15], one finds

$$\mathbf{\Upsilon}_n(\theta_{\mathbf{T}}) \to -\mathbf{S}, \quad \text{(a.s.)}. \tag{A.10}$$

We recall that $\mathbf{S}$ is in (3.2). Now, for any $\theta = \theta_{\mathbf{T}} + \mathbf{v}n^{-1/3}$ with $\|\mathbf{v}\| \leq 1$, consider a Taylor expansion of $\int_{\mathbf{I}} \mathbf{m}(\theta, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u})$ in θ around $\theta_{\mathbf{T}}$, and use (A.10), and (C.2.ii) to obtain

$$\begin{aligned} & n\int_{\mathbf{I}} \mathbf{m}(\theta, \mathbf{u}) d\mathbf{C}_n(\mathbf{u}) - n\int_{\mathbf{I}} \mathbf{m}(\theta_T, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) \\ = \; & n^{2/3}\mathbf{v}^\top \mathbf{\Phi}_n(\theta_{\mathbf{T}}) + 2^{-1}n^{1/3}\mathbf{v}\mathbf{S}\mathbf{v}^\top + O(1) \quad \text{(a.s.)} \end{aligned} \tag{A.11}$$

uniformly in $\mathbf{v}$ with $\|\mathbf{v}\| \leq 1$. On the other hand, since

$$\int_{\mathbf{I}} \frac{\partial}{\partial\theta_\ell}\mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})^2\, \mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u}) < \infty,$$

and $\frac{\partial}{\partial\theta_\ell}\mathbf{m}(\theta, \cdot)$ is of bounded variation by assumption (C.4)($\ell = 1, \ldots, p$), using Lemma A.1 we can show that

$$\int_{\mathbf{I}} \frac{\partial}{\partial\theta_\ell}\mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) = O\left(n^{-1/2}(\log\log n)^{1/2}\right) \quad \text{(a.s.)}. \tag{A.12}$$

Therefore, using (A.11) and (A.12), we obtain for any $\theta = \theta_{\mathbf{T}} + \mathbf{v}n^{-1/3}$ with $\|\mathbf{v}\| = 1$:

$$\begin{aligned} & n\int_{\mathbf{I}} \mathbf{m}(\theta, \mathbf{u})\ d\mathbf{C}_n(\mathbf{u}) - n\int_{\mathbf{I}} \mathbf{m}(\theta_T, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) \\ & \leq O(n^{1/6}(\log\log n)^{1/2}) - 2^{-1}\vartheta n^{1/3} + O(1) \quad \text{(a.s.)}, \end{aligned} \tag{A.13}$$

where ϑ is the smallest eigenvalue of the matrix $\mathbf{S}$. Observe that ϑ is positive since $\mathbf{S}$ is symmetric, positive and non singular by assumption (C.3). Using (A.13) and the fact that the function $\theta \mapsto \int_{\mathbf{I}} \mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u})$ is continuous, we conclude that as $n \to \infty$, with probability one, $\theta \mapsto \int_{\mathbf{I}} \mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u})$ reaches its maximum value at some point $\widehat{\theta}_n$ fulfills

$$\int_{\mathbf{I}} \frac{\partial}{\partial\theta}\mathbf{m}(\widehat{\theta}_n, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) = 0$$

and

$$\|\widehat{\theta}_n - \theta_{\mathbf{T}}\| = O(n^{-1/3}).$$

(2) Using the first part of Theorem 3.2, by a Taylor expansion of

$$\int_{\mathbf{I}} \frac{\partial}{\partial\theta}\mathbf{m}(\widehat{\theta}_n, \mathbf{u})\ \mathrm{d}\mathbf{C}_n(\mathbf{u}),$$

in $\widehat{\theta}_n$ around $\theta_{\mathbf{T}}$, we obtain

$$0 = \int_{\mathbf{I}} \frac{\partial}{\partial \theta}\mathbf{m}(\widehat{\theta}_n, \mathbf{u})\,\mathrm{d}\mathbf{C}_n(\mathbf{u})$$
$$= \int_{\mathbf{I}} \frac{\partial}{\partial \theta}\mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\,\mathrm{d}\mathbf{C}_n(\mathbf{u}) + (\widehat{\theta}_n - \theta_T)^\top \int_{\mathbf{I}} \frac{\partial^2}{\partial \theta^2}\mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\,\mathrm{d}\mathbf{C}_n(\mathbf{u}) + o(n^{-1/2}).$$

Hence,

$$\sqrt{n}(\widehat{\theta}_n - \theta_T) = -\left[\boldsymbol{\Upsilon}_n(\theta_{\mathbf{T}})\right]^{-1}\sqrt{n}\boldsymbol{\Phi}_n(\theta_{\mathbf{T}}) + o_P(1). \tag{A.14}$$

Using (A.8) and (A.10), by Slutsky theorem, we conclude then

$$\sqrt{n}(\widehat{\theta}_n - \theta_T) \to \mathcal{N}(0, \boldsymbol{\Xi}_\phi), \tag{A.15}$$

where we recall that $\boldsymbol{\Xi}_\phi$ is defined in (3.1). □

Proof of Theorem 4.2. (1) Assume that $\theta_{\mathbf{T}} = \theta_0$. Hence, from (A.14), using (A.6), we obtain

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_{\mathbf{T}}\right) = -\mathbf{S}^{-1}\sqrt{n}\boldsymbol{\Phi}_n(\theta_{\mathbf{T}}) + o_P(1). \tag{A.16}$$

On the other hand, expanding in Taylor series

$$\frac{2n}{\phi''(1)}\widehat{D_\phi}(\theta_0, \widehat{\theta}_n) = \frac{2n}{\phi''(1)}\int_{\mathbf{I}} \mathbf{m}(\widehat{\theta}_n, \mathbf{u})\,\mathrm{d}\mathbf{C}_n(\mathbf{u})$$

in $\widehat{\theta}_n$ around $\theta_{\mathbf{T}}$, in connection with the fact that $\int_{\mathbf{I}} \mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\,\mathrm{d}\mathbf{C}_n(\mathbf{u}) = 0$, we get

$$\frac{2n}{\phi''(1)}\widehat{D_\phi}(\theta_0, \widehat{\theta}_n) = \frac{2n}{\phi''(1)}\boldsymbol{\Phi}_n(\theta_{\mathbf{T}})(\widehat{\theta}_n - \theta_{\mathbf{T}}) - \frac{n}{\phi''(1)}(\widehat{\theta}_n - \theta_{\mathbf{T}})^\top \boldsymbol{\Upsilon}_n(\theta_{\mathbf{T}})(\widehat{\theta}_n - \theta_{\mathbf{T}}) + o_P(1).$$

Using (A.6), (A.16) and the fact that $\mathbf{S} = \phi^{''}(1)\mathbb{I}_{\theta_{\mathbf{T}}}$ ($\mathbb{I}_{\theta_{\mathbf{T}}}$ denotes the Fisher information matrix) when $\theta_{\mathbf{T}} = \theta_0$ to obtain

$$\frac{2n}{\phi''(1)}\widehat{D_\phi}(\theta_0, \widehat{\theta}_n) = \frac{1}{\phi''(1)}\sqrt{n}\boldsymbol{\Phi}_n(\theta_{\mathbf{T}})\mathbb{I}_{\theta_{\mathbf{T}}}^{-1}\sqrt{n}\boldsymbol{\Phi}_n(\theta_{\mathbf{T}}).$$

Finally, use the convergence in (A.8) and the fact that $\mathbf{M} = \phi^{''}(1)\mathbb{I}_{\theta_{\mathbf{T}}}$ when $\theta_{\mathbf{T}} = \theta_0$, to conclude that $\frac{2n}{\phi''(1)}\widehat{D_\phi}(\theta_0, \theta_{\mathbf{T}})$ converges in distribution to a χ^2 variable with p degrees of freedom when $\theta_{\mathbf{T}} = \theta_0$.

(2) Assume that $\theta_{\mathbf{T}} \neq \theta_0$, using Taylor expansion again of

$$\widehat{D_\phi}(\theta_{\mathbf{T}}, \theta_0) = \int_{\mathbf{I}} \mathbf{m}(\widehat{\theta}_n, \mathbf{u})\,\mathrm{d}\mathbf{C}_n(\mathbf{u})$$

in $\widehat{\theta}_n$ around $\theta_{\mathbf{T}}$, combined with the fact that

$$\int_{\mathbf{I}} \frac{\partial}{\partial \theta}\mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\,\mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u}) = 0,$$

we obtain from part (2) of Theorem 3.2

$$\int_{\mathbf{I}} \mathbf{m}(\widehat{\theta}_n, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) = \int_{\mathbf{I}} \mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) + o_P(n^{-1/2}).$$

Hence,

$$\begin{aligned} &\sqrt{n}\left(\widehat{D_\phi}(\theta_0, \theta_{\mathbf{T}}) - D_\phi(\theta_0, \theta_{\mathbf{T}})\right) \\ &= \sqrt{n}\left(\int_{\mathbf{I}} \mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\, \mathrm{d}\mathbf{C}_n(\mathbf{u}) - \int_{\mathbf{I}} \mathbf{m}(\theta_{\mathbf{T}}, \mathbf{u})\, \mathrm{d}\mathbf{C}_{\theta_{\mathbf{T}}}(\mathbf{u})\right) + o_P(1), \end{aligned}$$

which under assumption (C.2.iii) by Theorem 2.1 in [35] once more, converges to a centred normal variable with variance given in (4.2). □

ACKNOWLEDGEMENT

We would like to thank the Editor, an Associate editor, and the referees for their constructive criticism and helpful comments. We are also grateful to Professors Michel Broniatowski and Paul Deheuvels for their helpful discussions and suggestions leading to improvement of this paper.

(Received June 26, 2008)

REFERENCES

[1] M. Broniatowski and A. Keziou: Minimization of ϕ-divergences on sets of signed measures. Studia Sci. Math. Hungar. *43* (2006), 4, 403–442.

[2] S. Bouzebda and A. Keziou: A test of independence in some copula models. Math. Methods Statist. *17* (2008), 2, 123–137.

[3] M. Broniatowski and A. Keziou: Parametric estimation and tests through divergences and the duality technique. J. Multivariate Anal. *100* (2009), 1, 16–36.

[4] S. Bouzebda and S. Keziou: A new test procedure of independence in copula models via χ^2-divergence. Comm. Statist. Theory Methods *38* (2009), 20, to appear.

[5] N. Cressie and T. R. C. Read: Multinomial goodness-of-fit tests. J. Roy. Statist. Soc. Ser. B,*46* (1984), 3, 440–464.

[6] P. Deheuvels: La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. Acad. Roy. Belg. Bull. Cl. Sci. (5), *65* (1979), 6, 274–292.

[7] P. Deheuvels: Propriétés d'existence et propriétés topologiques des fonctions de dépendance avec applications à la convergence des types pour des lois multivariées. C. R. Acad. Sci. Paris Sér. A-B *288* (1979), 2, A145–A148.

[8] P. Deheuvels: Nonparametric test of independence. In: Proc. Nonparametric Asymptotic Statistics, Rouen 1979 (Lecture Notes in Mathematics 821). Springer, Berlin 1980, pp. 95–107.

[9] P. Deheuvels: A Kolmogorov-Smirnov type test for independence and multivariate samples. Rev. Roumaine Math. Pures Appl. *26* (1981), 2, 213–226.

[10] P. Deheuvels: A multivariate Bahadur–Kiefer representation for the empirical copula process. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) *364* (2009), 120–147.

[11] Z. Feng and C. E. McCulloch: Statistical inference using maximum likelihood estimation and the generalized likelihood ratio when the true parameter is on the boundary of the parameter space. Statist. Probab. Lett. *13* (1992), 4, 325–332.

[12] J.-D. Fermanian, D. Radulović, and M. Wegkamp: Weak convergence of empirical copula processes. Bernoulli *10* (2004), 5, 847–860.

[13] P. Gaenssler and W. Stute: Seminar on empirical processes. (Notes based on lectures given at the Seminar on Empirical Processes held at Schloß Mickeln, Düsseldorf, FRG, September 8–13, 1985). DMV Seminar, Bd. 9. Birkhäuser Verlag, Basel – Boston 1987.

[14] J. Galambos: Order statistics of samples from multivariate distributions. J. Amer. Statist. Assoc. *70* (1975) (351, part 1), 674–680.

[15] C. Genest, K. Ghoudi, and L.-P. Rivest: A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika *82* (1995), 3, 543–552.

[16] E. J. Gumbel: Bivariate exponential distributions. J. Amer. Statist. Assoc. *55* (1960), 698–707.

[17] J. Hüsler and R.-D. Reiss: Maxima of normal random vectors: between independence and complete dependence. Statist. Probab. Lett. *7* (1989), 4, 283–286.

[18] H. Joe: Parametric families of multivariate distributions with given margins. J. Multivariate Anal. *46* (1993), 2, 262–282.

[19] H. Joe: Multivariate Models and Dependence Concepts (Monographs on Statistics and Applied Probability, vol. 73). Chapman & Hall, London 1997.

[20] A. Keziou: Dual representation of ϕ-divergences and applications. C. R. Math. Acad. Sci. Paris *336* (2003), 10, 857–862.

[21] A. Keziou and S. Leoni-Aubin: On empirical likelihood for semiparametric two-sample density ratio models. J. Statist. Plann. Inference *138* (2008), 4, 915–928.

[22] G. Kimeldorf and A. Sampson: One-parameter families of bivariate distributions with fixed marginals. Comm. Statist. *4* (1975), 293–301.

[23] G. Kimeldorf and A. Sampson: Uniform representations of bivariate distributions. Comm. Statist. *4* (1975), 7, 617–627.

[24] K.-Y. Liang and S. G. Self: On the asymptotic behaviour of the pseudolikelihood ratio test statistic. J. Roy. Statist. Soc. Ser. B *58* (1996), 4, 785–796.

[25] F. Liese and I. Vajda: Convex Statistical Distances. Teubner, Leipzig 1987.

[26] F. Liese and I. Vajda: On divergences and informations in statistics and information theory. IEEE Trans. Inform. Theory *52* (2006), 10, 4394–4412.

[27] B. V. M. Mendes, E. F. L. de Melo, and R. B. Nelsen: Robust fits for copula models. Comm. Statist. Simulation Comput. *36* (2007), 4–6, 997–1017.

[28] D. S. Moore and M. C. Spruill: Unified large-sample theory of general chi-squared statistics for tests of fit. Ann. Statist. *3* (1975), 599–616.

[29] R. B. Nelsen: An Introduction to Copulas. (Lecture Notes in Statistics, vol. 139.) Springer-Verlag, New York 1999.

[30] D. Oakes: Multivariate survival distributions. J. Nonparametr. Statist. *3* (1994), 3–4, 343–354.

[31] R. T. Rockafellar: Convex Analysis. Princeton University Press, Princeton, N.J. 1970.

[32] L. Rüschendorf: Asymptotic distributions of multivariate rank order statistics. Ann. Statist. *4* (1976), 5, 912–923.

[33] L. Rüschendorf: On the distributional transform, Sklar's theorem, and the empirical copula process. J. Statist. Plann. Inference *139* (2009), 11, 3921–3927.

[34] F. H. Ruymgaart: Asymptotic normality of nonparametric tests for independence. Ann. Statist. *2* (1974), 892–910.

[35] F. H. Ruymgaart, G. R. Shorack, and W. R. van Zwet: Asymptotic normality of nonparametric tests for independence. Ann. Math. Statist. *43* (1972), 1122–1135.

[36] B. Schweizer: Thirty years of copulas. In: Advances in probability distributions with given marginals, Rome 1990, vol. 67, Math. Appl., pp. 13–50. Kluwer Academic Publishers, Dordrecht 1991.

[37] S. G. Self and K.-Y. Liang: Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J. Amer. Statist. Assoc. *82* (1987), 398, 605–610.

[38] J. H. Shih and T. A. Louis: Inferences on the association parameter in copula models for bivariate survival data. Biometrics *51* (1995), 4, 1384–1399.

[39] A. Sklar: Random variables, joint distribution functions, and copulas. Kybernetika *9* (1973), 449–460.

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

[41] W. Stute: The oscillation behavior of empirical processes: The multivariate case. Ann. Probab. *12* (1984), 361–379.

[42] H. Tsukahara: Semiparametric estimation in copula models. Canad. J. Statist. *33* (2005), 3, 357–375.

[43] W. Wang and A. A. Ding: On assessing the association for bivariate current status data. Biometrika *87* (2000), 4, 879–893.

Salim Bouzebda, LSTA-Université Paris 6, 175, rue du Chevaleret, Boîte 158, 75013 Paris. France.

e-mail: salim.bouzebda@upmc.fr

Amor Keziou, Laboratoire de Mathématiques (FRE 3111) CNRS, Université de Reims and LSTA-Université Paris 6. UFR Sciences, Moulin de la Housse, B.P. 1039, 51687 Reims. France.

e-mail: amor.keziou@upmc.fr