Estimation and tests of independence in copula models via divergences

Salim BOUZEBDA\textsuperscript{1,2} and Amor KEZIOU\textsuperscript{1}

L.S.T.A., Université Pierre et Marie Curie-Paris 6, 175, rue du Chevaleret, 8\textsuperscript{ème} étage, bâtiment A, 75013 PARIS\textsuperscript{3}

Transport and Traffic Engineering Laboratory
INRETS - ENTEPE, 25 avenue F. Mitterrand
69675 Bron Cedex\textsuperscript{2}

Abstract. We introduce new estimates and tests of independence in copula models with unknown margins using $\phi$-divergences and the duality technique. The asymptotic laws of the estimates and the test statistics are established both when the parameter is an interior point or not.

1 Introduction

Let $F(x_1, x_2) := P(X_1 \leq x_1, X_2 \leq x_2)$ be a 2-dimensional distribution function, and $F_i(x_i) := P(X_i \leq x_i)$, $i = 1, 2$ the marginal distribution of $F(\cdot, \cdot)$. It is well known since the work of [14] that there exists a distribution function $C(\cdot, \cdot)$ on $[0, 1]^2$ with uniform marginals such that

$$C(u_1, u_2) := P\{F_1(X_1) \leq u_1, F_2(X_2) \leq u_2\}.$$

For a systematic theory of copula, see [4, 5, 6] and [11]. Many useful multivariate models for dependence between $X_1$ and $X_2$ turn out to be generated by parametric families of copulas of the form $\{C_\theta; \theta \in \Theta\}$, typically indexed by a vector valued parameter $\theta \in \Theta \subseteq \mathbb{R}^d$ (see, e.g., [11] and [8] among others). In the sequel, we assume that $C_\theta(\cdot, \cdot)$ admits a density $c_\theta(\cdot, \cdot)$ with respect to the Lebesgue measure $\lambda$ on $\mathbb{R}^2$. In this paper, we consider the estimation and test problems for semiparametric copula models with unknown general margins. Let $(X_{1k}, X_{2k})$, $k = 1, \ldots, n$ be a bivariate sample with distribution function $F_{\theta_T}(\cdot, \cdot) = C_{\theta_T}(F_1(\cdot), F_2(\cdot))$ where $\theta_T \in \Theta$ is used to denote the true unknown value of the parameter. In order to estimate $\theta_T$, some semiparametric estimation procedures, based on the maximization, on the parameter space $\Theta$, of properly chosen pseudo-likelihood criterion, have been proposed by [12], [16] and [15] among others. In each of these papers, some asymptotic normality properties are established for $\sqrt{n}(\hat{\theta} - \theta_T)$, where $\hat{\theta} = \hat{\theta}_n$ denotes a properly chosen estimator of $\theta_T$. This is achieved, provided that $\theta_T$ lies in the interior, denoted by $\Theta$, of the parameter space $\Theta \subseteq \mathbb{R}^d$. On the other hand, the case where $\theta_T \in \partial \Theta := \partial \Theta - \Theta$ is a boundary value of $\Theta$, has not been studied in a systematic way until present. We find in [9] many examples of parametric copulas, for which marginal independence is verified for some specific values of the parameter $\theta$, on the boundary $\partial \Theta$ of the admissible parameter set $\Theta \subseteq \mathbb{R}^d$, $d \geq 1$. In the sequel, we denote by $\theta_0$ the value of the parameter which corresponds to the independence. Moreover, it turns out that, for the above-mentioned estimators, the asymptotic normality of $\sqrt{n}(\hat{\theta} - \theta_T)$, may fail to hold for $\theta_T \in \partial \Theta$. 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 $\phi$-divergences holds, even under the independence assumption, when, either, $\theta_0$ is an interior, or a boundary point of $\Theta$. The proposed test statistics of independence using $\phi$-divergences are also studied, under the null hypothesis $\mathcal{H}_0$ of independence, as well as under the alternative hypothesis $\mathcal{H}_1$. 

1
2 A new inference procedure

Recall that the φ-divergences between a bounded signed measure Q, and a probability P on \( \mathcal{D} \), when Q is absolutely continuous with respect to P, is defined by

\[
D_\phi(Q, P) := \int_{\mathcal{D}} \phi \left( \frac{dQ}{dP} \right) dP,
\]

where \( \phi \) is a proper closed convex function from \( ]-\infty, \infty[ \) to \( ]0, \infty[ \) with \( \phi(1) = 0 \) and such that the domain \( \text{dom}\phi := \{ x \in \mathbb{R} : \phi(x) < \infty \} \) is an interval with end points \( a_\phi < b_\phi \). The Kullback-Leibler, modified Kullback-Leibler, \( \chi^2 \), modified \( \chi^2 \), Hellinger and \( L^1 \) divergences are examples of \( \phi \)-divergences; they are obtained respectively for \( \phi(x) = x \log x - x + 1 \), \( \phi(x) = -\log x + x - 1 \), \( \phi(x) = \frac{1}{2}(x-1)^2 \), \( \phi(x) = 2(\sqrt{x}-1)^2 \) and \( \phi(x) = |x-1| \).

We refer to [10] for a systematic theory of divergences. In the sequel, for all \( \theta \), we denote by \( D_\phi(\theta, \theta_T) \) the \( \phi \)-divergences between \( C_\theta(\cdot, \cdot) \) and \( C_{\theta_T}(\cdot, \cdot) \), i.e.,

\[
D_\phi(\theta, \theta_T) := \int I \phi \left( \frac{dC_\theta}{dC_{\theta_T}} \right) dC_{\theta_T}(u_1, u_2) = \int I \phi \left( \frac{c_\theta}{c_{\theta_T}} \right) dC_{\theta_T}(u_1, u_2). \tag{1}
\]

Denote \( C_n(\cdot, \cdot) \) the empirical copula associated to the data, i.e.,

\[
C_n(u_1, u_2) := \frac{1}{n} \sum_{k=1}^{n} I \{ F_{1n}(X_{1k}) \leq u_1 \} I \{ F_{2n}(X_{2k}) \leq u_2 \}, \quad (u_1, u_2) \in I, \tag{2}
\]

and \( F_{jn}(t) := \frac{1}{n} \sum_{k=1}^{n} I \{ X_{jk} \leq t \} \), \( j = 1, 2 \). In order to estimate the divergences \( D_\phi(\theta, \theta_T) \) for a given \( \theta \in \Theta \) in particular for \( \theta = \theta_0 \), and the parameter \( \theta_T \), we will make use of the dual representation of \( \phi \)-divergences obtained by [3] Theorem 4.4. By this, we readily obtain that \( D_\phi(\theta_0, \theta_T) \) can be rewritten into

\[
D_\phi(\theta_0, \theta_T) := \sup_{f \in \mathcal{F}} \left\{ \int I f dC_{\theta_0} - \int I \phi^*(f) dC_{\theta_T} \right\}, \tag{3}
\]

where \( \phi^* \) is used to denote the convex conjugate of \( \phi \), namely, the function defined by

\[
\phi^* : t \in \mathbb{R} \mapsto \phi^*(t) := \sup_{x \in \mathbb{R}} \{ tx - \phi(x) \},
\]

and \( \mathcal{F} \) is an arbitrary class of measurable functions fulfilling the following two conditions : \( \forall f \in \mathcal{F}, \int |f| dC_{\theta_0} \) is finite and \( \phi^*(dC_{\theta_0}/dC_{\theta_T}) \in \mathcal{F} \). Furthermore, the sup in the above display is unique and is achieved at \( f = \phi^*(dC_{\theta_0}/dC_{\theta_T}) \). Note that the plug-in estimate \( \int I \phi(dC_{\theta_0}/dC_n) dC_n(u_1, u_2) \) of \( D_\phi(\theta, \theta_T) \) is not well defined since \( C_\theta(\cdot, \cdot) \) is not absolutely continuous with respect to \( C_n(\cdot, \cdot) \); the use of the dual representation, as we will show, avoids this problem.

By the above statement, taking the class of functions \( \mathcal{F} = \{ u \in I \mapsto \phi \left( 1/c_{\theta_0} \right) ; \theta \in \Theta \} \), we obtain the formula

\[
D_\phi(\theta_0, \theta_T) = \sup_{\theta \in \Theta} \left\{ \int I \phi^* \left( \frac{1}{c_{\theta}} \right) du_1 du_2 - \int I \left[ \frac{1}{c_{\theta}} \phi^* \left( \frac{1}{c_{\theta}} \right) - \phi \left( \frac{1}{c_{\theta}} \right) \right] dC_{\theta_T}(u_1, u_2) \right\}, \tag{4}
\]

whenever \( \int I |\phi^*(1/c_{\theta_0})| du_1 du_2 \) is finite for all \( \theta \in \Theta \). Furthermore, the sup is unique and reached at \( \theta = \theta_T \). Hence, the divergence \( D_\phi(\theta_0, \theta_T) \) and the parameter \( \theta_T \) can be estimated respectively by

\[
\sup_{\theta \in \Theta} \left\{ \int I \phi^* \left( \frac{1}{c_{\theta}} \right) du_1 du_2 - \int I \left[ \frac{1}{c_{\theta}} \phi^* \left( \frac{1}{c_{\theta}} \right) - \phi \left( \frac{1}{c_{\theta}} \right) \right] dC_n(u_1, u_2) \right\}. \tag{5}
\]
and
\[
\text{arg sup}_{\theta \in \Theta} \left\{ \int_I \phi' \left( \frac{1}{c_0} \right) \, du_1 du_2 - \int_I \left[ \frac{1}{c_0} \phi' \left( \frac{1}{c_0} \right) - \phi \left( \frac{1}{c_0} \right) \right] \, dC_n(u_1, u_2) \right\},
\]

(6)
in which \(C_{\theta_T}\) is replaced by \(C_n\). Note that this class of estimates contains the maximum pseudo-likelihood (MPL) estimator proposed by \cite{12}; it is obtained for the \(KL_m\)-divergence taking \(\phi(x) = -\log(x) + x - 1\). The results in \cite{1} show that, for \(\Theta = [\theta_0, \infty)\), and when the true value \(\theta_T\) of the parameter is equal to \(\theta_0\) the classical asymptotic normality property of the MPL estimate is no more satisfied. To circumvent this difficulty, in what follows, we enlarge the parameter space \(\Theta\) into a wider space \(\Theta_e \supset \Theta\). This is tailored to let \(\theta_0\) become an interior point of \(\Theta_e\). More precisely, set
\[
\Theta_e := \left\{ \theta \in \mathbb{R}^d \text{ such that } \int \phi'(1/c_0(u_1, u_2)) \, du_1 du_2 < \infty \right\}.
\]

(7)
Assume that \(\hat{\Theta}_e\) is non empty set and \(\Theta_e \supset \Theta\). So, applying (3), with the class of functions
\[
\mathcal{F} := \{ (u_1, u_2) \mapsto \phi'(1/c_0(u_1, u_2)); \theta \in \Theta_e \},
\]
we obtain
\[
D_\phi(\theta_0, \theta_T) = \sup_{\theta \in \Theta_e} \left\{ \int_I \phi' \left( \frac{1}{c_0} \right) \, du_1 du_2 - \int_I \left[ \frac{1}{c_0} \phi' \left( \frac{1}{c_0} \right) - \phi \left( \frac{1}{c_0} \right) \right] \, dC_{\theta_T}(u_1, u_2) \right\},
\]

(8)
Furthermore, the sup in this display is unique and reached in \(\theta = \theta_T\). Hence, we propose to estimate \(D_\phi(\theta_0, \theta_T)\) by
\[
\hat{D}_\phi(\theta_0, \theta_T) := \sup_{\theta \in \Theta_e} \int I m(\theta, u_1, u_2) \, dC_n(u_1, u_2),
\]

(9)
and to estimate the parameter \(\theta_T\) by
\[
\hat{\theta}_n := \text{arg sup}_{\theta \in \Theta_e} \left\{ \int_I m(\theta, u_1, u_2) \, dC_n(u_1, u_2) \right\},
\]

(10)
where \(m(\theta, u_1, u_2)\) is equal to
\[
\int I \phi' \left( \frac{1}{c_0(u_1, u_2)} \right) \, du_1 du_2 - \left\{ \phi' \left( \frac{1}{c_0(u_1, u_2)} \right) \frac{1}{c_0(u_1, u_2)} - \phi \left( \frac{1}{c_0(u_1, u_2)} \right) \right\}.
\]

In the sequel, we denote by \(\frac{\partial \phi}{\partial \theta} m(\theta, u_1, u_2)\) the \(d\)-dimensional vector with entries \(\frac{\partial \phi}{\partial \theta} m(\theta, u_1, u_2)\) and \(\frac{\partial^2 \phi}{\partial \theta^2} m(\theta, u_1, u_2)\) the \(d \times d\)-matrix with entries \(\frac{\partial^2 \phi}{\partial \theta^2} m(\theta, u_1, u_2)\).

3 The asymptotic behavior of the estimates

In this section, we provide both weak and strong consistency of the estimates (10). We also state their asymptotic normality and evaluate their limiting variance. We will use the following notations
\[
K_1(\theta, u_1, u_2) := \phi' \left( \frac{1}{c_0(u_1, u_2)} \right)
\]
and
\[
K_2(\theta, u_1, u_2) := \left\{ \phi' \left( \frac{1}{c_0(u_1, u_2)} \right) \frac{1}{c_0(u_1, u_2)} - \phi \left( \frac{1}{c_0(u_1, u_2)} \right) \right\}.
\]
Let \(\mathcal{Q}\) be the set of u-shaped functions, and \(\mathcal{R}\) the set of reproducing u-shaped functions (see e.g. \cite{13} p. 894 for definition). We make use of the following conditions.
There exists a neighborhood \( N(\theta_T) \subset \Theta_e \) of \( \theta_T \) such that the first and the second partial derivatives with respect to \( \theta \) of \( K_1(\theta, u_1, u_2) \) are dominated on \( N(\theta_T) \) by some \( \lambda \)-integrable functions;

There exists a neighborhood \( N(\theta_T) \) of \( \theta_T \), such that for all \( \theta \in N(\theta_T) \), the function \( \frac{\partial}{\partial \theta} m(\theta, u_1, u_2) : (0, 1)^2 \to \mathbb{R} \) is continuously differentiable and there exist functions \( r_i \in \mathcal{R}, \tilde{r}_i \in \mathcal{R} \) and \( q \in \mathcal{L} \) (\( i, j = 1, 2, \ i \neq j \) and \( \ell, \ell' = 1, \ldots, m \)) with

\[
\begin{align*}
(i) & \quad \left| \frac{\partial}{\partial \theta} m(\theta, u_1, u_2) \right| \leq r_1(u_1) r_2(u_2), \quad \left| \frac{\partial^2}{\partial \theta \partial u} m(\theta, u_1, u_2) \right| \leq \tilde{r}_i(u_i) r_j(u_j); \\
(ii) & \quad \left| m^2(\theta, u_1, u_2) \right| \leq \tilde{r}_i(u_i) r_j(u_j); \\
(iii) & \quad \left| \frac{\partial^3}{\partial \theta} K_2(\theta, u_1, u_2) \right| \leq \tilde{r}_i(u_i) r_j(u_j); \\
(iv) & \quad \left| m(\theta, u_1, u_2) \right| \leq \tilde{r}_i(u_i) r_j(u_j); \\
(v) & \quad \left| \frac{\partial^2}{\partial \theta^2} m(\theta, u_1, u_2) \right| \leq \tilde{r}_i(u_i) r_j(u_j); \\
(vi) & \quad \left| \frac{\partial^3}{\partial \theta^2} m(\theta, u_1, u_2) \right| \leq \tilde{r}_i(u_i) r_j(u_j)
\end{align*}
\]

\( \text{and } \int_1 \{r_1(u_1) r_2(u_2)\}^2 dC_{\theta_T}(u_1, u_2) < \infty, \int_1 \{q_i(u_i) \tilde{r}_i(u_i) r_j(u_j)\} dC_{\theta_T}(u_1, u_2) < \infty; \)

The matrix \( \int (\partial^2 / \partial \theta) m(\theta, u_1, u_2) dC_{\theta_T}(u_1, u_2) \) is non-singular;

The function \( (u_1, u_2) \in I \mapsto \frac{\partial}{\partial \theta} m(\theta, u_1, u_2) \) is of bounded variation on \( I \).

**Theorem 3.1** Assume that conditions (C.1-3) hold.

1. Let \( B(\theta_T, n^{-1/3}) := \{ \theta \in \Theta_e : |\theta - \theta_T| \leq n^{-1/3} \} \), then as \( n \) tends to infinity, with probability one, the function \( \theta \mapsto \int m(\theta, u_1, u_2) dC_n(u_1, u_2) \) attains its maximum value at some point \( \hat{\theta}_n \) in the interior of \( B(\theta_T, n^{-1/3}) \), which implies that the estimate \( \hat{\theta}_n \) is consistent almost surely and satisfies \( \int \frac{\partial}{\partial \theta} m(\hat{\theta}_n, u_1, u_2) dC_n(u_1, u_2) = 0 \).

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

\[
\Xi_{\phi} = S^{-1} MS^{-1},
\]

where

\[
S := -\int \frac{\partial^2}{\partial \theta^2} m(\theta_T, u_1, u_2) dC_{\theta_T}(u_1, u_2),
\]

and

\[
M := \text{Var} \left[ \frac{\partial}{\partial \theta} m(\theta_T, F_1(X_1), F_2(X_2)) + W_1(\theta_T, X_1) + W_2(\theta_T, X_2) \right],
\]

where

\[
W_i(\theta_T, X_i) := \int_1 \{F_i(X_i) \leq u_i\} \frac{\partial^2}{\partial \theta^2} m(\theta_T, u_1, u_2) c_{\theta_T}(u_1, u_2) du_1 du_2, \ i = 1, 2.
\]

**4 New tests of independence**

In the framework of the parametric copula model, the null hypothesis, i.e., the independence case \( C_0(u_1, u_2) = u_1 u_2 \) corresponds to \( \mathcal{H}_0 : \theta_T = \theta_0 \). We consider the composite alternative hypothesis \( \mathcal{H}_1 : \theta_T \neq \theta_0 \). Since \( \theta_0 \) is a boundary value of the parameter space \( \Theta \), we can see that the convergence in distribution of the corresponding pseudo-likelihood ratio statistic to a \( \chi^2 \) random variable does not hold; see [1]. We give now a solution to this problem. We propose the following statistics

\[
T_n := \frac{2n}{\phi'^{(1)}(1)} \hat{D}_{\phi}(\theta_0, \theta_T).
\]

We will use the following additional conditions
(C.5) We have
\[
\lim_{\theta \to \theta_0} \frac{\partial^2}{\partial \theta_i \partial u_i} m(\theta, u_1, u_2) = 0,
\]
and there exist \( M_1 > 0 \) and \( \delta_1 > 0 \) such that, for all \( \theta \) in some neighborhood of \( \theta_0 \), one has, for \( i = 1, 2, \)
\[
\left| \frac{\partial^2}{\partial \theta_i \partial u_i} m(\theta, u_1, u_2) c_{\theta_T}(u_1, u_2) \right| < M_1 r(u_i)^{-1.5+\delta_1} r(u_{3-i})^{0.5+\delta_1},
\]
where \( r(u) := u(1-u) \) for \( u \in (0,1) \).

**Theorem 4.1**

1. Assume that conditions \((C.1-5)\) hold. If \( \theta_T = \theta_0 \), then the statistic \( T_n \) converges in distribution to a \( \chi^2 \) variable with \( d \) degrees of freedom.
2. Assume that conditions \((C.1-4)\) hold. If \( \theta_T \neq \theta_0 \), then \( \sqrt{n} \left( \hat{D}_\phi(\theta_0, \theta_T) - D_\phi(\theta_0, \theta_T) \right) \) converges in distribution to a centered normal variable with variance
\[
\sigma^2_\phi(\theta_0, \theta_T) := \text{Var} \left[ m(\theta_T, F_1(X_1), F_2(X_2)) + Y_1(\theta_T, X_1) + Y_2(\theta_T, X_2) \right],
\]
where
\[
Y_i(\theta_T, X_i) := \int 1_{(F_i(X_i) \leq u_i)} \frac{\partial}{\partial u_i} m(\theta_T, u_1, u_2) c_{\theta_T}(u_1, u_2) \, du_1 du_2, \quad i = 1, 2.
\]

**Remark 1** The above regularity conditions are satisfied by a large number of parametric families of bivariate copulas; see for instance [15].

**Remark 2** The parameters \((11)\) and \((11)\) may be consistently estimated respectively by the sample mean of
\[
\frac{\partial^2}{\partial \theta^2} m(\hat{\theta}_n, F_1n(X_{1,k}), F_2n(X_{2,k})), \quad k = 1, \ldots, n,
\]
and the sample variance of
\[
\frac{\partial}{\partial \theta} m(\hat{\theta}_n, F_1n(X_{1,k}), F_2n(X_{2,k})) + W_1(\hat{\theta}_n, X_{1,k}) + W_2(\hat{\theta}_n, X_{2,k}), \quad k = 1, \ldots, n,
\]
as was done in [7]. The asymptotic variance \((12)\) can be consistently estimated in the same way.

5 Concluding remarks

We have introduced a new estimation and test procedure in parametric copula models with unknown margins. The methods 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. It will be interesting to investigate 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.

Références


