



# A semiparametric estimation procedure for multi-parameter Archimedean copulas based on the L-moments method

Fateh Benatia, Brahim Brahimi and Abdelhakim Necir

Laboratory of Applied Mathematics, Mohamed Khider University of Biskra, 07000, Algeria

Received 22 April 2011; Accepted 07 October 2011

Copyright © 2011, Journal Afrika Statistika. All rights reserved

**Abstract.** A new semiparametric estimation method for multi-parameters Archimedean copulas based on the L-moments theory is proposed. Consistency and asymptotic normality of the defined estimator are established. Extensive simulation study to compare estimators based on the L-moments, the maximum likelihood and the measures of concordance is carried out. We concluded that this method is quick and does not use the density function and therefore no boundary problems arise.

**Résumé.** Une nouvelle méthode d'estimation semi-paramétrique pour les copules Archimédenne a plusieurs paramètres basée sur la théorie des L-moments est proposée. La consistance et la normalité asymptotique de l'estimateur sont établies. Une étude de simulation, pour comparer les estimateurs basés sur les L-moments, le maximum de vraisemblance et les mesures de concordance, est effectuée. On conclut que cette méthode est rapide et ne utilise pas la fonction de densité et par conséquent les problèmes de frontière ne se posent pas.

**Key words:** L-moments; Copulas; Dependence; Concordance measures; Semiparametric estimation.

**AMS 2010 Mathematics Subject Classification :** 78M05, 62F12, 62H12, 62H20.

---

## 1. Introduction

Let  $\mathbf{X} = (X_1, \dots, X_d)$  be a  $d$ -dimensional random vector ( $d \geq 2$ ) defined on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with distribution function  $F(\mathbf{x}) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d)$  for  $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$  and marginal distribution functions  $F_j(x) = \mathbb{P}(X_j \leq x)$  for  $x \in \mathbb{R}$ ,  $j = 1, \dots, d$ .

---

Fateh Benatia: [fatahbenatia@hotmail.com](mailto:fatahbenatia@hotmail.com)

Brahim Brahimi: [brah.brahim@gmail.com](mailto:brah.brahim@gmail.com)

Abdelhakim Necir: [necirabdelhakim@yahoo.fr](mailto:necirabdelhakim@yahoo.fr)

If not stated otherwise, we always assume that the  $F_j$  are continuous functions. According to Sklar's Theorem (Sklar [12]), there exists a unique copula  $C : [0, 1]^d \rightarrow [0, 1]$  such that

$$F(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d)), \text{ for all } \mathbf{x} \in \mathbb{R}^d.$$

On the other terms, the copula  $C$  is the joint df of the random vector  $\mathbf{U} = (U_1, \dots, U_d)$ , with  $U_j = F_j(X^{(j)})$ ,  $j = 1, \dots, d$ , and is defined for all  $\mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d$ , by

$$C(\mathbf{u}) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)),$$

where  $F_j^{-1}$  is the generalized inverse function (or the quantile function) of the df  $F_j$ . For more details on copula theory, see for instance Nelsen [10]. In this paper we are dealing with an important class of copulas called the Archimedean family defined by

$$C(\mathbf{u}) = \varphi^{-1} \left( \sum_{j=1}^d \varphi(u_j) \right),$$

where  $\varphi : [0, 1] \rightarrow \mathbb{R}$  is a twice differentiable function called the generator, satisfying:  $\varphi(1) = 0$ ,  $\varphi'(x) < 0$  and  $\varphi''(x) \geq 0$  for any  $x \in (0, 1)$ . The notation  $\varphi^{-1}$  stands for the inverse function of  $\varphi$ . Archimedean copulas are easier to construct and they have nice properties and a variety of known copula families belong to this class including the models of Gumbel, Clayton, Frank, ... (see, Table 4.1 in Nelsen [10, page 116]).

In this paper, we are dealing with the statistical inference on multiparameter Archimedean copula models when the generator  $\varphi$  is unknown but assumed to belong to a class  $\Phi := \{\varphi_\theta, \theta \in \mathcal{O}\}$ , in other terms when the copula  $C$  belongs to a class  $\mathcal{C} := \{C_\theta, \theta \in \mathcal{O}\}$ , where  $\mathcal{O}$  is an open subset of  $\mathbb{R}^r$  for some integer  $r \geq 1$ . The estimation of parameters of copula models is one of main topics in multivariate statistical analysis and several parameter estimation have been developed in the literature. Namely, the methods of concordance, also called  $(\tau, \rho)$ -inversion methods (Oakes [11], Genest [3]), fully maximum likelihood (ML), pseudo maximum likelihood (PML) (Genest *et al.* [4]), inference function of margins (IFM) (Joe [7, 8]), and minimum distance (MD) (Tsukahara [13]). It is shown that PML's procedure performs well compared to the cited methods as proved by several authors. Indeed, Kim *et al.* [9] has concluded that the PML method is conceptually almost the same as the IFM one. The main advantage of PML's method is that, one would not lose any important statistical insights that would be gained by applying the IFM. Moreover, the PML method does not require modeling the marginal distributions explicitly.

In conclusion, the PML estimator is better than those of the ML and IFM in most practical situations. In time-consuming point of view the ML, PML, IFM and MD procedures require intensive computations for high dimensional copula models. It is worth mentioning that, when using these methods, we use the maximum likelihood function, and then the copula density is automatically involved. Therefore a serious boundry problems arise. There are many numerical methods to solve this issue, but they are still inefficient when for high dimensional copula models (see, Yan [16]).

The aim of this paper is to propose an alternative estimation method based on the L-moment theory (see, Hosking [5]) avoiding technical problems caused by copula density and

providing estimators with reasonable time-consuming, bias and root mean squared error (RMSE). The  $L$ -moments based estimation have been used as an alternative to the classical central moments  $\mu_k = \mathbf{E} \left[ (Y - \mu)^k \right]$  determined by the df  $F_Y$  of the underlying rv  $Y$ . An  $L$ -moment  $\lambda_k$  is defined as a specific linear combination of the expectations of the order statistics  $Y_{1:k} \leq \dots \leq Y_{k:k}$ . More precisely, the  $k$ th  $L$ -moment is defined by

$$\lambda_k = \frac{1}{k} \sum_{\ell=0}^{k-1} \frac{(-1)^\ell (k-1)!}{\ell! (k-1-\ell)!} \mathbf{E} [Y_{k-\ell:k}], k = 1, 2, \dots$$

By analogy with the classical moments, the first four  $L$ -moments  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  measure, respectively, location, scale, skewness and kurtosis features. The functional representation of  $\lambda_k$  is terms of the df  $F_Y$  is given by (see, Hosking [6]):

$$\lambda_k = \mathbf{E} [Y \mathbf{P}_{k-1} (F_Y (Y))] = \int_{\mathbb{R}} y \mathbf{P}_{k-1} (F_Y (y)) dF_Y (y), \tag{1}$$

where  $\mathbf{P}_k (u) := \sum_{\ell=0}^k p_{k,\ell} u^\ell$ , with  $p_{k,\ell} = (-1)^{k+\ell} (k+\ell)! / [(\ell^2)! (k-\ell)!]$  is the shifted Legendre polynomial. In particular, we have,  $\mathbf{P}_0 (u) = 1$ ,  $\mathbf{P}_1 (u) = 2u - 1$ ,  $\mathbf{P}_2 (u) = 6u^2 - 6u + 1$ . Let  $\mathbf{K}_C (s) := \mathbb{P} (C (\mathbf{U}) \leq s)$   $s \in [0, 1]$  of rv  $C (\mathbf{U})$ , then by the representation 1, the corresponding  $k$ th  $L$ -moment of the rv  $C (\mathbf{U})$  is

$$\begin{aligned} \lambda_k (C) &:= \mathbf{E} [C (\mathbf{U}) \mathbf{P}_{k-1} (\mathbf{K}_C (C (\mathbf{U})))] \\ &= \int_{[0,1]^d} C (\mathbf{u}) \mathbf{P}_{k-1} (\mathbf{K}_C (C (\mathbf{u}))) dC (\mathbf{u}), k = 1, 2, \dots \end{aligned}$$

This also may be rewritten into

$$\lambda_k (C) = \int_0^1 s \mathbf{P}_{k-1} (\mathbf{K}_C (s)) d\mathbf{K}_C (s), k = 1, 2, \dots \tag{2}$$

Suppose now, for unknown  $\theta$ , that  $\varphi = \varphi_\theta$ . Therefore  $C = C_\theta$ ,  $\mathbf{K}_C = \mathbf{K}_\theta$  and  $\lambda_k (C) = \lambda_k (\theta)$ , it follows that

$$\lambda_k (\theta) = \int_0^1 s \mathbf{P}_{k-1} (\mathbf{K}_\theta (s)) d\mathbf{K}_\theta (s), k = 1, 2, \dots$$

One nice property of Archimedean copulas, among others, is that the df of rv  $C (\mathbf{U})$  may be represented in terms of derivative of the generator. Indeed from Theorem 4.3.4 in Nelsen [10, page 127], we have for any  $s \in [0, 1]$ ,  $\mathbf{K}_\theta (s) = s - \varphi_\theta (s) / \varphi'_\theta (s)$ . It follows that the corresponding density is  $\mathbf{K}'_\theta (s) = \varphi''_\theta (s) \varphi_\theta (s) / (\varphi'_\theta (s))^2$ . Therefore the  $k$ th copula-moment may be rewritten into

$$\lambda_k (\theta) = \int_0^1 s \mathbf{P}_{k-1} \left( s - \frac{\varphi_\theta (s)}{\varphi'_\theta (s)} \right) \frac{\varphi''_\theta (s) \varphi_\theta (s)}{(\varphi'_\theta (s))^2} ds, k = 1, 2, \dots \tag{3}$$

In particular, the first two copula  $L$ -moments are

$$\lambda_1(\theta) = \int_0^1 s \frac{\varphi''_{\theta}(s) \varphi_{\theta}(s)}{(\varphi'_{\theta}(s))^2} ds,$$

$$\lambda_2(\theta) = \int_0^1 s \left( 2s - \frac{2\varphi_{\theta}(s)}{\varphi'_{\theta}(s)} - 1 \right) \frac{\varphi''_{\theta}(s) \varphi_{\theta}(s)}{(\varphi'_{\theta}(s))^2} ds.$$

From equations (3), we may consider  $\lambda : \theta \rightarrow \lambda(\theta)$  as a mapping from  $\mathcal{O} \subset \mathbb{R}^r$  to  $\mathbb{R}^r$ , that will be used as a mean to estimate the parameter  $\theta$ . More precisely, for a given sample  $(\mathbf{X}_1, \dots, \mathbf{X}_n)$  of the random vector  $\mathbf{X}$ , let us denote  $\hat{\theta}^{CLM}$  as the estimator of  $\theta$  defined by  $(\lambda_k)_{1 \leq k \leq r}$ . That is

$$\hat{\theta}^{CLM} := \lambda^{-1}(\hat{\lambda}_k), \quad k = 1, \dots, r, \tag{4}$$

where  $\hat{\lambda}_k$  is the empirical version of  $\lambda_k(C)$  and  $\lambda^{-1}$  is the inverse of the mapping  $\lambda$ , provided it exists. The rest of the paper is organized as follows. In Section 2, we present the main steps of the copula L-moment estimation procedure and establish the consistency and asymptotic normality of the proposed estimator. In Section 3, an application to multiparameter Archimedean copula models is given. In Section 4, an extensive simulation study is carried out to evaluate and compare the CLM based estimation with the PML and  $(\tau, \rho)$ -inversion methods. Comments and conclusion are given Section 5. The proofs are relegated to the Appendix.

## 2. Semi-parametric CLM-based estimation

In this section we present a semiparametric estimation procedure for the copula models based on the CLM's (3). First suppose that the underlying copula  $C$  belongs to a parametric family  $C_{\theta}$ , with  $\theta = (\theta_1, \dots, \theta_r)$ , and satisfies the *concordance ordering condition* of copulas (see, Nelsen [10, page 135]), that is:

$$\text{for every } \theta^{(1)}, \theta^{(2)} \in \mathcal{O} : \theta^{(1)} \neq \theta^{(2)} \implies C_{\theta^{(1)}} (> \text{ or } <) C_{\theta^{(2)}}. \tag{5}$$

It is clear that this condition implies the well-known *identifiability condition* of copulas: for every  $\theta_1, \theta_2 \in \mathcal{O} : \theta_1 \neq \theta_2 \implies C_{\theta_1} \neq C_{\theta_2}$ . The identifiability condition is a natural and even a necessary condition, since if the parameter is not identifiable then consistent estimator cannot exist (see, e.g., van der Vaart [15, page 62]). Consider now a random sample  $(\mathbf{X}_1, \dots, \mathbf{X}_n)$  from a  $d$ -dimensional random vector  $\mathbf{X} = (X_1, \dots, X_d)$  with joint distribution function  $F$  and copula  $C$  which are completely unknown. The marginal distribution functions  $F_j$  are estimated by their empirical counterparts  $F_{jn}(x_j) = n^{-1} \sum_{i=1}^n \mathbf{1}\{X_{ji} \leq x_j\}$ ,  $j = 1, \dots, d$ . By letting  $\hat{U}_{ji} := F_{jn}(X_{ji})$  and

$$\hat{\mathbf{U}}_i := (\hat{U}_{1i}, \dots, \hat{U}_{di}), \quad i = 1, \dots, n,$$

the copula  $C$  is estimated by the empirical copula (Deheuvels [2]) which is defined as

$$C_n(\mathbf{u}) = n^{-1} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1}\{\hat{U}_{ji} \leq u_j\}, \quad \text{for } \mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d.$$

The empirical df of the rv  $C(\mathbf{U})$  pertaining to the sample  $C_n(\widehat{\mathbf{U}}_1), \dots, C_n(\widehat{\mathbf{U}}_n)$  is defined by

$$\mathbf{K}_n(t) := n^{-1} \sum_{i=1}^n \mathbf{1} \left( C_n(\widehat{\mathbf{U}}_i) \leq t \right), \quad t \in [0, 1].$$

The empirical estimator of the  $k$ th copula L-moment  $\lambda_k$  may be done by substituting  $\mathbf{K}_n$  in formula (2), that is

$$\begin{aligned} \widehat{\lambda}_k &= \lambda_k(C_n) = \int_0^1 s \mathbf{P}_{k-1}(\mathbf{K}_n(s)) d\mathbf{K}_n(s), \\ &= n^{-1} \sum_{i=1}^n C_n(\widehat{\mathbf{U}}_i) \mathbf{P}_{k-1} \left( \mathbf{K}_n \left( C_n(\widehat{\mathbf{U}}_i) \right) \right), \quad k = 1, 2, \dots \end{aligned}$$

We call the solution  $\widehat{\theta}^{CLM} := (\widehat{\theta}_1, \dots, \widehat{\theta}_r)$  of the following system

$$\lambda_k(\theta_1, \dots, \theta_r) = \widehat{\lambda}_k, \quad k = 1, \dots, r, \tag{6}$$

the CLM estimator for  $\theta$ . Consistency and asymptotic normality of  $\widehat{\theta}^{CLM}$  are stated in Theorem 1 below whose proof is relegated to the appendix. For convenience we set

$$g_k(\mathbf{u}; \theta) := C_\theta(\mathbf{u}) \mathbf{P}_{k-1}(\mathbf{K}_\theta(C_\theta)) - \lambda_k(\theta) \tag{7}$$

and

$$\mathbf{g}(\mathbf{u}; \theta) = (g_1(\mathbf{u}; \theta), \dots, g_r(\mathbf{u}; \theta)).$$

Let  $\theta_0$  be the true value of  $\theta$  and assume that the following assumptions [A1] – [A3] hold.

- [A1]  $\theta_0 \in \mathcal{O} \subset \mathbb{R}^r$  is the unique zero of the mapping  $\theta \rightarrow \int_{[0,1]^d} \mathbf{g}(\mathbf{u}; \theta) dC_{\theta_0}(\mathbf{u})$  which is defined from  $\mathcal{O}$  to  $\mathbb{R}^r$ .
- [A2]  $\mathbf{g}(\cdot; \theta)$  is differentiable with respect to  $\theta$  such that the Jacobian matrix denoted by  $\dot{\mathbf{g}}(\mathbf{u}; \theta) := [\partial g_k(\mathbf{u}; \theta) / \partial \theta_\ell]_{r \times r}$ ,  $\dot{\mathbf{g}}(\mathbf{u}; \theta)$  is continuous both in  $\mathbf{u}$  and  $\theta$ , and the Euclidian norm  $|\dot{\mathbf{g}}(\mathbf{u}; \theta)|$  is dominated by a  $dC_\theta$ -integrable function.
- [A3] The  $r \times r$  matrix  $A_0 := \int_{[0,1]^d} \dot{\mathbf{g}}(\mathbf{u}; \theta_0) dC_{\theta_0}(\mathbf{u})$  is nonsingular.

**Theorem 1.** *Assume that the concordance ordering condition (5) and assumptions [A1] – [A3] hold. Then, there exists a solution  $\widehat{\theta}^{CLM}$  to the system (6) which converges in probability to  $\theta_0$ . Moreover*

$$\sqrt{n} \left( \widehat{\theta}^{CLM} - \theta_0 \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( \mathbf{0}, \mathcal{A}_0^{-1} \mathcal{D}_0 (\mathcal{A}_0^{-1})^T \right), \quad \text{as } n \rightarrow \infty,$$

where

$$\mathcal{D}_0 := \text{var} \{ \mathbf{g}(\xi; \theta_0) + \mathbf{S}(\xi; \theta_0) \}$$

and

$$\mathbf{S}(\xi; \theta_0) = (S_1(\xi; \theta_0), \dots, S_r(\xi; \theta_0)),$$

with

$$S_k(\xi; \theta_0) := \sum_{j=1}^d \int_{[0,1]^d} \frac{\partial(C_\theta(\mathbf{u})\mathbf{P}_{k-1}(\mathbf{K}_\theta(C_\theta)))}{\partial u_j} (\mathbf{1}\{\xi_j \leq u_j\} - u_j) dC_{\theta_0}(\mathbf{u}),$$

for  $k = 1, \dots, r$ , where  $\xi := (\xi_1, \dots, \xi_d)$  is a  $(0, 1)^d$ -uniform random vector with joint df  $C_{\theta_0}$ .

*Remark 1.* Following Genest et al. [4] and Tsukahara [13] in the case of PML’s estimator and Z-estimator respectively, one may consistently estimate the asymptotic variance  $\mathcal{A}_0^{-1}\mathcal{D}_0(\mathcal{A}_0^{-1})^T$  by the sample variance of the sequence of rv’s

$$\left\{ \widehat{\mathcal{A}}_i^{-1} \widehat{\mathcal{D}}_i (\widehat{\mathcal{A}}_i^{-1})^T, i = 1, \dots, n \right\},$$

where

$$\widehat{\mathcal{A}}_i := \int_{[0,1]^d} \dot{\mathbf{g}}(\mathbf{u}; \widehat{\theta}^{CLM}) dC_{\widehat{\theta}^{CLM}}(\mathbf{u})$$

and

$$\widehat{\mathcal{D}}_i := \mathbf{g}(\widehat{\mathbf{U}}_i; \widehat{\theta}^{CLM}) + \mathbf{V}(\widehat{\mathbf{U}}_i; \widehat{\theta}^{CLM}).$$

### 3. Illustrative example

As an illustrative example, we consider the Gumbel family, which is an Archimedean copula, defined by  $C_\beta(\mathbf{u}) = \exp\left(-\left(\sum_{j=1}^d (-\ln u_j)^\beta\right)^{1/\beta}\right)$ ,  $\beta \geq 1$ , with generator  $\varphi_\beta(t) = (-\ln t)^\beta$ ,  $\beta \geq 1$ . In general, it is better to fit high dimensional data with multi-parameters copula models than those of single one. For example, to construct a copula with two parameter from one parameter copula, we may use the transformed copula (Nelsen [10, page 96]) defined by  $C_\Gamma(\mathbf{u}) = \Gamma^{-1}(C(\Gamma(u_1), \dots, \Gamma(u_d)))$ , where  $\Gamma : [0, 1] \rightarrow [0, 1]$  is a continuous, concave and strictly increasing function with  $\Gamma(0) = 0$  and  $\Gamma(1) = 1$ . More precisely, suppose that  $\Gamma = \Gamma_\alpha$ , with  $\Gamma_\alpha(t) = \exp(1 - t^{-\alpha})$ ,  $\alpha > 0$  and the copula  $C_\beta$ , then the transformed copula  $C_{\alpha,\beta}(\mathbf{u}) = \Gamma_\alpha^{-1}(C_\beta(\Gamma_\alpha(u_1), \dots, \Gamma_\alpha(u_d)))$  is defined by

$$C_{\alpha,\beta}(\mathbf{u}) := \left( \left( \sum_{j=1}^d (u_j^{-\alpha} - 1)^\beta \right)^{1/\beta} + 1 \right)^{-1/\alpha}, \text{ for } \mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d, \quad (8)$$

which is also a two-parameter Archimedean copula with generator  $\varphi_{\alpha,\beta}(t) := (t^{-\alpha} - 1)^\beta$ . The copula family  $C_{\alpha,\beta}$  verifies the concordance ordering condition (5) (see, Nelsen [10, page 145]). Calculating the two first  $k$ th CLM we get

$$\begin{cases} \lambda_1(\alpha, \beta) := \frac{2\alpha + \alpha\beta - 1}{4\beta + 2\beta\alpha} \\ \lambda_2(\alpha, \beta) := \frac{2\alpha^3\beta^2 + 13\alpha^2\beta^2 - 2\alpha^2\beta + 27\alpha\beta^2 - \alpha\beta - 4\alpha + 18\beta^2 + 3\beta - 8}{6\beta^2(2\alpha^3 + 13\alpha^2 + 27\alpha + 18)}. \end{cases} \quad (9)$$

For given sample  $(\mathbf{X}_1, \dots, \mathbf{X}_n)$  form the random vector  $\mathbf{X} = (X_1, \dots, X_d)$ , we define the CLM estimator  $(\hat{\alpha}, \hat{\beta})$  of  $(\alpha, \beta)$  as the unique solution of the system

$$\begin{cases} \lambda_1(\alpha, \beta) = \hat{\lambda}_1 \\ \lambda_2(\alpha, \beta) = \hat{\lambda}_2 \end{cases}$$

Of course, in this case, we cannot give explicit formulas, in terms of  $\{\hat{\lambda}_1, \hat{\lambda}_2\}$ , for the parameters  $\{\alpha, \beta\}$ . However for given values of  $\{\hat{\lambda}_1, \hat{\lambda}_2\}$ , we can obtain the corresponding values of  $\{\alpha, \beta\}$  by solving the previous system by numerical methods.

#### 4. Simulation study

It is worth mentioning that, in this paper, all numerical computations are performed by using on a PC with a microprocessor speed of 2.4 GHz. In our simulation study we employ the copula  $C_{\alpha, \beta}$  given in formula (8), to evaluate and compare the performance of CLM's estimator with the PML and  $(\tau, \rho)$ -inversion ones. The evaluation of the performance procedures are based on the bias and the RMSE based on the values of estimator  $\hat{\theta}$  from the  $i$ th samples for  $N$  generated samples from copula  $C_{\alpha, \beta}$ . In both parts, we selected  $N = 1000$ . The procedure outlined in Section 2 is repeated for different sample sizes  $n$  with  $n = 30, 50, 100, 200$  to assess the improvement in the bias and RMSE of the estimators with increasing sample size. In addition, the simulation procedure is repeated for a large set of parameters of the true copula  $C_{\alpha, \beta}$ . For each sample, by using formulas (9), we obtain the CM-estimator  $(\hat{\alpha}_i, \hat{\beta}_i)$  of  $(\alpha, \beta)$  for  $i = 1, \dots, N$ , and the estimators  $\hat{\alpha}$  and  $\hat{\beta}$  are given by  $\hat{\alpha} = \frac{1}{N} \sum_{i=1}^N \hat{\alpha}_i$  and  $\hat{\beta} = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i$ . The choice of the true values of the parameter  $(\alpha, \beta)$  have to be meaningful, in the sense that each couple of parameters assigns a value of one of the dependence measure, that is weak, moderate and strong dependence. In other words, if we consider Kendall's  $\tau$  as a dependence measure, then we should select values for copula parameters that correspond to specified values of  $\tau$  by means of the equation  $\tau(\alpha, \beta) = 4 \int_{[0,1]^2} C_{\alpha, \beta}(u_1, u_2) dC_{\alpha, \beta}(u_1, u_2) - 1$ . The selected values of the true parameters are summarized in table 1.

| $\tau$ | $\alpha$ | $\beta$ |
|--------|----------|---------|
| 0.01   | 0.1      | 1.059   |
| 0.5    | 0.5      | 1.6     |
| 0.8    | 0.9      | 3.45    |

**Table 1.** The true parameters of transformed Gumbel copula used for the simulation study.

#### 5. Conclusion

From Table 5, we conclude that by considering three dependence cases: weak ( $\tau = 0.01$ ), moderate ( $\tau = 0.5$ ) and strong ( $\tau = 0.8$ ), the performance, in terms of bias and RMSE, of

the CM based estimation is well justified. In each case, for small and large samples, the bias and RMSE are sufficiently small. Moreover, in time-consuming point of view, we observe that for a sample size  $n = 30$  and for  $N = 1000$  replications, the central processing unit (CPU) time to process CM's method took 36.609 seconds, which is relatively small. For one replication  $N = 1$ , the CPU time (in seconds) for different sample sizes are summarized as follows:  $(n, CPU) = (30, 0.036), (100, 0.069), (200, 0.202), (500, 6.836)$ . Table 5 shows that both the PML and the CM based estimation perform better than the  $(\tau, \rho)$ -inversion method. However, in weak dependence case  $\tau = 0.01$ , the CM method provides better results than the PML one, mainly when the sample size increases. On the other hand, it is worth mentioning that our method is quick with respect to the PML one. The main advantage of our method is that it provides estimators with explicit forms, as far as Archimedean copula models are concerned. This is not the case of the other methods which require numerical procedures leading to possible problems in processing time and inaccuracy issues. In conclusion, the CM based estimation method performs well for the chosen model. Furthermore, its usefulness in the weak dependence case particularly makes it a good candidate for statistical tests of independence.

### Acknowledgements

The authors are grateful to the anonymous referee for a careful reading of the manuscript and making some useful suggestions.

| $n$ | $\tau = 0.01$  |       |                 |       | $\tau = 0.5$   |       |               |       |
|-----|----------------|-------|-----------------|-------|----------------|-------|---------------|-------|
|     | $\alpha = 0.1$ |       | $\beta = 1.059$ |       | $\alpha = 0.5$ |       | $\beta = 1.6$ |       |
|     | Bias           | RMSE  | Bias            | RMSE  | Bias           | RMSE  | Bias          | RMSE  |
| 30  | 0.156          | 0.450 | 0.103           | 0.296 | -0.138         | 0.874 | 0.962         | 1.909 |
| 50  | 0.097          | 0.305 | 0.052           | 0.182 | -0.014         | 0.608 | 0.326         | 0.664 |
| 100 | 0.048          | 0.194 | 0.023           | 0.106 | -0.024         | 0.386 | 0.161         | 0.361 |
| 200 | 0.010          | 0.130 | 0.010           | 0.071 | 0.015          | 0.268 | 0.059         | 0.226 |
| 500 | 0.005          | 0.074 | 0.007           | 0.038 | 0.012          | 0.162 | 0.012         | 0.083 |

| $n$ | $\tau = 0.8$   |       |                |       |             |
|-----|----------------|-------|----------------|-------|-------------|
|     | $\alpha = 0.9$ |       | $\beta = 3.45$ |       | CPU         |
|     | Bias           | RMSE  | Bias           | RMSE  |             |
| 30  | -0.744         | 1.976 | 1.437          | 2.322 | 36.609 secs |
| 50  | -0.531         | 1.319 | 1.396          | 2.196 | 1.165 mins  |
| 100 | -0.424         | 0.734 | 1.286          | 1.897 | 3.368 mins  |
| 200 | -0.218         | 0.458 | 0.626          | 0.936 | 15.917 mins |
| 500 | -0.087         | 0.283 | 0.187          | 0.428 | 1.899 hours |

**Table 2.** Bias, RMSE and CPU of the CLM estimator of two-parameters transformed Gumbel copula.

|                 | $\tau = 0.01$  |       |                 |       | $\tau = 0.5$   |       |               |       |
|-----------------|----------------|-------|-----------------|-------|----------------|-------|---------------|-------|
|                 | $\alpha = 0.1$ |       | $\beta = 1.059$ |       | $\alpha = 0.5$ |       | $\beta = 1.6$ |       |
|                 | Bias           | RMSE  | Bias            | RMSE  | Bias           | RMSE  | Bias          | RMSE  |
| $n = 30$        |                |       |                 |       |                |       |               |       |
| CLM             | -0.066         | 0.067 | 0.151           | 0.299 | 0.088          | 0.401 | -0.415        | 0.602 |
| PML             | 0.142          | 0.453 | -0.495          | 0.476 | -0.139         | 0.867 | 0.974         | 1.906 |
| $\rho$ - $\tau$ | 1.252          | 2.785 | -0.621          | 1.762 | -0.552         | 2.212 | -0.985        | 1.132 |
| $n = 50$        |                |       |                 |       |                |       |               |       |
| CLM             | -0.071         | 0.082 | 0.109           | 0.181 | -0.017         | 0.341 | 0.315         | 0.543 |
| PML             | -0.096         | 0.345 | -0.287          | 0.442 | -0.163         | 0.606 | -0.482        | 0.664 |
| $\rho$ - $\tau$ | 1.005          | 1.893 | -0.578          | 1.367 | -0.455         | 1.991 | -0.852        | 1.031 |
| $n = 100$       |                |       |                 |       |                |       |               |       |
| CLM             | 0.042          | 0.069 | 0.029           | 0.109 | 0.029          | 0.158 | 0.155         | 0.357 |
| PML             | -0.054         | 0.191 | -0.154          | 0.381 | -0.082         | 0.367 | -0.403        | 0.469 |
| $\rho$ - $\tau$ | 0.862          | 1.124 | -0.473          | 1.271 | -0.355         | 1.104 | -0.752        | 0.991 |
| $n = 200$       |                |       |                 |       |                |       |               |       |
| CLM             | 0.026          | 0.063 | 0.012           | 0.074 | 0.009          | 0.106 | 0.013         | 0.105 |
| PML             | -0.040         | 0.132 | -0.211          | 0.305 | -0.052         | 0.172 | -0.324        | 0.428 |
| $\rho$ - $\tau$ | 0.771          | 1.020 | -0.403          | 1.110 | -0.310         | 0.891 | -0.656        | 0.786 |

|                 | $\tau = 0.8$   |       |                |       |
|-----------------|----------------|-------|----------------|-------|
|                 | $\alpha = 0.9$ |       | $\beta = 3.45$ |       |
|                 | Bias           | RMSE  | Bias           | RMSE  |
| $n = 30$        |                |       |                |       |
| CLM             | -0.139         | 0.399 | 0.312          | 1.022 |
| PML             | -0.870         | 0.682 | 0.899          | 1.049 |
| $\rho$ - $\tau$ | -0.437         | 1.798 | 1.229          | 2.473 |
| $n = 50$        |                |       |                |       |
| CLM             | 0.085          | 0.395 | 0.310          | 0.772 |
| PML             | -0.765         | 0.488 | 0.782          | 0.901 |
| $\rho$ - $\tau$ | -0.402         | 1.326 | 1.174          | 2.140 |
| $n = 100$       |                |       |                |       |
| CLM             | -0.023         | 0.167 | 0.117          | 0.537 |
| PML             | -0.283         | 0.387 | 0.608          | 0.786 |
| $\rho$ - $\tau$ | -0.420         | 0.895 | 1.024          | 1.865 |
| $n = 200$       |                |       |                |       |
| CLM             | -0.019         | 0.139 | 0.101          | 0.411 |
| PML             | 0.216          | 0.349 | 0.480          | 0.525 |
| $\rho$ - $\tau$ | -0.221         | 0.321 | 0.721          | 0.989 |

**Table 3.** Bias and RMSE of the CLM, PML and  $\tau$ - $\rho$  estimators of two-parameters transformed Gumbel copula.

## Appendix A: Appendix

### A.1. Proof of Theorem 1

First notice that CLM’s estimator may be considered as a RAZ-estimator (van der Vaart [15, page 41]). Then by applying Theorem 1 in Tsukahara [13] we may easily establish the consistency and asymptotic normality of the considered estimator. Indeed, the existence of a sequence of consistent roots  $\hat{\theta}^{CLM}$  to (6), can be checked by using similar arguments as the proof of Theorem 1 in Tsukahara [13]. That is, we have to verify only the conditions in Theorem A.10.2 in Bickel *et al.* [1]. Firstly, recall the functions (7) and set

$$\Phi(\theta) := \int_{\mathbb{I}^d} \mathbf{g}(\mathbf{u}; \theta) dC_{\theta_0}(\mathbf{u}), \text{ and } \Phi_n(\theta) := n^{-1} \sum_{i=1}^n \mathbf{g}(\hat{\mathbf{U}}_i; \theta),$$

where  $\hat{\mathbf{U}}_i = (F_{1n}(X_{1i}), \dots, F_{dn}(X_{di}))$ , with  $(X_{j1}, \dots, X_{jn})$  is a given random sample from the rv  $X_j$ . From assumption [A2] the following derivatives exist

$$\dot{\Phi}(\theta) = \int_{\mathbb{I}^d} \dot{\mathbf{g}}(\mathbf{u}; \theta) dC_{\theta_0}(\mathbf{u}), \quad \dot{\Phi}_n(\theta) = \frac{1}{n} \sum_{i=1}^n \dot{\mathbf{g}}(\hat{\mathbf{U}}_i; \theta).$$

We have to verify

$$\sup \left\{ \left| \dot{\Phi}_n(\theta) - \dot{\Phi}(\theta) \right| : |\theta - \theta_0| < \epsilon_n \right\} \xrightarrow{\mathbb{P}} 0, \text{ as } n \rightarrow \infty, \quad (\text{A1})$$

for any real sequence  $\epsilon_n \rightarrow 0$ . Indeed, the continuity of function  $\dot{\mathbf{g}}$  in  $\theta$  implies that

$$\sup \left\{ \left| \dot{\mathbf{g}}(\hat{\mathbf{U}}_i; \theta) - \dot{\mathbf{g}}(\hat{\mathbf{U}}_i; \theta_0) \right| : |\theta - \theta_0| < \epsilon_n \right\} = o_{\mathbb{P}}(1), \quad i = 1, \dots, n.$$

Also, since

$$\left| \dot{\Phi}_n(\theta) - \dot{\Phi}_n(\theta_0) \right| \leq \frac{1}{n} \sum_{i=1}^n \left| \dot{\mathbf{g}}(\hat{\mathbf{U}}_i; \theta) - \dot{\mathbf{g}}(\hat{\mathbf{U}}_i; \theta_0) \right|,$$

then

$$\sup \left\{ \left| \dot{\Phi}_n(\theta) - \dot{\Phi}_n(\theta_0) \right| : |\theta - \theta_0| < \epsilon_n \right\} \xrightarrow{\mathbb{P}} 0, \text{ as } n \rightarrow \infty. \quad (\text{A2})$$

Making use of the law of the large number, we get  $n^{-1} \sum_{i=1}^n \dot{\mathbf{g}}(\mathbf{U}_i; \theta_0) \xrightarrow{\mathbb{P}} \dot{\Phi}(\theta_0)$ , as  $n \rightarrow \infty$ , where  $\mathbf{U}_i = \{F_j(X_{ji})\}_{1 \leq j \leq d}$ . Moreover, since  $\dot{\mathbf{g}}$  is continuous in  $\mathbf{u}$  then by Glivenko-Cantelli theorem, we infer that  $n^{-1} \sum_{i=1}^n \left| \dot{\mathbf{g}}(\hat{\mathbf{U}}_i; \theta_0) - \dot{\mathbf{g}}(\mathbf{U}_i; \theta_0) \right| \xrightarrow{\mathbb{P}} 0$ . It follows that  $\left| \dot{\Phi}_n(\theta_0) - \dot{\Phi}(\theta_0) \right| \xrightarrow{\mathbb{P}} 0$ , which together with (A2), implies (A1). Conditions (MG0) and (MG3) in Theorem A.10.2 in Bickel *et al.* [1] are trivially satisfied by our assumptions [H1] – [H3]. In view of the general theorem for  $Z$ -estimators (see, van der Vaart and Wellner [14], Theorem 3.3.1), it remains to prove that  $\sqrt{n} \left( \dot{\Phi}_n - \dot{\Phi} \right) (\theta_0)$  converges in law to the appropriate limit. But this follows from Proposition 3 in Tsukahara [13], which achieves the proof of Theorem 1.  $\square$

## References

- [1] Bickel, P. J., Klaassen, C. A. J., Ritov, Y. and Wellner, J. A., 1993. Efficient and Adaptive Estimation for Semiparametric Models. Baltimore, MD: *The Johns Hopkins University Press*.
- [2] Deheuvels, P., 1979. La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg. Bull. Cl. Sci.* **65**, 274-292.
- [3] Genest, C., 1987. Frank's family of bivariate distributions. *Biometrika* **74**, 549-555.
- [4] Genest, C., Ghoudi, K. and Rivest, L. P., 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543-552.
- [5] Hosking, J.R.M., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J. Royal Statist. Soc. Ser. B* **52**, 105-124.
- [6] Hosking, J.R.M., 1998. L-moments. in: Kotz, S., Read, C., Banks, D.L. (Eds.), *Encyclopedia of Statistical Sciences*, vol. **2**. Wiley, New York, 357-362.
- [7] Joe, H., 1997. *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.
- [8] Joe, H., 2005. Asymptotic efficiency of the two-stage estimation method for copula-based models. *J. Multivariate Anal.* **94**, 401-419.
- [9] Kim, G., Silvapulle, M. J. and Silvapulle, P., 2007. Comparison of semiparametric and parametric methods for estimating copulas. *Comm. Statist. Simulation Comput.* **51**, 2836-2850.
- [10] Nelsen, R.B., 2006. *An Introduction to Copulas*, second ed. Springer, New York.
- [11] Oakes, D., 1982. A model for association in bivariate survival data. *J. Roy. Statist. Soc. Ser.* 414-422.
- [12] Sklar, A., 1959. Fonctions de répartition à  $n$  dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris* **8**, 229-231.
- [13] Tsukahara, H., 2005. Semiparametric estimation in copula models. *Canad. J. Statist.* **33**, 357-375.
- [14] van der Vaart, A. W., Wellner, J. A., 1996. *Weak Convergence and Empirical Processes: With applications to Statistics*. Springer, New York.
- [15] van der Vaart, A.W., 1998. *Asymptotic Statistics*, Cambridge University Press.
- [16] Yan, J., 2007. Enjoy the Joy of Copulas: With a Package copula. *Journal of Statistical Software*, **21**(4), 1-21.