

Simulation techniques of Archimedean Copula Estimators: Parametric and Semi-Parametric Approaches

N. Idiou and F. Benatia

Abstract — In this paper, we look at two different approaches methodologies for copula estimation. The first is based on a parametric approach using MLE and IFM methods, while the second is entirely based on Kendall's tau and Spearman's rho in a semi-parametric context, where the margins are estimated non-parametrically. Interestingly, based on R software simulation techniques, the contribution of their algorithms, approach, and illustration was our main focus for this paper. As an application, a class of Archimedean copulas was notably chosen. This particular class of copulas was also presented for censored data to show the estimator's performance even better.

Key words — Archimedean copula, Parametric estimation, Simulation study, R software, Semi-parametric approaches.

I. INTRODUCTION

In mathematics, a copula is a function that serves as the primary link between the multivariate distribution function and its univariate margins, as proven by Sklar's theorem (1959). Regardless of the shape of the margins, the action of the copula is to represent the characteristics of dependence that are associated with each of the random variables. Under certain conditions, Sklar's theorem [14] shows that there is a unique copula function C , which is given by:

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (1)$$

where $F(x_1, \dots, x_n)$ is the joint distribution of $X = (X_1, \dots, X_n)$ and $F_1(x_1), \dots, F_n(x_n)$ are the margins.

Nelsen [14], Genest and MacKay [9] describe a variety of copula functions that can be used to fit a wide variety of dependence types. Our concentration in this paper will focus on the Archimedean copulas class for reasons of dependency modeling and its approaches, particularly in actuarial science and financial risk management, then our main reason for focusing on this class is:

- (1) they are easy to construct,
- (2) have interesting properties that further facilitate the modeling of dependency structures,
- (3) a large variety of copula families belonging to this class. They are often characterized by a generator, which is a function, thus reducing the search for a large dimensional distribution function (for more details about Archimedean copulas see part (2)).

Several methods for estimating copula parameters have been developed, among the most important, the minimum distance (MD) [17]), a function of margin inference (IFM) [28], [29], the concordance methods based on [15], Genest,

[23], also known as tau-inversion and rho-inversion, which are based on Kendall's tau and Spearman's rho rank correlation coefficients respectively, defined as:

$$\begin{aligned} \tau &= \frac{P[(X_1^{(1)} - X_1^{(2)})(X_2^{(1)} - X_2^{(2)}) > 0]}{P[(X_1^{(1)} - X_1^{(2)})(X_2^{(1)} - X_2^{(2)}) < 0]} \\ \rho &= 3 \left\{ \frac{P[(X_1^{(1)} - X_1^{(2)})(X_2^{(1)} - X_2^{(3)}) > 0]}{P[(X_1^{(1)} - X_1^{(2)})(X_2^{(1)} - X_2^{(3)}) < 0]} \right\} \end{aligned}$$

where $(X_1^{(1)}, X_2^{(1)}), (X_1^{(2)}, X_2^{(2)}), (X_1^{(3)}, X_2^{(3)})$ are three independent (iid) copies of the bivariate random vector $(X_1; X_2)$. However, they may be represented as follows:

$$\tau(C) = 4 \int_{[0,1]^2} C(u)du - 1 \quad (2)$$

$$\rho(C) = 12 \int_{[0,1]^2} C(u)du - 3 \quad (3)$$

which shows that Spearman's rho and Kendall's tau depend on the underlying copula C (they can be also considered as moments of the copula). Subsequently, many other authors were interested by copula families of the Archimedean class for the application of these methods, due to their mathematical structure and to their ease properties despite the high dimensional level.

In copula approaches viewpoint and concerning their estimate, if the margins F_1, \dots, F_d are known, then we bring to classic statistical inference methods. But, because the margins are generally unknown, mainly two approaches can be adopted for its estimations parametric and nonparametric. Nevertheless, in the first approach we estimate the margins parametrically, i.e., the resulting estimate of C will be entirely parametric, we suppose that the marginal belongs to a family indexed by a parameter, so to estimate the margins, it suffices to estimate their parameter. Although in the other case the margins will estimate non-parametrically, i.e., we do not assume that the margins belong to any family, then it will be semi-parametric copula estimation, for example with the version of the empirical estimator given by:

$$\hat{F}_i(x) = \frac{1}{n} \sum_{k=1}^n I_{\{X_k \leq x_i\}} \quad (4)$$

Therefore, the parameter of the copula must be estimated, whatever the way we estimate the margins. Our focus for the application is done on the Archimedean class of copulas in R software, which is free statistical calculation software. It works on all platforms and is available and has many existing features characterized by contributed packages.

Writing an R package for copulas was natural, this package (copula) mainly allows us to build and work easily with the Archimedean Copulas [20] [10].

Among the package copula advantages is the evaluation of the Archimedean copulas, currently widely used and well-known as Ali-Mikhail-Haq, Clayton, Frank, Gumbel, and Joe. Therefore, in regard to association measures, Kendall's tau and the coefficients of dependence are implemented.

In this paper, two approaches are presented for estimating copulas, the parametric and semi-parametric one. In the first approach, two methods are presented, the maximum likelihood method (also known as the MLE method) and the inference function method for margins (IFM), which simplify the estimation procedure when the MLE method is impracticable.

Even so, because the margins will be estimated non-parametrically, a semi-parametric approach is required. Nonetheless, for this type of estimation, the method of moments based on Kendall's tau and Spearman's rho is presented. For the reasons stated previously, the class of Archimedean copulas was selected as an application. However, the purpose of this paper is to introduce this type of copula, as well as its applications, methodology, and visualization in R software.

The remaining sections of the paper are organized as follows. In section two, some basic concepts concerning copula and their characterization have been presented, notably the Archimedean copulas. Two methods of copula estimation using parametric approaches are presented in section three, where the margins have been parametrically estimated. In section four, where the margins are estimated non-parametrically, a semi-parametric estimation was presented using the method of moments based on Kendall's tau and Spearman's rho, respectively. In part five of the paper, we used the MLE estimator to introduce the log-likelihood function and the estimation of a copula using two stages dedicated to this method using censored data. Part six includes a conclusion and discussion. For illustration, the Appendix interacts with a key in R.

II. ARCHIMEDEAN COPULAS AND THEIR CHARACTERIZATION

The mathematical properties of the Archimedean class are characterized by a generator function as we have already said. Let $u = (u_1, \dots, u_d)$ a d-dimensional unit vector, where $u \in [0,1]^d$ and assuming that φ a mapping defined by:

$$\varphi: [0, \infty] \rightarrow [0,1] \quad (4)$$

called a generator and satisfies the following conditions:

$$- \varphi(1) = 0 \quad (5)$$

- φ is continuous, convex, decreasing function, i.e.,

$$(\varphi'(t) < 0 \text{ and } \varphi''(t) > 0, \text{ for } 0 < t < 1) \quad (6)$$

An Archimedean copula is a copula of the form:

$$C(u) = \varphi(\varphi^{-1}(u_1) + \dots + \varphi^{-1}(u_d)) \quad (7)$$

where φ^{-1} is the inverse function of the generator φ , that can be indexed by the association parameter θ , thus a whole family of copulas can be Archimedean. Notice that, one of the nice properties of Archimedean copula is that the distributions function $K_\theta(t) = P(C_\theta \leq t)$ of C_θ can represent according to its generator, where C_θ is the empirical version of C. See Theorem 4.3.4 in Nelsen [14], then for any $t \in [0,1]$, we have:

$$K_\theta(t) = t - \frac{\varphi_\theta(t)}{\varphi'_\theta(t)}$$

Therefore, the corresponding density is:

$$K'_\theta(t) = \frac{\varphi''_\theta(t)\varphi_\theta(t)}{(\varphi'_\theta(t))^2}$$

In order for (7) to be a copula, the generator must be a d-monotonic function see Kimberling (1974) in particular reference to t-norms, Hofert [27] in respect of copula reworking.

Therefore, for any function φ which satisfies conditions (5) and (6) we can say that a function of the form (7) is an Archimedean copula if and only if $\varphi \in \Theta$, where

$\Theta = [0, \infty] \rightarrow [0,1]$ and φ is d-monotone. The advantage of using Archimedean copulas in mathematics is the realization of dimensionality reduction, while the copula of n-variables is a function with n-places, the generator φ never takes only one argument. In the bivariate case, the means by which φ generates the copula is according to:

$$\varphi(C(u, v)) = \varphi(u) + \varphi(v) \quad (8)$$

If $\varphi(0) = \infty$, the generator is termed strict, and the inverse function φ^{-1} exists. In this case, from (8), the copula is recovered by:

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$$

Hence, the non-strict generators are those for which ($\varphi(0) < \infty$), we say that the generators have a singular component, in this case we must begin by defining a pseudo-inverse function φ^{-1} . As an example:

$$\varphi(t) = 1 - t, \rightarrow \varphi^{-1}(t) = \max(1 - t; 0),$$

Noted that:

$$\varphi^{-1}(\varphi(u) + \varphi(v)) = \max(u + v - 1; 0) = W$$

The lower Frechet bound for (bivariate) copulas is Archimedean. For more details on Archimedean copulas

(strict and not strict), see Nelsen [14], Genest and MacKay [9], Genest and Rivet [8], Jouini and Clemen [30], Mari and Kotz [12]. Many applications on Archimedean copula have been discussed in finance, actuarial science, see [18] who uses Archimedean copulas to induce dependence among random variables.

Examples of families of Archimedean copulas are listed in Table I. An immediate consequence of the definition of d-monotonicity is that, in the bivariate case, C in (7) is a copula if and only if $\varphi \in \Theta$ is convex. Furthermore, a generator $\varphi \in \Theta$ generates an Archimedean copula in arbitrary dimensions if and only if φ is completely monotone.

We focus here on the latter case, where the complete monotonicity is fulfilled by several commonly used parametric generators (Archimedean copula for example).

TABLE I: FAMILIES OF BIVARIATE ARCHIMEDEAN COPULAS

Family	Copulas	Parameters pace
AMH	$\frac{uv}{(1 - \theta(1-u)(1-v))}$	$\theta \in (0,1)$
C	$\frac{(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}}{-\theta^{-1} \ln(1 - (e^{-\theta u} - 1)(e^{-\theta v} - 1))}$	$\theta \in (0, \infty)$
F	$\frac{1}{\exp(-(-\ln(u))^{\theta})}$	$\theta \in (0, \infty)$
G	$\frac{1}{1 + (-\ln(v))^{\theta}}$	$\theta \in [1, \infty)$
J	$\frac{1}{-((1-u)^{\theta} + (1-v)^{\theta}) - (1-u)^{\theta}(1-v)^{\theta}}$	$\theta \in [1, \infty)$

Table I represents particular Archimedean copulas families (completely monotone) and the expressions of their corresponding generators, the families of Ali-Mikhail-Haq (AMH), Clayton (C), Frank (F), Gumbel-Hougaard (G) and Joe (J) (See also Hofert [26]).

The parameter θ can vary only in $(0, \infty)$. The latter is the range for which the Clayton and Frank generators are completely monotonous. Also note that for the Ali-Mikhail-Haq family, Kendall's range of attainable tau is limited in $[0, 1/3]$.

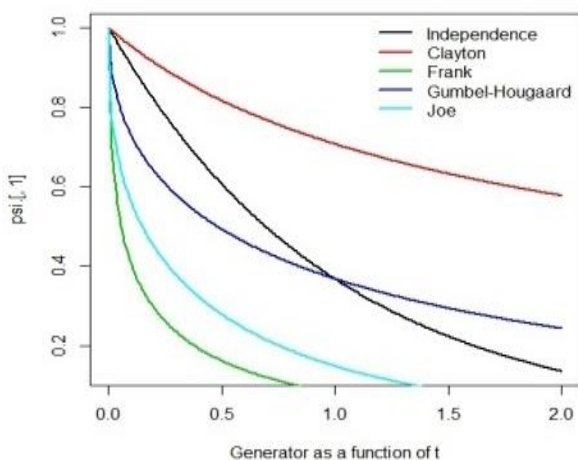


Fig. 1. Archimedean generators listed in Table I.

Measuring dependence between two random variables or more (also called a generalized correlation) often becomes necessary in most useful applications on copulas. These

measures are cited as association measures and are generally studied in the bivariate case. Kendall's tau is one of those measures, for a vector (X, Y) is given by:

$$\tau = E[\text{sing}((X_1 - X_2)(Y_1 - Y_2))]$$

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] -$$

$$P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

where (X_1, Y_1) and (X_2, Y_2) are independent identically distributed copies of the vector (X, Y) and called the signum function:

$$\text{sing}(x) = 1_{(0, \infty)}(x) - 1_{(-\infty, 0)}(x)$$

In a different way, the Kendall's tau is defined as the probability of concordance less the probability of discordance [16], it measures, as a number in $[-1, 1]$.

Therefore, by principle an Archimedean copula can typically be expressed in terms of their generators, i.e., generated by a twice continuously differentiable generator with:

$$\varphi(t) > 0, t \in [0, 1].$$

Then, Kendall's tau can be represented as:

$$\tau = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt \Rightarrow \tau = 1 + 4 \int_0^1 \frac{\varphi^{-1}(t)}{\varphi^{-1}(t)'} dt$$

For the proof see [14].

TABLE II: KENDALL'S TAU FOR THE ARCHIMEDEAN COPULAS

Family	Copulas	τ interval
AMH	$1 - (2(\theta + (1 - \theta)^2 \ln(1 - \theta)))/3\theta^2$	$-0.1817 \leq \tau < \frac{1}{3}$
C	$\frac{\theta/(\theta + 2)}{1 + (4(D_1(\theta) - 1))/\theta}$	$0 \leq \tau < 1$
F	$(\theta - 1)/\theta$	$-1 \leq \tau < 1$
G	$\frac{1}{\sum_{k=1}^{\infty} (k(\theta k + 2)(\theta(k - 1) + 2))}$	$0 \leq \tau < 1$
J	$1 - 4 \sum_{k=1}^{\infty} \frac{1}{(k(\theta k + 2)(\theta(k - 1) + 2))}$	$0 \leq \tau < 1$

Table II represent the Kendall's tau for each family given in Table I for Archimedean copulas.

For example, for the Clayton family, $\tau = \frac{\theta}{\theta + 2}$, then this family covers $0 \leq \tau < 1$, where the quantity D_1 is the Debye function of order one, defined by:

$$D_1(\theta) = \int_0^{\theta} \frac{t}{\exp(t) - 1} \frac{dt}{\theta}, \theta \in]0, \infty[.$$

III. RESULTS

The estimation of copula parameters is the main focus of this section, even by way, over the first case, the estimation of the copula C is completely parametric, and it is assumed that the margins belong to a parameter-indexed family. Hence, in order to estimate their parameters, it is sufficient to estimate the margins. To hold this estimate, it is desirable to parameterize the copula function for statistical modeling

purposes, for a reason that data can be used to shed light on the extent of the association between the random variables.

Assume that the copula C belongs to a continuous parametric family $\{C_\theta: \theta \in \Theta\}$, where Θ the open subset of \mathbb{R}^p for some integer $p \geq 1$ represent the parameter space θ , denote the bivariate distribution association parameter (possibly a vector). Then the copula is assumed to be parameterized as C_θ (i.e., a copula that belongs to this family).

This notation designates a family of copulas, where the margins F_1 and F_2 are generally independent of θ .

Even so, estimate C_θ amounts to estimating the vector of unknown parameters, generally carried out in two steps:

- Step1: margins-based estimation.
- Step 2: copula-based estimation to estimate its parameters.

By using the maximum likelihood estimator (MLE) firstly then certain likelihood functions are to be maximized. Sometimes, this likelihood cannot be maximized. However, another strategy can propose on the margins called the Inference function method for margins (IFM), which overcame this problem.

A. Maximum Likelihood Estimator (MLE)

Let β be the vector of marginal parameters and θ be the vector of copula parameters. Given the relatively simple functional form the self-selection likelihood function under an Archimedean copula, MLE can be employed to jointly estimate all parameters of the unknown parameters vector $(\beta_1, \dots, \beta_d, \theta)$ at the same time. Assume that we observe d -independent realizations (X_{i1}, \dots, X_{ip}) , $i = 1, \dots, d$, specified by pmargins with (CDF) cumulative distribution function F_i . However, the density of F defined by the formula (1) is given by:

$$f(x_1, \dots, x_d; \theta) = c_\theta(F_{1,\beta_1}(x_1), \dots, F_{d,\beta_d}(x_d)); \theta) \prod_{i=1}^d f_{i,\beta_i}(x_i) \quad (9)$$

That is associated with a sample (X_{i1}, \dots, X_{ip}) , where c_θ is a density of a parametric copula C_θ and f_{i,β_i} is a density of F_{i,β_i} .

Both approaches presented seek to maximize a likelihood approximation based on (9). Consequently, the parameter vector to be estimated in the parametric approach is $\alpha = (\beta, \theta)$ and the function of loglikelihood is given by:

$$l(\beta_1, \dots, \beta_d, \theta) = \sum_{i=1}^n \log c_\theta\{F_{1,\beta_1}(X_{i1}), \dots, F_{d,\beta_d}(X_{id}); \theta\} + \sum_{j=1}^d \sum_{i=1}^n \log f_{j,\beta_j}(X_{ij})$$

However, the MLE of θ will given by:

$$\hat{\theta}_{ML} = \text{argmax}_l(\theta), \text{ where } \theta \in \Theta.$$

See, Lehmann and Casella [11] for more details. Under a specific copula model, the likelihood function is to be

evaluated in R software by using “loglikCopula” and “loglikMvdc” packages. However, the “loglikis” most commonly used for a model fitted by maximum likelihood.

To illustrate, let choose C as a bivariate Archimedean copula (dimension $d = 2$), where the margins F_1 and F_2 are distributed under Normal $(N(\mu, \sigma^2))$ and exponential $(\exp(\lambda))$ distributions respectively. By the way, the representation of these multivariate distributions F_j ($j = 1; 2$ in our case) can be done by using the “mvdc()” packages.

Since, the “fitMvdc” function is a wrapper for routine optimization in R, then to obtain the maximum likelihood estimator, a function “fitMvdc ()” was selected.

The results obtained in TABLE II are established for the four most famous families of Archimedean copulas. From a sample of 1000 independent observations of such distribution with parameters margins:

$$\beta_1 = (\mu = 0, \sigma^2 = 1) \text{ and } \beta_2 = \lambda = 1,$$

where the copula parameter is chosen $\theta = 2$.

TABLE III: ARCHIMEDEAN COPULA PARAMETERS ESTIMATION VIA MLE

Family	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\theta}$	Maximized log likelihood
Clayton	-0.0047	0.9752	1.0290	1.955	-1935
Frank	-0.0021	0.9762	0.9597	1.778	-2398
Gumbel	-0.0556	1.0010	1.0040	1.990	-2040
Joe	-0.0235	1.0098	0.9984	2.030	-2189

A richer performance includes standard marginal and copula parameter errors estimates that can be acquired by asking summary. See the following table:

TABLE IV: STANDARD ERRORS ESTIMATES PARAMETERS VIA MLE

Family	$\hat{\mu}$ Std. Error	$\hat{\sigma}$ Std. Error	$\hat{\lambda}$ Std. Error	$\hat{\theta}$ Std. Error
Clayton	0.026	0.016	0.032	0.09
Frank	0.030	0.022	0.03	0.196
Gumbel	0.031	1.019	0.031	0.059
Joe	0.032	1.021	0.031	0.08

In the parametric approach, through the use of a maximum likelihood estimator, the margins are estimated parametrically. Sometimes, this likelihood can be complicated and impossible to calculate, where the numerical optimization is available, and it might be too difficult. In these situations, another technique can propose to overcome this complication, generally called the IFM method which proceeds in two stages (the reader is invited to refer to the (IFM) method and the parametric MLE method).

B. Margin Inference Function Method (IFM)

This method was introduced by Xu [29] and Joe [16] in a general framework. He called the inference function method for margins (IFM) because estimation functions relate to likelihood score functions (univariate or multivariate).

The IFM was used primarily for multivariate models in which a multi-parameter numerical optimization for maximum likelihood estimation is unattainable or takes a too long time if we talk about a time-consuming viewpoint. Therefore, the estimation by the IFM method can be decreased the computational load potentially associated with

the estimation of the maximum likelihood. Hence, this method proceeds in two stages:

- Stage 1: Estimate the margins parameters firstly.
- Stage 2: fix the marginal parameters obtained in the first stage, and then estimate the copula parameters.

See Joe [28], [29]. By analogy, the unknown margins parameter vectors $(\beta_1, \dots, \beta_d)$ are first estimated by:

$$\beta_{n,j} = \operatorname{argsup} \left(\sum_{i=1}^n \log f_{j,\beta_j}(X_{ij}) \right), \text{ where } \beta_j \in \mathbb{R}^{p_j}$$

For $j = 1, \dots, d$. By the way in the second stage, estimating the vector of the unknown copula parameter θ is performed by:

$$\hat{\theta}_{\text{IFM}} = \operatorname{argsup} \left(\sum_{i=1}^n \log c_{\theta} \left(F_{1,\beta_{n,1}}(x_{1i}), \dots, F_{d,\beta_{n,d}}(x_{di}); \theta \right) \right),$$

where $\theta \in \Theta$. The main advantage of this approach is that this method (IFM) is done in two stages. Wherever, under the MLE estimate, when models are usually multivariate, the number of parameters increases, and the numerical optimization becomes more complicated. Also, for some models, multi-dimensional numerical integration is needed, and this becomes increasing difficult, for the theoretical properties of the IFM method see [28].

To illustrate, we take into account the same setting used in the MLE estimation. The key R function is “fitCopula ()” with argument “method = “ml””.

TABLE V: ARCHIMEDEAN COPULA PARAMETERS ESTIMATION VIA IFM

Family	$\hat{\theta}$	Std. Error	Maximized
			Log likelihood
Clayton	1.956	0.086	425.8
Frank	1.779	0.195	41.77
Gumbel	1.98	0.051	3600
Joe	2.021	0.065	2334

If the margins are carelessly specified, the estimate will be biased. Hence, despite the advantages of the IFM, we mention that this method does not prevent it from suffering the same disadvantage as an estimate of the maximum likelihood. The proof of this phenomenon was obtained in Fermanian and Scaillet [31] and Kim et al. [32].

In the semi-parametric approach, we do not assume that the margins (F_1, \dots, F_d) belong to any parametric family. Then, we estimate them directly by the nonparametric estimator given by (4). Therefore, we replace the margins by their estimate in (9), to estimate θ , and we maximize the part of the likelihood involving θ . For more details, see [7].

The consistency of the estimator it is asymptotically normal under some conditions. However, despite these properties of convergence, it is not, in general efficient, except in the case of the bivariate Gaussian copula. See [6], [32].

IV. SEMI-PARAMETRIC ESTIMATION METHODS FOR COPULA

In the previous section, we established parametric methods to estimate copula parameters, where the estimation constructed on the margins is completely parametric. In this section, we focus on non-parametric estimation of these margins by using a moment’s estimator, performed on coefficients of dependency (Kendall’s tau and Spearman’s rho), which leads us to pass automatically to a semi-parametric approach. By the same way, the application is made on Archimedean copulas.

A. Copula Estimation via Moment Estimator Based on Kendall’s Tau and Spearman’s Rho

The estimation methods based on the correlation coefficients of Spearman’s rho or Kendall’s tau rank are called tau-inversion methods (respectively, rho-inversion) or concordance method (mentioned in part one), take advantage of the relationship between these coefficients of dependency and the parameter of the copula θ . Early references on the former in a copula setting are, among others, Oakes [15], Genest [23], Genest and Rivest [8].

We have seen the definition of the dependency coefficients given by formulas (2) and (3). However, for the pair (X_i, X_j) the empirical version of these coefficients is given by:

$$\hat{\tau}_{i,j} = \binom{n}{2}^{-1} \sum \operatorname{sign}(X_i^{(p)} - X_i^{(l)})(X_j^{(p)} - X_j^{(l)})$$

$$\hat{\rho}_{i,j} = \frac{\sum_{p=1}^n (\hat{U}_i^{(p)} - \bar{U}_i)(\hat{U}_j^{(p)} - \bar{U}_j)}{\sum_{p=1}^n (\hat{U}_i^{(p)} - \bar{U}_i)^2 \sum_{p=1}^n (\hat{U}_j^{(p)} - \bar{U}_j)^2}^{\frac{1}{2}}$$

where $\hat{U}_i^{(p)} = \hat{F}_i(X_i^{(p)})$ and $\bar{U}_i = \sum_{p=1}^n \frac{\hat{U}_i^{(p)}}{n}$, $i = 1, \dots, d$ and:

$$\begin{cases} \operatorname{sign}(x) = 1 & | x > 0 \\ \operatorname{sign}(x) = -1 & | x < 0 \\ \operatorname{sign}(x) = 0 & | x = 0 \end{cases}$$

Then, the estimators $\hat{\theta}$ obtained by tau-inversion methods (respectively, rho-inversion methods) are given respectively by:

$$\begin{cases} \tau(\hat{\theta}) = \hat{\tau}_{i,j} \text{ then } \hat{\theta} = \tau^{-1}(\hat{\tau}_{i,j}) \\ \rho(\hat{\theta}) = \hat{\rho}_{i,j} \text{ then } \hat{\theta} = \rho^{-1}(\hat{\rho}_{i,j}) \end{cases}$$

where τ^{-1} and ρ^{-1} are the inverses of τ and ρ respectively, if they exist.

In the bivariate case ($d = 2$) and when there is only one real parameter to estimate, the tau-inversion method applies to match the estimate under the model with its empirical estimate. Therefore, the estimator $\hat{\theta}$ satisfies:

$$\tau(\hat{\theta}) = \hat{\tau}_{1,2}, \text{ then } \hat{\theta} = \tau^{-1}(\hat{\tau}_{1,2})$$

These two estimators are consistent and asymptotically unbiased. Noted that the inversion method of Kendall’s tau or Spearman’s rho is belonging to a semi-parametric approach,

since the expressions of their empirical version for margins F_1, \dots, F_d are implicitly estimated non-parametrically by (4).

Therefore, if we suppose that C is the family of Gumbel copulas, then the estimator of θ by the tau-inversion method:

$$\tau(C_\theta) = 1 - \frac{1}{\theta} \text{ and } \widehat{\theta} = \tau^{-1}(\widehat{\tau}) = \frac{1}{1-\widehat{\tau}}$$

In R software, the name of the function corresponding to τ^{-1} is `iTau()`.

To illustrate, we suggest that the copula selected is belonging to a bivariate Archimedean family with a single parameter $\theta = 2$, where the margins are distributed under a standard normal. To applied, let a 1000 independent observations the following table summarizes our application in R:

TABLE VI: ESTIMATION OF ARCHIMEDEAN COPULA PARAMETERS VIA THE METHOD OF MOMENTS BASED ON KENDALL'S TAU

Family	τ theo value	Std. Error	θ estimate
Clayton	0.5026	0.068	2.027
Frank	0.1909	0.031	1.7708
Gumbel	0.4939	0.065	1.9760
Joe	0.3551	0.048	2.000

By the use of a standard error, we find the values corresponding across each copula indicated in the last columns of the previous table. By analogy using the same propositions as the previous one when the estimate is based on Spearman's rho. The results of the application are summarized in Table VII:

TABLE VII: ESTIMATION OF ARCHIMEDEAN COPULA PARAMETERS VIA THE METHOD OF MOMENTS BASED ON SPEARMAN'S RHO

Family	τ theo value	Std. Error	θ estimate
Clayton	0.6840	0.019	2.0078
Frank	0.2843	0.030	1.7763
Gumbel	0.6756	0.019	1.9746

The censored and uncensored samples (artificial data) for a type of non-Archimedean copula (bivariate Gaussian copula) are presented in the following figures.

V. APPLICATION: MLE ESTIMATOR FOR CENSORED DATA

A. Application for non-Archimedean Copula

When the variables contain censored observations (either left, right, or interval censoring), the copula is illustrated also by a two-step parametric MLE approach:

- Estimate the marginal parameters.
- Set the marginal parameters to the estimated values in the first step, then estimate the copula parameters.

A bivariate (non-Archimedean) copula fits the R code presented in the appendix to the data where one of the variables contains censored observations. However, all functions are required by the two parametric stages MLE estimation method to fit a copula.

The packages `copula`, `"MASS"`, `"survival"`, `"SPREDA"` are needed by all the procedures in R, therefore we make them installed and loaded by the `"install.packages(.)"` and `"library"` functions respectively.

The function for computing the log-likelihood in censoring is given by "loglikCensored". To illustrate, we generate some artificial data using a non-Archimedean copula (bivariate Gaussian copula), for a censored observation that will subsequently be treated as right-censored cases.

Assume that the first marginal is a normal distribution $N(\text{mean} = 5, \text{sd} = 0.8)$, and the second marginal is a Weibull distribution with parameters (location=2.4 and scale=1.3). The censored and uncensored samples (artificial data) for a type of non-Archimedean copula (bivariate Gaussian copula) are presented in the following figure.

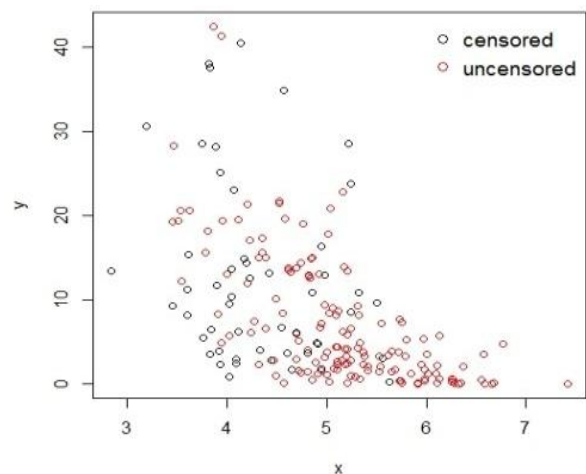


Fig. 2. Censored and uncensored artificial data for a bivariate Gaussian copula.

We are now interested by illustrating a Gaussian copula based on the MLE parametric estimation method, also known as the functions of the margin method inference (MFI) in some situations. Where we adjust the marginal in the first stages, then, in stage two we fix the marginal parameters and estimate the parameter of the copula.

In order to visualize the Gaussian copula log-likelihood functions, Fig. 3 shows the visualization of the logarithmic likelihood function used to estimate bivariate Gaussian copula parameters under right artificial censoring data. By the way, we include the MLE, confidence interval, and the true value for the copula parameter.

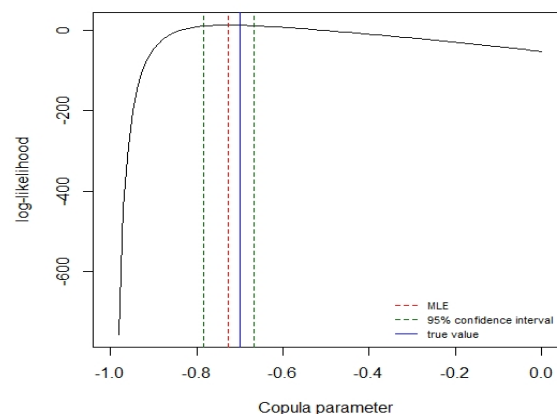


Fig. 3. Fitting a bivariate Gaussian copula under right censoring data.

B. Application for an Archimedean Copula

In this part of the paper, we are focused on an Archimedean copula family, where the data is arbitrarily censored by the same way using a bivariate Gumbel copula with parameter 3 to generate artificial data.

Assuming that the first marginal is distributed under normal distribution $N(\text{mean} = 5, \text{sd} = 0.8)$, and the second marginal is distributed under a Weibull distribution with parameters (location=2.4 and scale=1.3).

Then, the same procedure will apply to the rest of the programs we saw previously, see appendix.

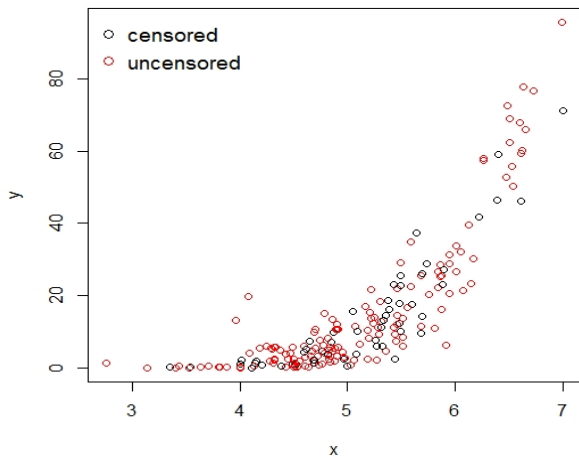


Fig. 4. Censored and incensored artificial data for a bivariate Gumbel copula.

However, to visualize Gumbel's copula log-likelihood functions. Fig. 5 shows the visualization of a logarithmic likelihood function used to estimate bivariate Gumbel copula parameters under censoring observations (arbitrary censoring).

By the way, we are including the MLE, confidence interval, and the true value for the copula parameter.

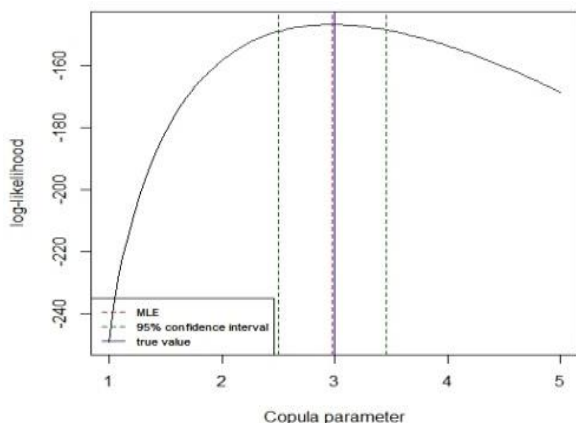


Fig. 5. Fitting a bivariate Gumbel's copula under arbitrary censoring.

When we use the two-stages of the IFM estimation method, the standard error for the copula parameter is normally underappreciated (i.e., too small). The reason for this underappreciated is that the marginal IFM was set during the second step of the adjustment process. Therefore, when estimating the copula parameter at this second stage, it does not consider the margins.

VI. CONCLUSION

Two approaches are generally used for estimating copulas, the parametric and semi-parametric one. In the first case, the resulting estimate of C will be completely parametric, we assume that the marginal belongs to a parameter-indexed family, so it is sufficient to estimate their parameter in order to estimate the margins parametrically. Even, if it will be a semi-parametric approach, we don't assume that the margins belong to any family, so the margins will be estimated non-parametrically.

In this paper, we are interested by the two approaches, where two estimators for this approach are proposed. The maximum-likelihood is called also the MLE method and the inference function method for margins (IFM) which simplifies the estimation procedure when it is impracticable by the MLE method.

Even so, for a semi-parametric approach, the method of moments based on Kendall's tau and Spearman's rho is presented. An extension of the copula moment method, regarded as a generalization of the tau-inverse method, has been presented in a generalized context.

As an application of the MLE estimator, two classes of Archimedean and non-Archimedean copulas were introduced. The first class was chosen because of its number of interesting properties and the efficiency of its results. In addition, the second class was chosen to be compared with the first.

Hence the purpose of this paper is to present a particular class of copulas (Archimedean), their uses, and their approach in statistics. However, their illustration in the R language-focused mainly on the semi-parametric approach, it was discussed in part five of the paper under censored data to introduce the log-likelihood function and the estimation of a copula using two stages consecrated for this method.

One of the perspectives is the modeling of this class of copula-based on the CM's estimator in the case of censored data.

APPENDIX

As an illustration in R software, the key function for the maximum likelihood:

- The illustration of the maximum likelihood estimator to illustrate the Gumbel copula needs at first to assert the function **gumbelCopula()** with the real value of parameter.
- To illustrate the margins as $N(\mu = 0, \sigma^2 = 1)$ and $exp(\lambda = 1)$, it is enough to assert the **mvdc()** function.
- To maximize a loglikelihood function you need to use the **fitMvdc()** for illustration.

The key function for the moment estimator Based on Kendall's Tau and Spearman's Rho:

- The illustration of the moment estimator Based on Kendall's Tau and Spearman's Rho needs the same procedure given above for margins, except that the function used to the corresponding copula parameter estimate is **iTau()** and **iRho()** respectively.
- However, the standard error is returned by the function **summary()**.

Application on non-Archimedean copula under censoring data:

- For a censored and incensored artificial data using a bivariate Gaussian copula, we generate the margins distribution by `mvdc(copula=..., c("norm", "myWeibull"), list(list(mean=..., sd=...), list(location=..., scale=...)))`. Where the first marginal is a normal distribution, and the second marginal is a Weibull distribution.
- To visualize the copula: we draw sample, inspect the generated data and we apply random censoring to the observations by select `rmyWeibull()` function, then `ifelse(xy[,2]> censored, 0, 1)` where `xy[,2] <- pmin(xy[,2], censored)`. Hence, the censored observations will subsequently be treated as right censored cases.
- To plot the censored data we assert the `plot()` function, as well as, we choose the legend and the col.

To fit a bivariate Gaussian copula to the artificial data by means of the two-stage estimation method.

- Stage 1: fit the marginal's, and obtain the marginal parameters by:

```
(normMLE<- fitdistr())
normMean1 <- normMLE$estimate
normSD1 <- normMLE$estimate
(weibullMLE<- survreg())
weibullLocation2 <- coef(weibullMLE)
weibullScale2 <- weibullMLE$scale
Compute the cumulative probabilities for the
marginal's by cbind(pnorm()),
pmyWeibull() function.
```

- Stage 2: fix the parameters of the marginal's and estimate the copula parameter by:

```
Starting value for the copula parameter:
tauEst<- sin(cor(..., ..., method="kendall")*pi/2)).
fitGaussian<- optim(tauEst, loglikCensored,
method="BFGS", copula=...,)
Apply MLE for the copula parameter using
fitGaussian$par function.
```

- To visualize the logarithmic likelihood function used to estimate the parameter of the copula, we need to select the Kendall's tau value `taus<- seq(-1, 0, length.out=100)`.
- By the way, when data is censored the R key for copula is `loglikCensored()`, `liks<- sapply(taus, function() loglikCensored(param=..., copula=...))`. Then we call the function `plot(taus, liks, type=..., xlab="...", ylab="...")`.
- To include the MLE for the copula parameter we declare `abline(v=fitGaussian$par)` function.
- To include the confidence interval for the copula parameter we declare `abline(v=ci)` function.
- To include the true value for the copula parameter we assert `abline(v=myCop@parameters)` function.
- To add a legend we assert `legend()` function.

Application for an Archimedean copula under censoring:

By the same way, using in non-Archimedean copula we just need to change the Gaussian copula by the corresponding Archimedean copula.

ACKNOWLEDGMENT

The authors would like to extend their gratitude to the editor of the journal and to the reviewers for their valuable advice.

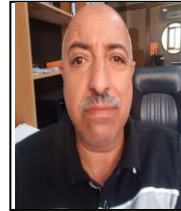
REFERENCES

- [1] Benatia. F & al. A semi parametric estimation procedure for multi parameter Archimedean copulas based on the L-moments method. 2011. Africa statistika, Numero: 335-345.
- [2] Benatia, F, Chine, A. Bivariate copulas estimation using the trimmed L-moments methode. 2017. Africa Statistika, Numéro: 1185-1197.
- [3] Brahimi, B et al. A semiparametric estimation of copula models based on the method of moments. 2012. Statistical Methodology Volume 9, Issue 4, July 2012, Pages 467-477.
- [4] Elamir, E.A.H, SeheultA, H. Trimmed L-moments 2003. Computational statistics. Numéro: 299-314.
- [5] Genest, C and Favre, A. Everything you always wanted to know about copula modeling but were afraid to ask. 2007. Journal of Hydrologic Engineering, 12(4),347-368.
- [6] Genest, C and Werker, B. J. M. Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models. In C. M. Cuadras and J. A. R. Lallena, editors, Proceedings of the Conference on Distributions with Given Marginals and Statistical Modelling, pages 103-112, Kluwer Academic Publishers, Dordrecht, 2002.
- [7] Genest, C, Ghoudi, K. and Rivest, L. P. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika, 82(3):543-552, 1995.
- [8] Genest, C and Rivet, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. Journal of the American Statistical Association 88, 1034-43.
- [9] Genest, C, MacKay, J (1986). The Joy of Copulas: Bivariate distributions with Uniform Marginals. The American Statistician, 40, 280-283.
- [10] Kojadinovic, I, Yan, J (2010). Modeling Multivariate Distributions with Continuous Margins Using the copula R Package. Journal of Statistical Software.
- [11] Lehmann, E. L., & Casella, G. (1998). Theory of point estimation. New York: Springer.
- [12] Mari, D. D. and S. Kotz (2001). Correlation and Dependence. London: Imperial College Press.
- [13] Marius, H, Martin, M, Alexander J.M. Likelihood inference for Archimedean copulas in high dimensions under known margins. Journal of Multivariate analysis, September 2012.
- [14] Nelsen R. B., 2006. An Introduction to Copulas, second ed. Springer, New York.
- [15] Oakes D., 1982. A model for association in bivariate survival data. J. Roy. Statist. Soc. Ser. B 44, no. 3, 414-422.
- [16] Scarsini M (1984). On Measures of Concordance Stochastica, 8(3), 201 [218].
- [17] Tsukahara, H., 2005. Semiparametric estimation in copula models. Canad. J. Statist. 33, 357-375.
- [18] Valdez E. A. (2001). Copula Models for Sums of dependent Risks, School of Actuarial Studies, University of New South Wales.
- [19] Xu. J.J, Statistical Modelling and Inference for Multivariate and Longitudinal Discrete Response Data. Ph.D. Thesis, Department of Statistics, University of British Columbia, 1996.
- [20] Yan, J (2007). Enjoy the Joy of Copulas: With a Package copula. Journal of Statistical Software.
- [21] Fermanian, J.-D., & Scaillet, O. (2005). Some statistical pitfalls in copula modelling for financial applications. In E. Klein (Ed.), Capital formation, governance and banking (pp. 59-74). Hauppauge, NY: Nova Science.
- [22] Kim G., Silvapulle, M. J., & Silvapulle P. (2007). Comparison of semiparametric and parametric methods for estimating copulas. Computational Statistics & Data Analysis, 51(6), 2836-2850.
- [23] Genest C., 1987. Frank's family of bivariate distributions. Biometrika, 74, 549-555.

- [24] Hoeffding, W. A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, 19(3) :293–325, 1948.
- [25] Hofert, M et al. 2017. *Elements of Copula Modeling* with R. Springer.
- [26] Hofert, M. (2008). Sampling Archimedean copulas. *Computational Statistics & Data Analysis*, 52, 5163–5174.
- [27] Hofert, M (2010). Sampling Nested Archimedean Copulas with Applications to CDO Pricing. N978-3-8381-1656-3. PhD thesis.
- [28] Joe, H., 2005. Asymptotic efficiency of the two-stage estimation method for copula-based models. *J. multivariate Anal.* 94, 401-419.
- [29] Joe, H. 1997. *Multivariate Models and Dependence concepts*, Chapman & Hall, London.
- [30] Jouini, M. N. and R. T. Clemen (1996). Copula models for aggregating expert opinions. *Operations Research* 44, 444–57.
- [31] Kimberling, CH (1974). A Probabilistic Interpretation of Complete Monotonicity. *Aequationes Mathematicae*, 10, 152{164}.
- [32] Klaassen, C. A. J. and J. A. Wellner. Efficient estimation in the bivariate normal copula model: normal margins are least favourable. *Bernoulli*, 55–77, 1997.



Nesrine Idiou Postgraduate student preparing a Ph.D. degree affiliated to the Applied Mathematics Laboratory of Mathematics Department, Mohamed Khider University, Biskra, Algeria. Graduated and have a master's degree in Mathematics, on the option of economics and finance applied mathematics (2017) from the University of Mentouri Constantine1, Constantine, Algeria. A statistician research scholar, her research interests focus on copulas estimation and their applications for complete and incomplete data. She has presented several national and international conference papers. She has published research papers in international journals.



Fatah Benatia Currently working as a professor at a department of mathematics and a member in a laboratory of applied mathematics, Mohamed Khider University, Biskra, Algeria. Received his D.E.A from Pierre et Marie Curie University, Jussieu, Paris 6, France. Have Academic Master in Applied Mathematics from University of Constantine 1, Constantine, Algeria. Have a Ph.D. in Applied Mathematics (Statistics and probability) from Mohamed Khider University, Biskra, Algeria. Associated researcher at Mac Gill University, Montreal, Canada (2001). His research interests are copula estimation for complete and incomplete data. He has published several research papers in various international journals and supervised many master's and doctoral theses.