REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000-000

ON CONSTRUCTION OF BERNSTEIN-BÉZIER TYPE BIVARIATE ARCHIMEDEAN COPULA

Authors: Selim Orhun Susam
– Department of Statistics, Dokuz Eylul University, Izmir, Turkey (orhun.susam@deu.edu.tr)

Burcu Hudaverdi
– Department of Statistics, Dokuz Eylul University, Izmir, Turkey (burcu.hudaverdi@deu.edu.tr)

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

- In this paper, a new class of bivariate multi-parameter Archimedean copula based on Kendall distribution using Bernstein-Bézier polynomials is introduced. The new class copula has flexible dependence properties depending on the polynomial degree and the control points. Some dependence characteristics such as Kendall's tau, upper tail and lower tail dependence of the new Archimedean copula class are derived. The simulation procedure based on these desired dependence characteristics is presented. Also, a parameter estimation process based on minimum Cramér-von-Mises distance is also given and its estimation performance is investigated through Monte Carlo simulation study.

Key-Words:

- *Archimedean copula; Kendall distribution; Bernstein-Bézier polynomials; Kendall's tau; Tail-dependence coefficients.*

AMS Subject Classification:

- 62G05, 60E05.

1. INTRODUCTION

Copula models are popular tools for describing multivariate data where the univariate distribution functions are combined to joint distribution function by Sklar's theorem (Sklar, 1959). Let X and Y be random variables with joint distribution function H and the marginal distribution functions F and G, respectively. Then, there exists a copula C such that $H(x, y) = C(F(x), G(y))$, for all x, y in $\Re$. As an advantage of the copula models, the dependence structure can be modelled separately from the marginal distributions. If F and G are continuous, then C is unique. Otherwise, the copula C is uniquely determined on $Ran(F) \times Ran(G)$. There are various families of copulas. One of the most popular families is Archimedean copula family of which the dependence structure can be characterized by an univariate distribution function (Nelsen, 2006, Section 4). The important feature that separates this class from the others is that it has a generator function φ which is used to construct an Archimedean copula.

Definition 1. A generator function φ is a continuous, strictly decreasing convex function defined from $\mathbf{I}$ to $[0, \infty)$ such that $\varphi(1) = 0$. If $\varphi(0) = \infty$, then the generator is called as a strict generator. The pseudo inverse of φ is the function $\varphi^{[-1]}$, defined on $[0, \infty)$ to $\mathbf{I}$ is given by

$$\varphi^{[-1]} = \begin{cases} \varphi^{-1}(t) & 0 \leq t \leq \varphi(0) \\ 0 & \varphi(0) \leq t < \infty \end{cases}$$

A bivariate Archimedean copula with generator function $\varphi, C : \mathbf{I}^2 \to \mathbf{I}$ is defined by

$$C(u, v) = \varphi^{[-1]}\{\varphi(u) + \varphi(v)\}. \tag{1.1}$$

where $u = F(x)$ and $v = G(y)$.

An Archimedean copula function can be reduced to an univariate distribution function through generator function. Genest et al. (1993) showed that the function $\varphi(t)$ can be obtained by the univariate distribution function $K(t) = Pr(C(u, v) \leq t)$. Remarkably, there is a link between the function $\varphi(t)$ and $K(t)$ such as

$$K(t) = t - \frac{\varphi(t)}{\varphi'(t)} = t - \lambda(t). \tag{1.2}$$

$K(t)$ called as Kendall distribution function identifies the generator function $\varphi(t)$ and so the dependence structure of the Archimedean copula family. Dependence measures such as Kendall's tau, upper and lower tail dependence coefficients can be obtained by using Kendall distribution function. For a bivariate Archimedean copula with Kendall distribution function $K(t)$, Genest and MacKay (1986) defined Kendall's Tau (τ) as

$$\tau = 3 - 4 \int_0^1 K(t)dt. \tag{1.3}$$

And also, Michiels et al. (2011) defined lower λ_L and upper λ_U tail dependence as

$$\lambda_L = 2^{\lim_{t \to 0^+} \left(t - K(t)\right)'}, \tag{1.4}$$

$$\lambda_U = 2 - 2^{\lim_{t \to 1^-} \left(t - K(t)\right)'} . \tag{1.5}$$

and they investigated a general method for constructing bivariate Archimedean copula families using λ function. They worked with polynomials to construct multi-parameter copula families. Genest et al. (1998) proposed several ways to generate bivariate Archimedean copula models via smooth transformations of existing generator function. Dimitrova et al. (2008) defined an estimation method of Kendall distribution using B-spline functions. In addition, they defined sufficient conditions for the B-spline estimator to possess the properties of the Kendall distribution function. So, the function can be considered as a proper Kendall distribution function and associated with the multivariate Archimedean copula. Cooray (2018) introduced two-parameter strict Archimedean generator function based on Clayton copula. Najjari et al. (2014) constructed a new generator function $\varphi(t)$ using hyperbolic functions as generators of Archimedean copulas. The majority of the papers proposed some methods based on generator function φ for constructing a new Archimedean family of copulas. In this study, we propose constructing a multi-parameter Archimedean copula using Kendall distribution function $K(t)$. We use Bernstein-Bézier polynomials to create the new Archimedean class. Kendall's tau, lower and upper tail dependence coefficients are also obtained according to the polynomial degree and the control points. This new multi-parameter Archimedean copula family is contributed to the expansion of the existing Archimedean copula family.

The contribution of this study is two fold: First, a new Archimedean copula class based on Bernstein-Bézier polynomial is proposed. Different values of Kendall's tau (negative or positive), lower and upper tail dependence coefficients can be obtained by changing the polynomial degree and the control points, so the proposed class has flexible dependence structure. It is possible to create a new distribution function which has desirable dependence characteristics. This is quite useful in power analysis of goodness-of-fit test statistic. Second, an algorithm is proposed to create different distributions with the same dependence level by changing the control points for poynomial degree. Also, an estimation process based on minimizing Cramér-von Mises distance is presented and a Monte Carlo simulation study is employed to measure the performance of the parameter estimates.

The rest of the paper is organized as follows. In Section 2, Bernstein-Bézier type Archimedean copula is given and some dependence characteristics are investigated. A simulation procedure of this new class for different polynomial degrees is given in Section 3. Parameter estimation procedure which is based on minumum Cramér-von-Mises measure is given and parameter estimates are obtained in Section 4. And the last section is devoted to the conclusion.

2. BERNSTEIN BÉZIER TYPE BIVARIATE ARCHIMEDEAN COPULA

A Kendall distribution function, $K(t)$ should satisfy the following properties (1-4) described in Nelsen (2006).

1. $K(0) = 0$
2. $K(1) = 1$
3. $K^{'}(t) > 0$
4. $K(t) > t\ ,\ \ t \in (0,1)$

Let $K(m,\alpha;t)$ be a Bernstein-Bézier type Kendall distribution function with polynomial degree m and control points α defined as

$$K(m,\alpha;t) = \sum_{k=0}^{m} \alpha_k B_{k,m}(t) \tag{2.1}$$

where $B_{k,m}(t) = \binom{m}{k} t^k (1-t)^{m-k}$ for $t \in [0,1]$.

Lemma 2.1. *A Bernstein-Bézier type Kendall distribution function $K(m,\alpha;t)$ satisfies the properties (1-4) if the following constraints hold:*

1. $\alpha_0 = 0 < \alpha_1 < \alpha_2 < \ldots < \alpha_m = 1$

2. $\alpha_k > \frac{k}{m}\ ,\ k = 1,\ldots,m-1$

Proof: $K(m,\alpha,t=0,) = \sum_{k=0}^{m} \alpha_k B_{k,m}(t=0) = 0$ holds since $\alpha_0 = 0$. Similarly, $K(m,\alpha,t=1) = \sum_{k=0}^{m} \alpha_k B_{k,m}(t=1) = 1$ holds since $\alpha_m = 1$.

Also, $K(m,\alpha,t)^{'} = m \sum_{k=0}^{m-1} (\alpha_{k+1} - \alpha_k) P_{k,m-1}(t) \geq 0$. See, Duncan (2005). So, $\alpha_0 = 0 < \alpha_1 < \alpha_2 < \ldots < \alpha_m = 1$.

If the Bézier control points $\alpha_k > \frac{k}{m}, k = 1,\ldots,m-1$ where $\alpha_k = k/m + \epsilon_k$, then,

$$
\begin{aligned}
K(m,\alpha,t) &= \sum_{k=0}^{m} \alpha_k \binom{m}{k} t^k (1-t)^{m-k} \\
&= \sum_{k=0}^{m} (\frac{k}{m} + \epsilon_k) \binom{m}{k} t^k (1-t)^{m-k} \\
&= \sum_{k=0}^{m} (\frac{k}{m}) \binom{m}{k} t^k (1-t)^{m-k} + \sum_{k=0}^{m} (\epsilon_k) \binom{m}{k} t^k (1-t)^{m-k} \\
&= t \sum_{k=1}^{m} \binom{m-1}{k-1} t^{k-1} (1-t)^{m-k} + \sum_{k=0}^{m} (\epsilon_k) \binom{m}{k} t^k (1-t)^{m-k} \\
&= t \sum_{p=0}^{m-1} t^p (1-t)^{m-p-1} \binom{m-1}{p} + \sum_{k=0}^{m} (\epsilon_k) \binom{m}{k} t^k (1-t)^{m-k} \\
&= t + \sum_{k=0}^{m} (\epsilon_k) \binom{m}{k} t^k (1-t)^{m-k} > t.
\end{aligned}
$$

□

We also obtain Kendall's tau, lower and upper tail dependence of the Bernstein-Bézier type Archimedean copula class using the following lemmas.

Lemma 2.2. *Kendall's tau for Bernstein-Bézier type Archimedean copula is obtained as*

$$\tau = 3 - 4 \sum_{k=0}^{m} \alpha_k \binom{m}{k} \beta(k+1, m-k+1)$$

where $\beta(.,.)$ *is the beta function defined as* $\beta(v_1, v_2) = \int_0^1 t^{v_1 - 1}(1-t)^{v_2-1} dt$ *for* v_1, v_2 *positive integers.*

Proof: τ is easily derived from equation $\tau = 3 - 4 \int_0^1 K(t) dt$. □

Lemma 2.3. *The lower tail* λ_L *and the upper tail* λ_U *dependence for Bernstein-Bézier type Archimedean copula are obtained by*

$$
\begin{aligned}
\lambda_L &= 2^{1-m\alpha_1} \\
\lambda_U &= 2 - 2^{1-m(1-\alpha_{m-1})}
\end{aligned}
$$

Proof: λ_U and λ_L are easily derived from equation $\lambda_L = 2^{\lim_{t\to 0^+}\left(t-K(t)\right)'}$, $\lambda_U = 2 - 2^{\lim_{t\to 1^-}\left(t-K(t)\right)'}$. □

It is seen that λ_L and λ_U are affected by only the control points α_1 and α_{m-1}, respectively. We can create Bernstein-Bézier type Archimedean copula using λ_L and λ_U, setting up the control points α_1 and α_{m-1}.

The following inequalities given in the next lemma provide an information for proper selection of λ_U and λ_L.

Lemma 2.4. *Let λ_L and λ_U be lower and upper tail dependence of Bernstein-Bézier type Archimedean copula with polynomial degree m. Then,*

$$1 > \lambda_L > \frac{2^{2-m}}{2-\lambda_U}$$

holds for all values of polynomial degree m.

Proof: It can ve proved using the inequality $\alpha_1 < \alpha_{m-1}$. Also, $0 < \lambda_U, \lambda_L < 1$, see Charpentier and Segers (2008). □

Suppose that the parameters α_k are defined as $\alpha_k > \frac{k}{m}$ for for $k = 1, \ldots, m-1$, then $K(m, \alpha; t) > t$. See, Lemma 2.1. Also, we note that if the control points are selected as $\alpha_k \to \frac{k}{m}$, then the dependence coefficients $(\tau,\ \lambda_U,\ \lambda_L)$ approximate 1. In other words, the Bernstein-Bézier type Archimedean copula approximates comonotonic dependence when the control points are closely distributed uniform.

The Bernstein-Bézier type Archimedean copula with higher degree can represent various dependence forms. However, they may have some disadvantages.

1. As the degree increases, the complexity and therefore the processing time increase.
2. Because of the complexity, the curves of higher degree are more sensitive to round off errors.

As opposed to these disadvantages, we can combine several Bernstein-Bézier type Kendall distribution functions, mostly of degree three and four. We note that the Bernstein- Bézier polynomials are invariant under barycentric combinations (Farin (2001), p. 61). So, we obtain the following Bernstein-Bezier type Archimedean copulas for $\theta \in [0,1]$:

$$\begin{aligned} K(m,\alpha;t) &= \sum_{k=0}^{m} \big(\theta\alpha_{1,k} + (1-\theta)\alpha_{2,k}\big) B_{k,m}(t) \\ &= \theta \sum_{k=0}^{m} \alpha_{1,k} B_{k,m}(t) + (1-\theta) \sum_{k=0}^{m} \alpha_{2,k} B_{k,m}(t) \\ &= \theta K(m,\alpha_{1,.};t) + (1-\theta) K(m,\alpha_{2,.};t). \end{aligned}$$

We can construct the weighted average of two Bernstein-Bézier Archimedean copulas either by taking the weighted average of corresponding points on the distribution, or by taking the weighted average of corresponding parameters α.

Dependence coefficients of two barycentric combinations of Bernstein-Bézier type Archimedean copula are given by

$$\begin{aligned} \tau =& 3 - 4\sum_{k=0}^{m} \alpha_{2,k} \beta(k+1, m-k+1) \binom{m}{k} \\ &+ 4\theta \Big(\sum_{k=0}^{m} (\alpha_{2,k} - \alpha_{1,k}) \beta(k+1, m-k+1) \binom{m}{k} \Big) \end{aligned}$$

$$\lambda_U = 2 - 2^{1+\theta m\alpha_{1,m-1} + (1-\theta) m \alpha_{2,m-1} - m}$$

$$\lambda_L = 2^{1-\left(\theta m \alpha_{1,1} + (1-\theta) m \alpha_{2,1}\right)}$$

Note that if θ is selected as 1, then the classical Bernstein-Bézier type Archimedean copula is obtained.

3. SIMULATING DATA FROM BERNSTEIN BÉZIER TYPE ARCHIMEDEAN COPULA

In this section, data simulation from Bernstein-Bézier type Archimedean copula is given. Construction of a new distribution function which has desirable Kendall's tau and tail dependence coefficients are investigated.

The following procedure is used to create a distribution with the dependence characteristics represented by Kendall's tau and tail dependence coefficients.

1. The arbitrary value of the upper tail dependence λ_U is determined primarily.
2. λ_L is determined arbitrarily by using Lemma 2.4.
3. The value of Kendall's tau τ is determined for the distributions with polynomial degrees 2 and 3. For the distributions having polynomial degree $m \geq 4$, an interval of Kendall's tau is determined. Then, Kendall's tau is selected arbitrarily from this interval.
4. Bivariate data is simulated using the following algorithm. See, Nelsen (2006).

The algorithm based on Michiels et al. (2011) allows one to simulate $C(u, v)$ by Kendall distribution function $K(t)$ given as;

- Simulate uniformly distributed random pair (s, t) on $[0, 1]$.
- Set $w = K^{-1}(t)$.
- Set u such that $\int_w^u \frac{1}{t-K(t)} dt - \ln(s) = 0$.
- Set v such that $\int_w^v \frac{1}{t-K(t)} dt - \ln(1 - s) = 0$.

The range of the parameters and the dependence coefficients depending on the Bernstein-Bézier polynomial degree m are summarized in Table 1. It is observed that as the degree of the polynomial increases, the range of the dependence coefficients gets wider.

Kendall's tau, upper and lower tail dependence coefficients obtained by the Bernstein-Bézier type Archimedean copula with control points for degree ($m = 3, 4, 5$) are summarized in Table 2. Also, different distributions having the same dependence level at the control points α_2 and α_3 for poynomial degree 5 are given. All the Bernstein-Bézier control points and dependence coefficients are obtained by applying the simulation procedure (1-4). All cases in Table 2 are examined in the subsection (3.1-3.3).

m	α_0	α_1	α_2	α_3	α_4	α_5	τ	λ_U	λ_L
3	0	$(\frac{1}{3}, 1)$	$(\max(\frac{2}{3}, \alpha_1), 1)$	1	-	-	$(0, 1)$	$(0, 1)$	$(\frac{1}{4}, 1)$
4	0	$(\frac{1}{4}, 1)$	$(\max(\frac{2}{4}, \alpha_1), 1)$	$(\max(\frac{3}{4}, \alpha_2), 1)$	1	-	$(-0.2, 1)$	$(0, 1)$	$(\frac{1}{8}, 1)$
5	0	$(\frac{1}{5}, 1)$	$(\max(\frac{2}{5}, \alpha_1), 1)$	$(\max(\frac{3}{5}, \alpha_2), 1)$	$(\max(\frac{4}{5}, \alpha_3), 1)$	1	$(-0.33, 1)$	$(0, 1)$	$(\frac{1}{16}, 1)$

Table 1: Range of parameters and dependence coefficients

Degree	$K(t)$	α_0	α_1	α_2	α_3	α_4	α_5	τ	λ_U	λ_L
$m = 3$	K_1	0	0.7173	0.7928	1	-	-	0.4899	0.7	0.45
$m = 4$	K_2	0	0.3537	0.5828	0.9815	1	-	0.68	0.1	0.75
$m = 5$	K_3	0	0.4	0.43	0.8531	0.9169	1	0.6	0.5	0.5
	K_4	0	0.4	0.63	0.6531	0.9169	1	0.6	0.5	0.5

Table 2: Parameters and dependence coefficients

3.1. Bernstein-Bézier type Archimedean copula with degree three

A Bernstein-Bézier type Archimedean copula with degree 3 has the following distribution function,

$$K(m=3,\alpha;t)=\sum_{k=0}^{3}\alpha_k\binom{3}{k}t^k(1-t)^{3-k}, t\in[0,1]$$

From Lemma 2.1, $\alpha_0 = 0, \alpha_3 = 1$, $\alpha_0 < \alpha_1 < \alpha_2 < \alpha_3$ and $\alpha_1 > \frac{1}{3}, \alpha_2 > \frac{2}{3}$. Kendall's tau of the distribution is given as

$$\tau = 3-4\sum_{k=0}^{3}\alpha_k\binom{3}{k}\beta(k+1,3-k+1)=2-\alpha_1-\alpha_2$$

and lower and upper tail dependence coefficients are

$$\lambda_L = 2^{1-3\alpha_1}, \lambda_U = 2-2^{3\alpha_2-2}$$

$(1-4)$ procedure is applied to determine the Kendall's tau and the tail dependence coefficients of the distribution. The arbitrary value of the upper tail dependence λ_U is determined primarily in the range $\lambda_U \in (0,1)$. We select λ_U as 0.7, so α_2 is equal to 0.7928 . From Lemma 2.4, $1 > \lambda_L > 0.3846$. Then, λ_L is determined arbitrarily as 0.45. So, α_1 is equal to 0.7173. The stage conditions for control points given Lemma 1 are satisfied. Finally, Kendall's tau is 0.4899. $K(3,\alpha;t)$ with control points $\alpha_0 = 0, \alpha_1 = 0.7173, \alpha_2 = 0.7928$ and $\alpha_3 = 1$ has the Kendall's tau value as $\tau = 0.4899$ and the value tail dependence coefficients as $\lambda_L = 0.45$ and $\lambda_U = 0.7$. Simulated data and $K(m=3,\alpha;t)$ with the sample of size 150 are visualized in Figure 1.

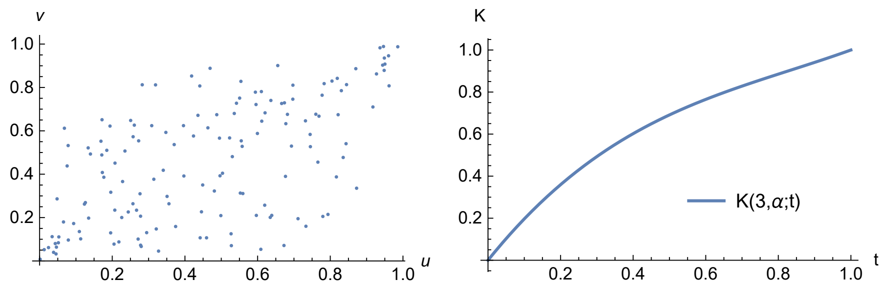

Figure 1: Simulated data from $K(3,\alpha;t)$ with $\tau = 0.4899, \lambda_L = 0.45, \lambda_U = 0.7$

3.2. Bernstein-Bézier type Archimedean copula with degree four

Bernstein-Bézier type Archimedean copula with degree 4 has the following distribution function with the dependence characteristics, Kendall's tau, lower and upper tail dependence,

$$K(4,\alpha;t) = \sum_{k=0}^{4} \alpha_k \binom{4}{k} t^k (1-t)^{4-k}, t \in [0,1]$$

$$\tau = \frac{1}{5}\Bigg(11 - 4(\alpha_1 + \alpha_2 + \alpha_3)\Bigg)$$

$$\lambda_L = 2^{1-4\alpha_1}, \lambda_U = 2 - 2^{4\alpha_3 - 3}$$

(1−4) procedure is applied to determine the Kendall's tau and the tail dependence values of the distribution. The arbitrary value of the upper tail dependence λ_U is determined primarily in range $\lambda_U \in (0,1)$. We select λ_U as 0.1 and so α_3 is equal to 0.9815. From Lemma 2.4, $1 > \lambda_L > 0.1315$. Then, λ_L is determined arbitrarily as 0.75 . So, α_1 is equal to 0.3537. Finally from Lemma 2.1, Kendall's tau should be selected in the range $\tau \in (0.3610, 0.7462)$. We determine Kendall's tau arbitrarily as 0.68. So, α_2 is 0.5828 . $K(4,\alpha;t)$ with control points $\alpha_0 = 0, \alpha_1 = 0.3537, \alpha_2 = 0.5828, \alpha_3 = 0.9815$ and $\alpha_4 = 1$ has the value of Kendall's tau $\tau = 0.68$ and the values of tail dependences as $\lambda_L = 0.75$ and $\lambda_U = 0.1$. Simulated data and $K(m = 4,\alpha;t)$ with the sample of size 150 is visualized in Figure 2.

3.3. Bernstein-Bézier type Archimedean copula with degree five

Bernstein-Bézier type Archimedean copula with degree 5 has the following distribution function with the dependence characteristics Kendall's tau, lower

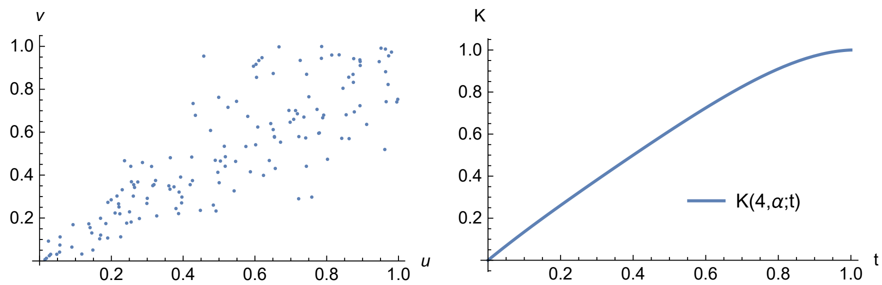

Figure 2: Simulated data from $K(4, \alpha; t)$ with $\tau = 0.68, \lambda_L = 0.75, \lambda_U = 0.1$

and upper tail dependence,

$$K(5, \alpha; t) = \sum_{k=0}^{5} \alpha_k \binom{5}{k} t^k (1-t)^{5-k}, t \in [0, 1]$$

$$\tau = \frac{1}{3}\Big(7 - 2(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)\Big)$$

$$\lambda_L = 2^{1-5\alpha_1}, \lambda_U = 2 - 2^{5\alpha_4 - 4}$$

$(1-4)$ procedure is again applied to determine the Kendall's tau and the tail dependence values of the distribution. The arbitrary value of the upper tail dependence λ_U is determined primarily in range $\lambda_U \in (0, 1)$. We select λ_U as 0.5 and so α_4 is equal to 0.9169. From Lemma 2.4, $1 > \lambda_L > 0.0833$. Then, λ_L is determined arbitrarily as 0.5. So, α_1 is equal to 0.4. Finally from Lemma 2.1, Kendall's tau should be selected in the range $\tau \in (0.2328, 0.6220)$. We determine Kendall's tau arbitrarily as 0.6. α_2 and α_3 can be derived from solving equations $\alpha_2 + \alpha_3 = 1.2831$. From the last equation and Lemma 2.1, α_2 and α_3 should be selected in the range $\alpha_2 \in (0.4, 0.6415)$ and $\alpha_3 \in (0.6415, 0.8831)$, respectively. Different α_2 and α_3 values can be selected in order to provide $\alpha_2+\alpha_3 = 1.2831$ in the range of α_2 and α_3. This case is important, because we can create different distributions with the same dependence level by selecting different α_2 and α_3 values. One possible selection is $\alpha_2 = 0.43$ and $\alpha_3 = 0.8531$. Another possible selection is $\alpha_2 = 0.63$ and $\alpha_3 = 0.6531$. $K_1(5, \alpha; t)$ with control points $\alpha_0 = 0, \alpha_1 = 0.4, \alpha_2 = 0.43, \alpha_3 = 0.8531, \alpha_4 = 0.9169, \alpha_5 = 1$ and $K_2(5, \alpha; t)$ with control points $\alpha_0 = 0, \alpha_1 = 0.4, \alpha_2 = 0.63, \alpha_3 = 0.6531, \alpha_4 = 0.9169, \alpha_5 = 1$ with the same dependence level are visualized in Figure 3.

For the higher order polynomial degree, for example $m = 6$, the range of τ, λ_L and λ_U are determined as the same as for degree $m < 6$. But the range of α_2, α_3 and α_4 for the solutions of $\alpha_2 + \alpha_3 + \alpha_4 = a$ cannot be determined easily.

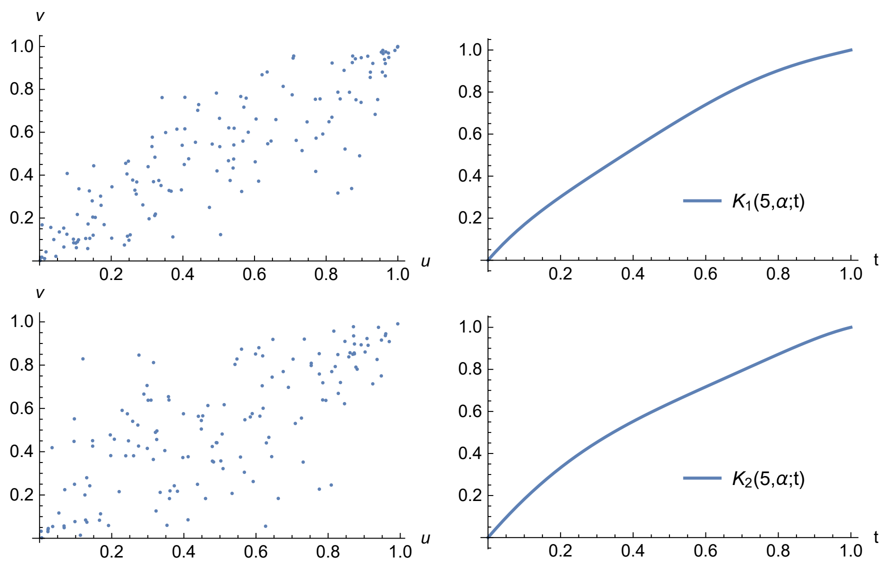

Figure 3: Simulated data from $K_1(5, \alpha; t)$ and $K_2(5, \alpha; t)$ with the same $\tau = 0.6, \lambda_L = 0.5, \lambda_U = 0.5$

4. PARAMETER ESTIMATION BASED ON CRAMÉR-VON-MISES MEASURE

Genest and Rivest (1993) proposed a nonparametric procedure using empirical estimate K_n of K. The psuedo observations of $\hat{T}_i$ were obtained by

$$\hat{T}_i = \sum_{j=1}^{n} I(X_i < X_j, Y_i < Y_j)/(n-1), i = 1, \dots, n.$$

Then, $K(t)$ was estimated by the empirical distribution function as

$$\hat{K}_n(t) = \sum_{i=1}^{n} (\hat{T}_i \leq t)/n. \tag{4.1}$$

Barbe et al. (1996) investigated consistency of $\hat{K}_n(t)$. Alternatively, Susam and Ucer (2018) defined the empirical Bernstein estimator of order $(m_1 > 0)$ for the Kendall distribution function as,

$$\hat{K}_{m_1,n}(t) = \sum_{k=0}^{m_1} \hat{K}_n(k/m_1) P_{k,m_1}(t) \tag{4.2}$$

where $P_{k,m_1}(t) = \binom{m_1}{k} t^k (1-t)^{m_1-k}$ is the binomial probability. Also, they showed that the Bernstein Kendall distribution function outperforms the empirical Kendall distribution function according to its performance by Monte Carlo simulation study.

In this study, through the parameter estimation process, we first estimate the Bernstein-Bézier type Archimedean copula parameters by using empirical estimate of $\hat{K}_n$. Then, Cramér-von-Mises (CvM) distance between the empirical Kendall distribution function and the Bernstein-Bézier type Kendall distribution function is obtained as

$$\begin{aligned} CvM_{\hat{K}_n} &= \int_0^1 n\big(\hat{K}_n(t) - K(\alpha, m_2; t)\big)^2 d\hat{K}_n(t) \\ &= \frac{1}{n}\sum_{i=1}^{n} \big(\hat{K}_n(\hat{T}_i) - K(\alpha, m_2; \hat{T}_i)\big)^2. \end{aligned}$$

Then the parameters are estimated by

$$\hat{\alpha}_{\hat{K}_n} = \underset{\alpha \in \Theta}{\operatorname{argmin}} \left\{ CvM_{\hat{K}_n} \right\}$$

where $\Theta = \left\{ \alpha_k > \frac{k}{m_2},\, \alpha_{k+1} > \alpha_k \,; k = 1, \ldots, m_2 - 1 \right\}$ and $\alpha_0 = 0$, $\alpha_{m_2} = 1$.

Secondly, the Bernstein-Bézier type Archimedean copula parameters are estimated by using empirical Bernstein estimator $\hat{K}_{m_1,n}(t)$. Since the empirical Bernstein Kendall distribution function is a continuous approximation of the empirical Kendall distribution function $\hat{K}_n$, we use empirical Bernstein Kendall distribution function which is upgraded version of $\hat{K}_n$ to obtain Cramér-von-Mises (CvM) distance as

$$CvM_{\hat{K}_{n,m}} = \int_0^1 n\big(\hat{K}_{n,m_1}(t) - K(\alpha, m_2; t)\big)^2 dt. \tag{4.3}$$

The estimation of the dependence parameter α_i for $i = 0, \ldots, m_2$ can be selected as the value that minimizes the CvM distance.

Lemma 4.1. *Let $K(\alpha, m_2; t)$ be the Bernstein-Bézier type Kendall distribution function with order ($m_2 > 0$) and let $\hat{K}_{m,n}(t)$ be the empirical Bernstein estimator of Kendall distribution function with order ($m_1 > 0$). Then the Cramér-von-Mises distance is defined as*

$$
\begin{aligned}
CvM &= n\sum_{k=0}^{m_1}\binom{m_1}{k}^2\hat{K}_n^2(\frac{k}{m_1})\beta(2k+1,2m_1-2k+1)\\
&+2n\sum_{k=0}^{m_1-1}\sum_{s=k+1}^{m_1}\binom{m_1}{k}\binom{m_1}{s}\hat{K}_n(\frac{k}{m_1})\hat{K}_n(\frac{s}{m_1})\beta(k+s+1,2m_1-k-s+1)\\
&+n\sum_{k=0}^{m_2}\binom{m_2}{k}^2\alpha_k^2\beta(2k+1,2m_2-2k+1)\\
&+2n\sum_{k=0}^{m_2-1}\sum_{s=k+1}^{m_2}\binom{m_2}{k}\binom{m_2}{s}\alpha_k\alpha_s\beta(k+s+1,2m_2-k-s+1)\\
&-2n\sum_{k=0}^{m_1}\sum_{s=0}^{m_2}\hat{K}_n(\frac{k}{m_1})\alpha_s\binom{m_1}{k}\binom{m_2}{s}\beta(k+s+1,m_1+m_2-k-s+1)
\end{aligned}
$$

where $\beta(.,.)$ is the beta function defined as $\beta(v_1,v_2)=\int_0^1 t^{v_1-1}(1-t)^{v_2-1}dt$ for v_1,v_2 positive integers.

Proof:

$$
\begin{aligned}
CvM &= n\int_0^1(\hat{K}_{n,m_1}(t)-K(\alpha,m_2;t))^2dt\\
&= n\int_0^1\hat{K}_{n,m_1}^2(t)dt+n\int_0^1(K(\alpha,m_2;t))^2dt-2n\int_0^1\hat{K}_{n,m_1}(t)K(\alpha,m_2;t)dt\\
&= n\int_0^1\Big(\sum_{k=0}^{m_1}\binom{m_1}{k}t^k(1-t)^{m_1-k}\hat{K}_n(\frac{k}{m_1})\Big)^2dt\\
&+n\int_0^1(\sum_{k=0}^{m_2}\alpha_k t^k\binom{m_2}{k}t^k(1-t)^{m_2-k})^2dt\\
&-2n\sum_{k=0}^{m_1}\sum_{s=0}^{m_2}\hat{K}_n(\frac{k}{m_1})\alpha_s\binom{m_1}{k}\binom{m_2}{s}\int_0^1 t^{k+s}(1-t)^{m_1+m_2-k-s}dt\\
&= I_1+I_2-I_3
\end{aligned}
$$

Now, we calculate part of I_1. We know that, $(a_1+a_2+\ldots+a_n)^2=\sum_{i=1}^n a_i^2+2\sum_{i=1}^{n-1}\sum_{j=i+1}^n a_ia_j$, then we can write

$$
\begin{aligned}
I_1 &= n\sum_{k=0}^{m_1}\binom{m_1}{k}^2\hat{K}_n^2(\frac{k}{m_1})\int_0^1 t^{2k}(1-t)^{2m_1-2k}dt\\
&+2\sum_{k=0}^{m_1-1}\sum_{s=k+1}^{m_1}\binom{m_1}{k}\hat{K}_n(\frac{k}{m_1})\binom{m_1}{s}\hat{K}_n(\frac{s}{m_1})\int_0^1 t^{k+s}(1-t)^{2m_1-k-s}dt\\
&= n\sum_{k=0}^{m_1}\binom{m_1}{k}^2\hat{K}_n^2(\frac{k}{m_1})\beta(2k+1,2m_1-2k+1)\\
&+2n\sum_{k=0}^{m_1-1}\sum_{s=k+1}^{m_1}\binom{m_1}{k}\hat{K}_n(\frac{k}{m_1})\binom{m_1}{s}\hat{K}_n(\frac{s}{m_1})\beta(k+s+1,2m_1-k-s+1)
\end{aligned}
$$

proof of the parts of I_2 and I_3 are the same as proof of part I_1.

$\square$

Then, the parameter estimate which gives the minimum value of Cramér-von-Mises distance based on Bernstein empirical distribution is defined for Bernstein-Bézier type Archimedean copula by

$$\hat{\alpha}_{\hat{K}_{n,m}} = \underset{\alpha \in \Theta}{\operatorname{argmin}} \left\{ CvM_{\hat{K}_{n,m}} \right\}$$

where $\Theta = \left\{ \alpha_k > \frac{k}{m_2},\ \alpha_{k+1} > \alpha_k \ \ ; k = 1, \ldots, m_2 - 1 \right\}$ and $\alpha_0 = 0$, $\alpha_{m_2} = 1$.

Genest et al. (1993) introduced a method-of-moment estimator for bivariate Archimedean copula based on empirical Kendall distribution function $\hat{K}_n(t)$. For one-parameter families, the parameter can be estimated by only using the first moment. However, for more than one parameters, we need the moments as much as the number of parameters.

We note that the estimation procedure explained in this section are not only available for Archimedean copulas but also available for all continuous copula classes. The empirical Kendall distribution function can also be used for all continuous copula classes. See Genest et al. (1993)

A Monte Carlo simulation study is conducted to measure the performance of the estimation method with several values of Kendall's tau, lower and upper tail dependence coefficients.

1.000 Monte Carlo samples of sizes $n = 50, 150$ are generated from each type of Bernstein-Bézier type Archimedean copulas given in Table 2 and investigated the performances of two parameter estimation methods as $\alpha_{\hat{K}_n}$ and $\alpha_{\hat{K}_{n,m}}$. For the empirical Bernstein estimator, we select the polynomial degree as $m_1 = 15$ for sample size $n = 50$ and $m_1 = 30$ for sample size $n = 150$.

Simulation results are shown in Table 3 and Table 4. When the results are examined, the minumum Cramér-von-Mises method based on Kendall distribution using Bernstein polynomials outperforms the method based on empirical Kendall distribution in almost all cases for all sample sizes.

5. CONCLUSION

In this study, we propose a new family of Archimedean copulas based on Kendall distribution function $K(t)$. We use Bernstein-Bézier polynomials to construct this new multi-parameter distribution. The method is illustrated for polynomial degree $m = 3, 4, 5$. There are several advantages of this new Archimedean

Dist.	Est. Mth.	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$
K_1	$\hat{\alpha}_{\hat{K}_n}$	0.00684	0.00431	-	-
	$\hat{\alpha}_{\hat{K}_{n,15}}$	0.00575	0.00313	-	-
K_2	$\hat{\alpha}_{\hat{K}_n}$	0.00903	0.01116	0.00221	-
	$\hat{\alpha}_{\hat{K}_{n,15}}$	0.00324	0.00688	0.00585	-
K_3	$\hat{\alpha}_{\hat{K}_n}$	0.00633	0.01580	0.01428	0.00349
	$\hat{\alpha}_{\hat{K}_{n,15}}$	0.00342	0.00925	0.01192	0.00193
K_4	$\hat{\alpha}_{\hat{K}_n}$	0.01544	0.00957	0.00992	0.00266
	$\hat{\alpha}_{\hat{K}_{n,15}}$	0.00534	0.01422	0.00923	0.00356

Table 3: MSE of the parameter estimations for four Bernstein-Bézier type copula with sample size $n = 50$

Dist.	Est. Mth.	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$
K_1	$\hat{\alpha}_{\hat{K}_n}$	0.00261	0.00151	-	-
	$\hat{\alpha}_{\hat{K}_{n,30}}$	0.00303	0.00141	-	-
K_2	$\hat{\alpha}_{\hat{K}_n}$	0.00209	0.00437	0.00096	-
	$\hat{\alpha}_{\hat{K}_{n,30}}$	0.00123	0.00384	0.00177	-
K_3	$\hat{\alpha}_{\hat{K}_n}$	0.00177	0.00661	0.00827	0.00242
	$\hat{\alpha}_{\hat{K}_{n,30}}$	0.00229	0.00589	0.00614	0.00091
K_4	$\hat{\alpha}_{\hat{K}_n}$	0.00516	0.00775	0.00650	0.00144
	$\hat{\alpha}_{\hat{K}_{n,30}}$	0.00224	0.00753	0.00670	0.00165

Table 4: MSE of the parameter estimations for four Bernstein-Bézier type copula with sample size $n = 150$

copula class. It is shown that while working with the Bernstein-Bézier polynomial structures, a multi-parameter copula family can be constructed in an organized way. It is possible to create a new distribution function which has desirable dependence characteristics using Kendall's tau, lower and upper tail dependence. The parameters of the new model can be interpreted in terms of these dependence characteristics. And also, it is possible that we can create different distributions with the same dependence structures. Also, we obtain the parameter estimates minimizing the Cramér-von-Mises distance which is based on Bernstein-Bézier type Archimedean copulas. We measure the performance of the estimation method with several values of Kendall's tau, lower and upper tail dependence coefficients by a Monte Carlo simulation study. We can conclude that the minimum Cramér-von-Mises method based on Kendall distribution using Bernstein polynomials outperforms the method based on empirical Kendall distribution function.

ACKNOWLEDGMENTS

We thank the anonymous referees and the editor for their helpful suggestions which improved the presentation of the paper.

REFERENCES

[1] Barbe P, Genest C, Ghoudi K, Rémillard B. (1996) On Kendalls Process. *Journal of Multivariate Analysis*, **58**, 197-229.

[2] Charpentier A, Segers J. (2009) Tails of multivariate Archimedean copulas. Journal pf *Multivariate Analysis*,**100**, 1521-1537.

[3] Cooray K. (2018) Strictly Archimedean copulas with complete association for multivariate dependence based on the Clayton family. *Dependence Modelling*, **6**,1-18.

[4] Dimitrova D, Kaishev K, Penev I. (2008) GeD spline estimation of multivariate Archimedean copulas.*Computational Statistics and Data Analysis*, **52**, 3570-3582.

[5] Duncan M. (2005) *Applied Geometry for Computer Graphics and CAD*, Springer Verlag, London.

[6] Farin G. (2001) *Curves and Surfaces for CAGD: A Practical Guide*, Morgan Kaufmann Publishers, San Francisco.

[7] Genest C, Mackay RJ. (1986) Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian journal of Statistics*, 14, 145-149.

[8] Genest C, Rivest L . (1993) Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*,**88**, 423, 1034-1043.

[9] Genest C, Ghoudi K, Rivest L. (1998). Discussion to “Understanding Relationships Using Copulas” by E. Frees and E. Valdez. *North American Actuarial Journal* **2**(1), 143149.

[10] Michiels F, Koch I, De Schepper A. (2011) A New Method for the Construction of Bivariate Archimedean Copulas Based on the λ Function. *Communications in Statistics - Theory and Methods* **40**, 15, 2670–2679.

[11] Najjari V,Bacigal T, Bal H.(2014) An Archimedean copula family with hyperbolic cotangent generator. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*,**20**, 5, 761–768.

[12] Nelsen RB. (2006) *An introduction to copulas*, Springer Verlag, NewYork (NY) .

[13] Sklar A. (1959) Fonctions de répartition á n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*,**8**, 229-231.

[14] Susam SO, Ucer Hudaverdi B. (2018) Testing independence for Archimedean copula based on Bernstein estimate of Kendall distribution function. *Journal of Statistical Computation and Simulation*, **88**, 13, 2589–2599.