Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 51 (2) (2022), 618 – 631
DOI : 10.15672/hujms.993698

Research Article

A multi-parameter Generalized Farlie-Gumbel-Morgenstern bivariate copula family via Bernstein polynomial

Selim Orhun Susam

Department of Econometrics, Munzur University, Tunceli, TURKEY

Abstract

In this paper, we are proposing a flexible method for constructing a bivariate generalized Farlie-Gumbel-Morgenstern (G-FGM) copula family. The method is mainly developed around the function $\phi(t)$ $(t \in [0,1])$, where ϕ is the generator of the G-FGM copula. The proposed construction method has useful advantages. The first of which is the direct relationship between the ϕ function and Kendall's tau. The second advantage is the possibility of constructing a multi-parameter G-FGM copula which allows us to better harmonize empirical instruction with the model. The construction method is illustrated by three real data examples.

Mathematics Subject Classification (2020). 62H10

Keywords. Bernstein polynomial, Copula, Kendall's tau, FGM copula

1. Introduction

When the researchers are interested in statistical modeling in many applications, the main objective is to determine the best fit for observed random variables within its dependence structure. One possible solution is to model the observed data among existing copulas that perform best according to the statistical tests. Another possibility is constructing an ideal copula with multi-parameter that can be estimated from observed data.

In this paper, the contribution is made on the last possibility. For this purpose, we aim to construct a multi-parameter G-FGM copula based on the generator function ϕ. In the literature, there are many papers for constructing G-FGM copula with only one or two parameters. For a review, see [2]. Contrary to existing methods, in this paper, we aim to construct a multi-parameter G-FGM copula family in order to achieve the best-fitted model for observed data according to the goodness of fit tests. Our generator function is developed using Bernstein polynomials which has useful properties. For instance, we may easily determine the shape of Bernstein polynomials and also its derivatives by managing the control points. Moreover, the first and the last points of Bernstein polynomial coincide with the first and last control points. Thus, we can easily define the Bernstein polynomials which satisfy all properties of G-FGM copula generator function. For usage of Bernstein polynomial in various ways in the copula theory, see ([3, 12, 14, 17–21]).

Email address: orhunsusam@munzur.edu.tr (S.O. Susam)
Received: 10.09.2021; Accepted: 13.02.2022

The paper is organized as follows: In Section 2, the basic concept of G-FGM copula, its definition and dependence properties are given. In section 3, the generator function type of G-FGM copula constructed from Bernstein polynomial is given and some dependence characteristics are investigated. In Section 4, new G-FGM copula is applied to three real data sets. And the last section is devoted to the conclusion.

2. Basic concepts

Let X and Y be random variables having a joint cumulative distribution function (c.d.f.) $H(x, y) = P(X \leq x, Y \leq y)$ and margins $F(x) = P(X \leq x)$, $G(y) = P(Y \leq y)$, respectively. Sklar [15] defines a copula representation of H as given by $H(x, y) = C(F(x), G(y))$, where C is a unique c.d.f. having uniform margins on unit interval. A copula must satisfy the following properties:

Definition 2.1. A bivariate copula is a function with following properties:

(1) C is 2-increasing function for all $x_1 \leq x_2$, $y_1 \leq y_2 \in [0, 1]$ such that

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0,$$

(2) C is grounded such that $C(x, 0) = C(0, y) = 0$ for all $x, y \in [0, 1]$,

(3) C has uniform margins such that $C(x, 1) = x$ and $C(1, y) = y$ for all $x, y \in [0, 1]$.

For any bivariate copula C and margins F and G, $H = C(F, G)$ is a c.d.f.. For more details about the copula, see [10].

This paper mainly focuses on the G-FGM copula families which have quite natural forms. This class is mainly characterized by their generator function ϕ that makes us possible to construct copulas from this class. Rodriguez Lallena [13] introduced the G-FGM copula class with a generator function ϕ defined on $[0, 1]^2$ by

$$C_{\theta,\phi}(u, v) = uv + \theta\phi(u)\phi(v), \quad \theta \in [-1, 1], \tag{2.1}$$

where ϕ is a function on $\mathbf{I}$. Also, Amblard and Girard [1] investigated sufficient and necessary conditions on ϕ to ensure that C_θ is a copula by following theorem:

Theorem 2.2. *ϕ generates a parametric family of copulas $C_{\theta,\phi}$, $\theta \in [-1, 1]$, if and only if it satisfies the following conditions:*

(1) *$\phi(0) = \phi(1) = 0$.*

(2) *ϕ is a 1-Lipschitz, such that $|\phi(u) - \phi(v)| \leq |u - v|$, $u, v \in \mathbf{I}$.*

Furthermore, C_θ is absolutely continuous.

The following theorem provided in [1] provides a new characterization of the generator functions ϕ constructing G-FGM copulas.

Theorem 2.3. *ϕ generates a parametric family of copulas $C_{\theta,\phi}$, $\theta \in [-1, 1]$, if and only if it satisfies the following conditions:*

(1) *ϕ is absolutely continuous.*

(2) *$|\phi'(x)| \leq 1$ almost everywhere in the unit range,*

(3) *$|\phi(x)| \leq \min(x, 1 - x)$, $x \in [0, 1]$.*

In such a case, C_θ is absolutely continuous.

In view of Definition 2.1 and Theorem 2.2, it is clearly obvious that Theorem 2.3 (1) and 2.3 (2) are both satisfied whenever $C_{\theta,\phi}$ is a copula. Assuming that ϕ is a 1-Lipschitz, we put $t_2 = 1$ and $t_1 = 0$ in the equation $|\phi(t_2) - \phi(t_1)| \leq |t_2 - t_1|$ then Theorem 2.3 (3) is satisfied. We note that, from Theorem 2.3 (3) the graph of the concave ϕ lies underneath a triangle in s 1, 3.

For dependence measures and coefficients, Amblard and Girard [1] defined the association coefficients by the following theorem:

Theorem 2.4. *Let X and Y be a random variable with copula $C_{\theta,\phi}$ given by (1). Kendall's tau (τ) and Spearman's rho (ρ) can be defined as*

$$\tau_{\theta,\phi} = 8\theta\Big(\int_0^1 \phi(x)dx\Big)^2$$

and

$$\rho_{\theta,\phi} = 12\theta\Big(\int_0^1 \phi(x)dx\Big)^2 = \frac{3}{2}\tau_{\theta,\phi}.$$

As a conclusion from the last theorem, there is a clear link between Kendalls tau and the G-FGM copula generator function. Thus, for an estimated value of Kendalls tau from observed data and for any feasible generator function ϕ, the possible G-FGM copula dependence structures can be explored.

We end this section with the introduction of a G-FGM copula based on a generator function which is the main focus of this article. We intend to create a new G-FGM copula model allowing high dependence association, and to create a multi-parameter G-FGM copula to increase modeling freedom, hence these makes it possible to work with powerful models that may provide much better goodness-of-fit results.

3. Proposed concave generator function

In many statistical applications, researchers usually work with bivariate one-parameter copula families. For applications of one-parameter copulas to survival analysis and quality control, see the books of [7] and [16]. This type of copula family has simpler forms and it is practical to make calculations on these models. However, the multi-parameter copula models will probably have a better fit for data since they are more flexible in terms of adjusting. Therefore, in this section, we explain how the generator function ϕ can be used in order to construct multi-parameter bivariate G-FGM copulas with the help of Bernstein polynomials. We construct a feasible generator function ϕ such that the properties defined in Theorem 1 are satisfied.

Let ϕ_m be a Bernstein polynomial with degree ($m > 0$) and control points α_k defined as

$$\phi_m(t) = \sum_{k=0}^{m} \alpha_k P_{k,m}(t), \;\; t \in [0,1],$$

where $P_{k,m}(t)$ are the binomial coefficients defined as $P_{k,m}(t) = \binom{m}{k} t^k(1-t)^{m-k}$. The next proposition shows that ϕ_m is a L-Lipschitz function where L is the Lipschitz constant. It will be of great help later.

Proposition 1. Let ϕ_m be a Bernstein function with order $m > 0$. Then ϕ_m is a Lipschitz with Lipschitz constant L.

Proof. We should prove that ϕ_m is L-Lipschitz, where L is Lipschitz constant. Let $t_2 \geq t_1$ be any points of $[0,1]$. We show that

$$|\phi_m(t_2) - \phi_m(t_1)| \leq L|t_2 - t_1|,$$

where L is a Lipschitz constant. Then

$$\begin{aligned}\phi_m(t_2) &= \sum_{j=0}^{m} \alpha_j \binom{m}{j} (1-t_2)^{m-j} \big(t_1 + (t_2 - t_1)\big)^j \\ &= \sum_{j=0}^{m} \alpha_j \binom{m}{j} (1-t_2)^{m-j} \Big(\sum_{k=0}^{j} \binom{j}{k} t_1^k (t_2 - t_1)^{j-k} \Big) \\ &= \sum_{j=0}^{m} \sum_{k=0}^{j} \alpha_j \frac{m! t_1^k (t_2 - t_1)^{j-k} (1-t_2)^{m-j}}{k!(j-k)!(n-j)!}.\end{aligned}$$

We can invert the order of the summation and write $k + l = j$, then

$$\phi_m(t_2) = \sum_{k=0}^{m} \sum_{l=0}^{m-k} \alpha_{k+l} \frac{m!}{k!l!(m-k-l)!} t_1^k (t_2 - t_1)^l (1-t_2)^{m-k-l}.$$

In the similar way, we can construct $\delta_{\alpha,m}(t_1)$,

$$\begin{aligned}\phi_m(t_1) &= \sum_{k=0}^{m} \alpha_k \binom{m}{k} t_1^k \big((t_2 - t_1) + (1-t_2)^{m-k}\big) \\ &= \sum_{k=0}^{m} \alpha_k \binom{m}{k} t_1^k \Big(\sum_{l=0}^{m-k} (t_2 - t_1)^l (1-t_2)^{m-k-l} \Big) \\ &= \sum_{k=0}^{m} \sum_{l=0}^{m-k} \alpha_k \frac{m!}{k!l!(m-k-l)!} t_1^k (t_2 - t_1)^l (1-t_2)^{m-k-l}.\end{aligned}$$

Then,

$$\begin{aligned}&\big|\phi_m(t_2) - \phi_m(t_1)\big| \\ &= \Big| \sum_{k=0}^{m} \sum_{l=0}^{m-k} \frac{m!}{k!l!(m-k-l)!} t_1^k (t_2 - t_1)^l (1-t_2)^{m-k-l} (\alpha_{k+l} - \alpha_k) \Big| \\ &\leq L \Big| \sum_{k=0}^{m} \sum_{l=0}^{m-k} \frac{m!}{k!l!(m-k-l)!} t_1^k (t_2 - t_1)^l (1-t_2)^{m-k-l} \Big| \\ &= L \Big| \sum_{l=0}^{m} \sum_{l=0}^{m-k} \frac{(t_2 - t_1)^l m!}{l!(m-l)!} \frac{l}{m} \Big(\sum_{k=0}^{m-l} \binom{m-l}{k} t_1^k (1-t_2)^{m-l-k} \Big) \Big| \\ &= L \Big| \sum_{l=0}^{m} \binom{m}{l} (t_2 - t_1)^l \frac{l}{m} (t_1 + 1 - t_2)^{m-l} \Big| \\ &= L \big| P_{l,m}(t_2 - t_1) \big| \\ &\leq L |t_2 - t_1|.\end{aligned}$$

It is obvious that Lipschitz constant L depends on the control points α. Also, Lipschitz constant L can be written as $L = \sup_{t \in [0,1]} \big|\phi'_m(t)\big|$. □

The next proposition shows that any ϕ_m function satisfying following requirements will provide a valid concave generator function for G-FGM copula.

Proposition 2. Let ϕ_m be a concave Bernstein polynomial with order $m > 0$. Then, ϕ_m is valid a generator function for G-FGM copula if the following constraints hold:

(1) If $\alpha_0 = \alpha_m = 0$, then $\phi_m(0) = \phi_m(1) = 0$,
(2) If $\alpha_{k+2} - 2\alpha_{k+1} + \alpha_k \leq 0,\ k = 0, \ldots, m-2$, then ϕ_m is concave,

(3) If

$$|m(\alpha_m - \alpha_{m-1})| \leq 1,$$
$$|m(\alpha_1 - \alpha_0)| \leq 1,$$

then, ϕ_m is 1-Lipschitz.

Proof. $\phi_m(0) = \sum_{k=0}^{m} \alpha_k P_{k,m}(0) = 0$ holds since $\alpha_0 = 0$. Similarly,

$$\phi_m(1) = \sum_{k=0}^{m} \alpha_k P_{k,m}(1) = 0$$

holds since $\alpha_m = 1$. Also,

$$\phi''_m(t) = m(m-1) \sum_{k=0}^{m-2} (\alpha_{k+2} - 2\alpha_{k+1} + \alpha_k) P_{k,m-2}(t) \leq 0$$

if Proposition 2 (2) satisfied. See, [5].

Also, as a result of Proposition 1, ϕ_m is Lipschitz with Lipschitz constant $L = \sup_{t\in[0,1]} |\phi'_m(t)|$. The inequality $L \leq 1$ ensures that the condition defined in Theorem 1.2 is satisfied. Because ϕ is a concave function with $\phi_m(0) = \phi_m(1) = 0$, then $\phi'_m(t) > 0, t \in [0, c)$, $\phi'_m(t) < 0, t \in (c, 1]$, $\phi'(c) = 0$ and also ϕ'_m is decreasing function. Hence it is obvious that $\phi'_m(0) = m(\alpha_1 - \alpha_0) > 0$ and $\phi'_m(1) = m(\alpha_m - \alpha_{m-1}) < 0$. Thus ϕ_m is 1-Lipschitz when both of the following conditions are satisfied:

$$|m(\alpha_m - \alpha_{m-1})| \leq 1,$$
$$|m(\alpha_1 - \alpha_0)| \leq 1.$$

□

The next proposition reveals that the proposed copula C_{θ,ϕ_m} generalizes the classical FGM copula with a generator function $\phi_{FGM}(t) = t(1-t)$.

Proposition 3. Let ϕ_m be a concave Bernstein type generator function with order $m > 0$. If the control points are determined as

$$\alpha_k = \frac{k^2}{m} - \frac{k(k-1)}{m-1},\ k = 0, \ldots, m.$$

Then, copula C_{θ,ϕ_m} is reduced to the classical FGM copula.

Proof. Primarily, we note that $P_{k,m}(t) = P(T = k)$ where T is a binomial random variable with parameters m and t. Then,

$$E(T^r) = \sum_{k=0}^{m} k^r P_{k,m}(t). \tag{3.1}$$

We know that mean and variance of the binomial variable T are mt and $mt(1-t)$, respectively. Also,

$$E(T^2) = Var(T) + E(T)^2 = mt(1-t) + m^2t^2.$$

If we determine the control points as $\alpha_k = \frac{k^2}{m} - \frac{k(k-1)}{m-1}$, $k = 0, \ldots, m$ then ϕ_m can be written as following equation:

$$\phi_m(t) = \sum_{k=0}^{m} (\frac{k^2}{m} - \frac{k(k-1)}{m-1}) P_{k,m}(t) = E(\frac{T^2}{m}) - E(\frac{T^2}{m-1}) + E(\frac{T}{m}) = t(1-t).$$

□

The following theorem defined in [8] helps us to prove the uniform convergence of the $\phi_m(t)$ to any generator function $\phi(t)$ which satisfies to properties defined in Theorem 2.2.

Theorem 3.1. *If $f(t)$ is a bounded and continuous function on the interval $[0,1]$, then as $m \to \infty$*

$$f_m^*(t) = \sum_{k=0}^{m} f(\frac{k}{m}) P_{k,m}(t) \to f(t)$$

uniformly for $t \in [0,1]$.

The α_k can be interpreted as $f(\frac{k}{m})$, $k = 0, \ldots, m$ for a suitable function f over $[0,1|$.

Proposition 4. Let ϕ_m be a Bernstein type generator function with order $m > 0$. For any valid generator function $\phi(t)$, $\phi_m(t)$ converges uniformly to $\phi(t)$ as m goes to infinity if the control points are determined as $\alpha_k = \phi(\frac{k}{m})$, $k = 0, \ldots, m$.

The proof can be easily established using the Theorem 3.1. Moreover, it can be easily shown that ϕ_m with control points $\alpha_k = \phi(\frac{k}{m})$, $k = 0, \ldots, m$ satisfies the properties of the generator function defined in Theorem 2.2. If the control points are determined as

$$\alpha_k = \phi_{FGM}(\frac{k}{m}) = \frac{k}{m}(1 - \frac{k}{m}),\ k = 0, \ldots, m.$$

Then, recalling the Equation (3.1), ϕ_m converges uniformly to ϕ_{FGM} as shown below:

$$\lim_{m\to\infty} \phi_m(t) = \lim_{m\to\infty} \sum_{k=0}^{m} \Big(\frac{k}{m}(1 - \frac{k}{m})\Big) P_{k,m}(t) = \lim_{m\to\infty} t(1-t)\frac{m-1}{m} = t(1-t).$$

The Proposition 4 reveals that this generator function based method is different from the existing methods in terms of obtaining G-FGM but also it is generalized form of the G-FGM and can be reduced FGM when m goes to infinity.

Now, we derive the most common measures of concordance between the components of a pair of random variables with copula C_{θ,ϕ_m}

Proposition 5. Let ϕ_m be a generator function for G-FGM copula. Then Kendall's tau and Spearman's rho based on ϕ_m are given by

$$\tau_{\theta,\phi_m} = 8\theta \sum_{k=0}^{m} \sum_{p=0}^{m} \alpha_k \alpha_p \binom{m}{k} \binom{m}{p} \beta(k+p+1, 2m-k-p+1)$$

and

$$\rho_{\theta,\phi_m} = \frac{3}{2} \tau_{\theta,\phi_m},$$

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.

Table 1. Maximum value of the parameters α_m and Kendall's tau τ.

Copula	θ	α_0	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	τ
C_{θ,ϕ_3}	± 1	0	$\frac{1}{3}$	$\frac{1}{3}$	0	—	—	—	—	—	—	—	± 0.2666
C_{ϕ_4}	± 1	0	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	0	—	—	—	—	—	—	± 0.3936
C_{ϕ_5}	± 1	0	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$	$\frac{1}{5}$	0	—	—	—	—	—	± 0.3936
C_{ϕ_6}	± 1	0	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	0	—	—	—	—	± 0.4586
C_{ϕ_7}	± 1	0	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{3}{7}$	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	0	—	—	—	± 0.4586
C_{ϕ_8}	± 1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{3}{8}$	$\frac{4}{8}$	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0	—	—	± 0.4981
C_{ϕ_9}	± 1	0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	0	—	± 0.4981
$C_{\phi_{10}}$	± 1	0	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$	$\frac{5}{10}$	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	0	± 0.5248

As a consequence of Proposition 4, above Kendall's tau and Spearman's rho can be used as an approximation to the true Kendall's tau and Spearman's rho of the copula defined in Equation (2.1). If $\alpha_k = 0$; $k = 0, \dots, m$, then $\tau_{\theta,\phi_m} = \rho_{\theta,\phi_m} = 0$, and bivariate independent copula is obtained. Maximum and minimum value Kendall's tau obtained by the concave Bernstein generator function for the copula C_{θ,ϕ_m} with maximum value of control points and $\theta = \pm 1$ for degree $m = 3, \dots, 10$ are summarized in Table 1. It is obvious that the range of the Kendall's tau increases when the degree of ϕ_m increases. The classical FGM copula has the limited Kendall's tau range ($\tau \in [-0.22, 0.22]$). However by the use of our new G-FGM copula, the Kendall's tau has improved greatly. A wider range for Kendall's tau (± 0.5248) were obtained for $C_{\theta,\phi_{10}}$. This enables us to modeling bivariate data sets with higher dependence structures.

Now, we pay our attention to the visual behavior of the generator function ϕ_m when it comes to the construction of copulas. In Figure 1 graphs are shown from ϕ_{10} with different control points. In Figure 1(a) graph of symmetric ϕ_{10} with control points

$$\alpha = \Big(0,\ 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5,\ 0.4,\ 0.3,\ 0.2,\ 0.1,\ 0\Big),$$

and $\theta = 1$, in Figure 1(c) graph of right skewed ϕ_{10} with control points

$$\alpha = \Big(0,\ 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.35,\ 0.27,\ 0.19,\ 0.16,\ 0.06,\ 0\Big),$$

and $\theta = 1$, in Figure 1(e) graph of left skewed ϕ_{10} with control points

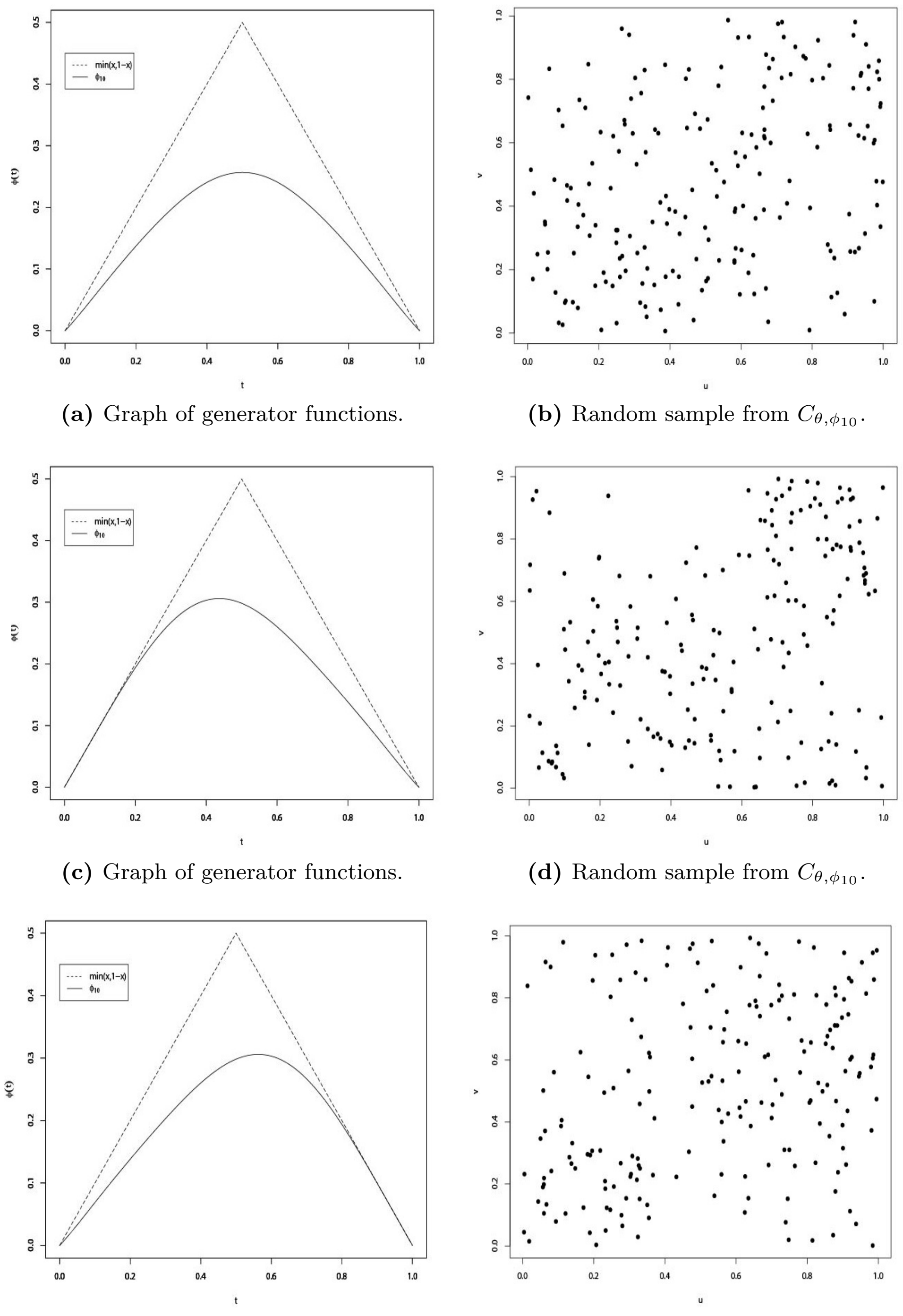
$$\alpha = \Big(0,\ 0.06,\ 0.16,\ 0.19,\ 0.27,\ 0.35,\ 0.4,\ 0.3,\ 0.2,\ 0.1,\ 0\Big)$$

and $\theta = 1$ are shown. The latter figures are represented by means of simulating observations from $(U, V) \sim C_{\theta,\phi_{10}}$. The idea here is to show how well the generator function ϕ_m behaves as a univariate representative of the affiliated copula. There are clear relationships between the visual behavior of the generator function ϕ_m and simulating observations from $(U, V) \sim C_{\theta,\phi_{10}}$.

4. Case study

In this section, we compare the new G-FGM copula with the commonly used one-parameter copulas from Archimedean (Clayton, Gumbel, Frank) and elliptical (normal, student-t) and a classical FGM, Minimum and Sinus copulas. Their formulae and coefficients range are given in Table 2. Especially, in this case study, it is aimed to investigating the goodness of fit performance of proposed G-FGM copulas under the different dependence structures. We use uranium data set available in R package "copula". For more detail about "copula" package, see [9]. According the this package "These data consist of log concentrations of 7 chemical elements in 655 water samples collected near Grand Junction, CO (from the Montrose quad-rangle of Western Colorado). Concentrations were measured for the following elements: Uranium (U), Lithium (Li), Cobalt (Co), Potassium (K), Cesium (Cs), Scandium (Sc), And Titanium (Ti)." We prefer to modeling the pairs of variables K-Ti, Co-Ti and K-Sc.

To avoid a decision about marginal distributions, the observations were transformed to pseudo-observation (normalized ranked data) by their corresponding empirical distribution functions. Figure 2(a), 2(c) and 2(e) show the scatter plots of pseudo-observation for the pairs K-Ti, Co-Ti and K-Sc, respectively. Looking at the data, strong positive dependence structure with $\tau = 0.3647$, mild positive dependence structure with $\tau = 0.0406$ and mild negative dependence structure with $\tau = -0.1368$ can be observed for the pairs of Co-Ti, K-Ti and K-Sc, respectively.



(a) Graph of generator functions. **(b)** Random sample from $C_{\theta,\phi_{10}}$.

(c) Graph of generator functions. **(d)** Random sample from $C_{\theta,\phi_{10}}$.

(e) Graph of generator functions. **(f)** Random sample from $C_{\theta,\phi_{10}}$.

Figure 1. Illustration of a G-FGM copula with ϕ_{10}.

In order to asses goodness of fit we use Cramér-von Mises distance which measure the distance between empirical copula and hypothesized parametric copula distribution functions are given by

$$CvM = \int_0^1 \int_0^1 n\Big(C_n(u,v) - C_{\theta,\phi_m}(u,v)\Big)^2 dC_n(u,v). \tag{4.1}$$

Table 2. Copulas definition and dependence range.

Copula	$C_\theta(u,v)$	$\theta \in$	$\tau \in$
Gumbel	$\exp\Big(-\big((-\log u)^{\theta+1}+(-\log v)^{\theta+1}\big)^{\frac{1}{\theta+1}}\Big)$	$[0,\infty)$	$[0,1)$
Clayton	$\Big(u^{-\theta}+v^{-\theta}-1\Big)^{-\frac{1}{\theta}}$	$(0,\infty)$	$[0,1)$
Frank	$-\frac{1}{\theta}\log\Big(1+\frac{\big(\exp(-\theta u)-1\big)\big(\exp(-\theta v)-1\big)}{\big(\exp(-\theta)-1\big)}\Big)$	$(-\infty,\infty)\backslash\{0\}$	$(-1,1)$
Normal	$\Phi_\theta\Big(\Phi^{-1}(u)+\Phi^{-1}(v)\Big)$	$[-1,1]$	$[-1,1]$
FGM	$uv+\theta(1-u)u(1-v)v$	$[-1,1]$	$[-0.22,0.22]$
SINUS	$uv+\theta\frac{1}{\pi}\sin(\pi u)\frac{1}{\pi}\sin(\pi v)$	$[-1,1]$	$[-0.33,0,33]$
MIN	$uv+\theta\min(u,1-u)\min(v,1-v)$	$[-1,1]$	$[-0.5,0.5]$

*Φ, Φ^{-1}, denote c.d.f. and the quantile function of normal distribution whereas Φ_θ denotes the c.d.f. of joint normal distribution with parameter θ.

where C_n is the empirical copula defined as

$$C_n(u,v)=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{I}(U_i\leq u, V_i\leq v).$$

Thus, the test statistic defined in Equation (4.1) allows us to compare the distances from the empirical copula among null hypothesis copulas (the smaller the better) while the p-value, simulated in 1000 Monte Carlo samples of bootstrap procedure, evaluates the null hypothesis that a copula is suitable for modeling the dependence structure in data. Also, the parameters of the C_{ϕ_m} are estimated by minimizing the Equation (4.1) under the consideration of constraints defined in Proposition 2. We note that Weiß [23] states minimum distance (MD) estimators suffer from large biases for smaller samples sizes. If this is the case, we recommend to use Maximum likelihood estimation (MLE) method to estimate the parameters of the proposed G-FGM copula for smaller sample sizes. For more details and the performance comparisons of the minimum distance and MLE methods, see [23].

Table 3. Goodness-of-Fit results for K-Ti.

Copula	$\hat\theta$	$\hat{\alpha_0}$	$\hat{\alpha_1}$	$\hat{\alpha_2}$	$\hat{\alpha_3}$	$\hat{\alpha_4}$	$\hat{\alpha_5}$	$\hat{\alpha_6}$	$\hat{\alpha_7}$	$\hat{\alpha_8}$	$\hat{\alpha_9}$	$\hat{\alpha_{10}}$	$\hat\tau$	CvM	P-Value
C_{FGM}	0.1827	—	—	—	—	—	—	—	—	—	—	—	0.0406	0.0555	0.0428
C_{SIN}	0.1506	—	—	—	—	—	—	—	—	—	—	—	0.0494	0.0523	0.0405
C_{MIN}	0.0937	—	—	—	—	—	—	—	—	—	—	—	0.0468	0.0500	0.0491
C_{Frank}	0.3659	—	—	—	—	—	—	—	—	—	—	—	0.0405	0.0553	0.0414
C_{Gumbel}	0.0423	—	—	—	—	—	—	—	—	—	—	—	0.0405	0.0671	0.0014
$C_{Clayton}$	0.0846	—	—	—	—	—	—	—	—	—	—	—	0.0405	0.0458	0.0574
C_{Normal}	0.0637	—	—	—	—	—	—	—	—	—	—	—	0.0405	0.0573	0.0414
C_{θ,ϕ_3}	0.4474	0	0.3333	0.1666	0	—	—	—	—	—	—	—	0.0681	0.0427	0.0625
C_{θ,ϕ_4}	0.7240	0	0.25	0.1666	0.0833	0	—	—	—	—	—	—	0.0715	0.0374	0.1052
C_{θ,ϕ_5}	1	0	0.2	0.1500	0.1000	0.0500	0	—	—	—	—	—	0.0692	0.0353	0.1123
C_{θ,ϕ_6}	1	0	0.1666	0.1689	0.1266	0.0844	0.0422	0	—	—	—	—	0.0705	0.0352	0.1217
$C_{\theta,\phi_{10}}$	0.9393	0	0.1	0.1895	0.1675	0.1437	0.1198	0.0958	0.0719	0.0479	0.0239	0	0.0729	0.0341	0.1296

Table 3 shows the goodness of fit results and estimated parameters for the pair K-Ti. According the this table, $C_{\phi_{10}}$ is the best performing copula model since it possesses the greatest p-value (0.1296) and lowest CvM (0.0341) values. Also, the p-values and CvM test statistics decease when the degree of generator function ϕ_m increases for the copula C_{ϕ_m}. Similarly, from Tables 4 and 5, the best fit among all copulas for the pairs of Co-Ti and K-Sc are Clayton (P-val:0.1837) and $C_{\phi_{10}}$ (p-val:0.0903), respectively. Although the best fit for the pair of Co-Ti is Clayton copula according to GoF results represented in Table 4, G-FGM copula C_{ϕ_m} for $m > 4$ has the acceptable fits according to p-values (p>0.05).

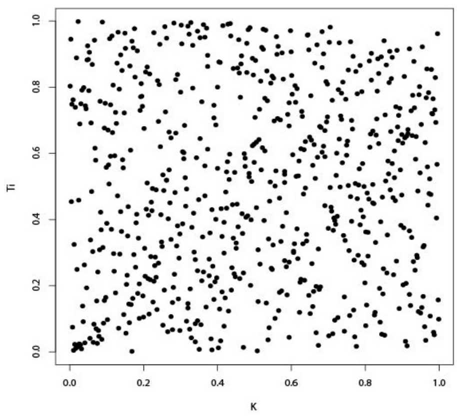
Table 4. Goodness-of-Fit results for Co-Ti.

Copula	$\hat{\theta}$	$\hat{\alpha_0}$	$\hat{\alpha_1}$	$\hat{\alpha_2}$	$\hat{\alpha_3}$	$\hat{\alpha_4}$	$\hat{\alpha_5}$	$\hat{\alpha_6}$	$\hat{\alpha_7}$	$\hat{\alpha_8}$	$\hat{\alpha_9}$	$\hat{\alpha_{10}}$	$\hat{\tau}$	CvM	P-Value
C_{FGM}	1	—	—	—	—	—	—	—	—	—	—	—	0.2222	0.3659	0.0005
C_{SIN}	1	—	—	—	—	—	—	—	—	—	—	—	0.3285	0.0416	0.0568
C_{MIN}	0.5893	—	—	—	—	—	—	—	—	—	—	—	0.2946	0.1238	0.0001
C_{Frank}	3.6941	—	—	—	—	—	—	—	—	—	—	—	0.3647	0.0222	0.1873
C_{Gumbel}	0.5742	—	—	—	—	—	—	—	—	—	—	—	0.3647	0.0757	0.0005
$C_{Clayton}$	1.1484	—	—	—	—	—	—	—	—	—	—	—	0.3647	0.1411	0.0005
C_{Normal}	0.5421	—	—	—	—	—	—	—	—	—	—	—	0.3647	0.0351	0.0664
C_{θ,ϕ_3}	1	0	0.3333	0.3333	0	—	—	—	—	—	—	—	0.2666	0.2653	0.0000
C_{θ,ϕ_4}	1	0	0.25	0.4999	0.25	0	—	—	—	—	—	—	0.3936	0.0426	0.0514
C_{θ,ϕ_5}	1	0	0.2	0.3999	0.3999	0.2	0	—	—	—	—	—	0.3936	0.0426	0.0563
C_{θ,ϕ_6}	1	0	0.1666	0.3333	0.4151	0.3333	0.1666	0	—	—	—	—	0.4031	0.0415	0.0591
$C_{\theta,\phi_{10}}$	1	0	0.1	0.1999	0.2999	0.3217	0.3429	0.3640	0.2998	0.1999	0.0999	0	0.4036	0.0391	0.0638

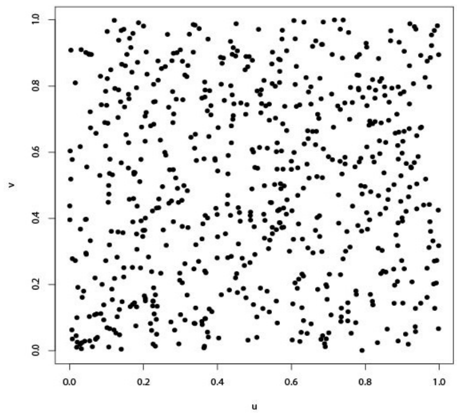
Table 5. Goodness-of-Fit results for K-Sc.

Copula	$\hat{\theta}$	$\hat{\alpha_0}$	$\hat{\alpha_1}$	$\hat{\alpha_2}$	$\hat{\alpha_3}$	$\hat{\alpha_4}$	$\hat{\alpha_5}$	$\hat{\alpha_6}$	$\hat{\alpha_7}$	$\hat{\alpha_8}$	$\hat{\alpha_9}$	$\hat{\alpha_{10}}$	$\hat{\tau}$	CvM	P-Value
C_{FGM}	-0.6157	—	—	—	—	—	—	—	—	—	—	—	-0.1368	0.1362	0.0001
C_{SIN}	-0.4414	—	—	—	—	—	—	—	—	—	—	—	-0.1450	0.1325	0.0002
C_{MIN}	-0.2623	—	—	—	—	—	—	—	—	—	—	—	-0.1311	0.1515	0.0004
C_{Frank}	-1.2506	—	—	—	—	—	—	—	—	—	—	—	-0.1368	0.1346	0.0004
C_{Gumbel}	0	—	—	—	—	—	—	—	—	—	—	—	0	0.3930	0.0000
$C_{Clayton}$	-0.2407	—	—	—	—	—	—	—	—	—	—	—	-0.1368	0.2109	0.0000
C_{Normal}	-0.2133	—	—	—	—	—	—	—	—	—	—	—	-0.1368	0.1365	0.0004
C_{θ,ϕ_3}	-1	0	0.1996	0.3333	0	—	—	—	—	—	—	—	-0.1719	0.0967	0.0278
C_{θ,ϕ_4}	-1	0	0.1412	0.2825	0.25	0	—	—	—	—	—	—	-0.1769	0.0845	0.0453
C_{θ,ϕ_5}	-1	0	0.1023	0.2046	0.3070	0.2	0	—	—	—	—	—	-0.181	0.0786	0.0543
C_{θ,ϕ_6}	-1	0	0.0785	0.1571	0.2357	0.3143	0.1666	0	—	—	—	—	-0.1849	0.0606	0.0795
$C_{\theta,\phi_{10}}$	-0.9929	0	0.0456	0.0892	0.1313	0.1726	0.2135	0.2542	0.2942	0.1978	0.1	0	-0.1863	0.0503	0.0903

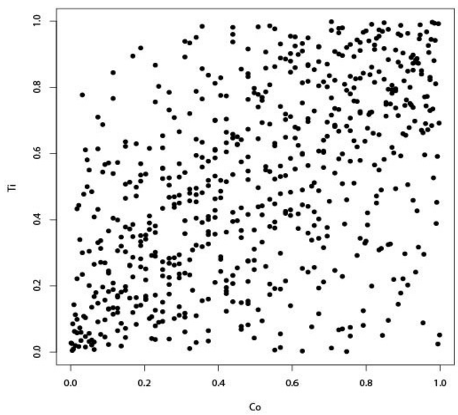
In order to assess graphical goodness of fit to pairs of Co-Ti, K-Ti and K-Sc for G-FGM copula, we generate random samples of size 655 using the estimated dependence parameters in Tables 3–5. Figures 2(b)-2(d) and 2(f) display the simulated random sample from the G-FGM copula functions with $m = 10$. these figures reveal that the G-FGM copulas provide an acceptable fit to the actual pairs of Co-Ti, K-Ti and K-Sc. In addition to that, Figure 3 represents to parametric estimation of the generator functions for the C_{ϕ_m}. In each graph, generator functions are visualized for the degrees $m = 3$, 10. According the Figure 3, left-skewed, symmetric and right-skewed generator functions posses to pairs K-Ti, Co-Ti and K-Sc, respectively. We may conclude from this figure that if the data has mild negative dependence association, the graph of the generator function of C_{ϕ_m} is a left-skewed. On the contrary, if the data has mild positive dependence association, the graph of the generator function of C_{ϕ_m} is a right-skewed. For the strong positive dependence structure as could be observed in the pair Co-Ti, the generator function for the C_{ϕ_m} has the symmetric.



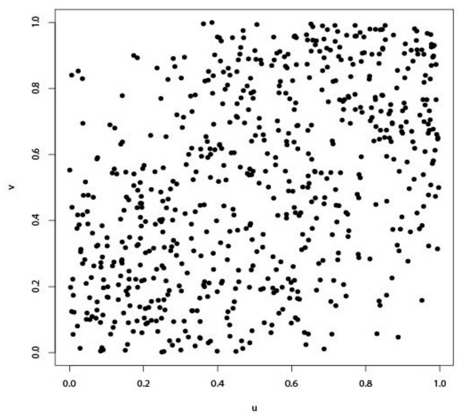
(a) Scatterplot for K-Ti.



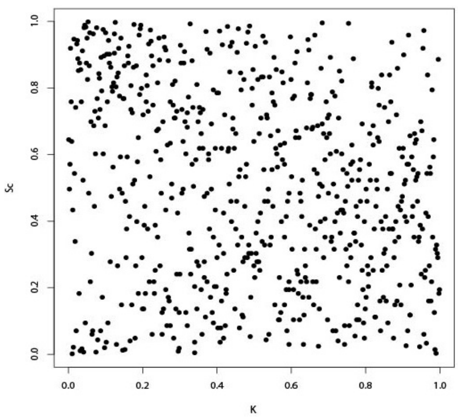
(b) $C_{\phi_{10}}$ with paramaters estimated from K-Ti.



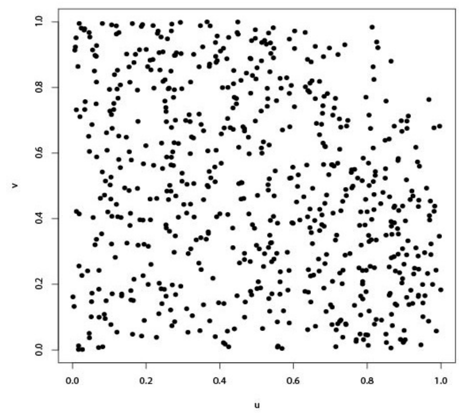
(c) Scatterplot for Co-Ti.



(d) $C_{\phi_{10}}$ with paramaters estimated from Co-Ti.

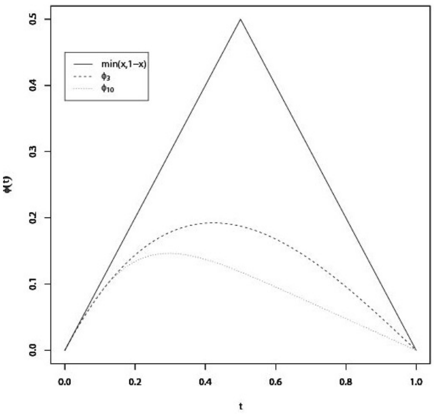


(e) Scatterplot for K-Sc.

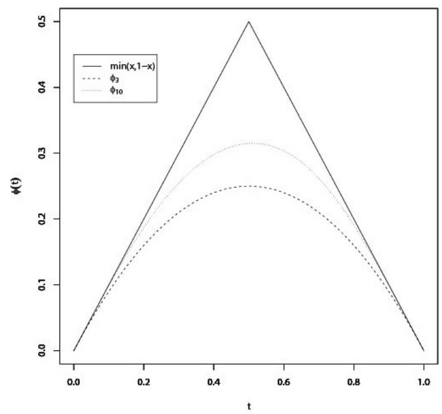


(f) $C_{\phi_{10}}$ with paramaters estimated from K-Sc.

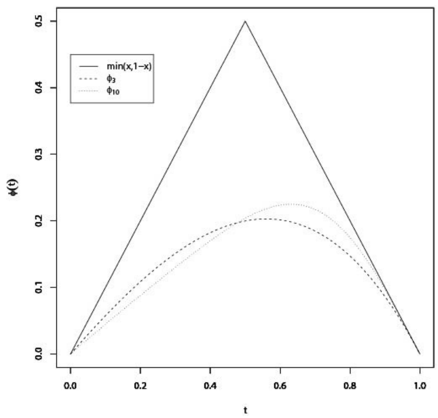
Figure 2. Scatter plots of Pseudo-observations for real datasets.



(a) Generator functions of C_{ϕ_m} for K-Ti.



(b) Generator functions of C_{ϕ_m} for Co-Ti.



(c) Generator functions of C_{ϕ_m} for K-Sc.

Figure 3. Graph of the generator functions for real datasets.

5. Conclusion

We have introduced the Bernstein polynomial type generator function of a new multi-parametric G-FGM family of copulas, describing its Kendall's tau with closed-form. The method is illustrated using $m = 3, 4, 5, 6, 10$ parameters of G-FGM copula models. It is also shown that, when constructing generator function with Bernstein polynomials, a multi-parameter G-FGM copula family can be constructed in an organized way. The proposed G-FGM copula makes it possible to work with powerful models that can provide a much better goodness-of-fit results according to its flexibility.

In the case study, researchers may use the MLE method to estimate the parameters of G-FGM copula. We know that MLE is an efficient estimator, also even for the small samples. Furthermore, the MLE can also facilitates the Akaike information criterion for model selection accounting for the number of parameters. In this study, we use MD estimator(s) since we study on the large samples, and these estimators have also asymptotically minimum variance, see [23].

Notice that for identically distributed random variables, exchangeability is equivalent to the symmetry of the copula: $C(u, v) = C(v, u)$, $u, v \in [0, 1]$ [6, 11]. When this is not satisfied, then copula C is said to be non-exchangeable [4]. In this framework, this paper is mainly developed around the G-FGM copula with exchangeable random variables.

But, researches might be interested to modeling the dependence structure of the non-exchangeable random variables using the asymmetric G-FGM copula given by

$$C(u,v) = uv + \theta a(u)b(v),\ \theta \in [-1,1]$$

where a and b are functions defined on the unit interval. [22] states that necessary and sufficient conditions for $C(u,v)$ to be a valid copula are $a(0) = a(1) = b(1) = b(0) = 0$ and

$$|a(x_2) - a(x_1)|\,|b(y_2) - b(y_1)| \leq |x_2 - x_1|\,|y_2 - y_1|,$$

for all x_2, x_1, y_2, y_1 on unit interval. Our further work will consist in constructing the asymmetric G-FGM copula with the generator function constructed from the Bernstein polynomials.

Acknowledgment. The author would like to thank the referee for his helpful comments and suggestions which improved the presentation of the paper.

References

[1] C. Amblard and S. Girard, *Symmetry and dependence properties within a semi-parametric family of bivariate copulas*, J. Nonparametr. Stat. **14** (6), 715-727, 2002.

[2] I. Bairamov and K. Bairamov, *From the Huang-Kotz FGM distribution to Bakers bivariate distribution*, J. Multivariate Anal. **113**, 106-115, 2013.

[3] J. Carnicero, M. Wiper and M. Ausin, *Density estimation of circular data with Bernstein polynomials*, Hacet. J. Math. Stat. **47** (2), 273-286, 2018.

[4] R. Cerqueti, and L. Claudio, *Non-exchangeable copulas and multivariate total positivity*, Inform. Sci. **360**, 163-169, 2016.

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

[6] F. Durante, E.P. Klement, C. Sempi and M. Ubeda-Flores, *Measures of non-exchangeability for bivariate random vectors*, Statist. Papers, **51** (3), 687-699, 2010.

[7] T. Emura, S. Matsui and V. Rondeau, *Survival Analysis with Correlated Endpoints, Joint Frailty-Copula Models*, JSS Research Series in Statistics, Springer, 2019.

[8] W. Feller, *An Introduction to Probability Theory and its Applications*, Wiley, 1965.

[9] Y. Jun, *Enjoy the joy of copulas: with a package copula*, J. Stat. Softw. **21** (4), 1-21, 2007.

[10] R.B. Nelsen, *An Introduction to Copulas*, Springer, 2007.

[11] R.B. Nelsen, *Extremes of nonexchangeability*, Statist. Papers, **48** (2), 329-336, 2007.

[12] D. Pfeifer and O. Ragulina, *Adaptive Bernstein copulas and risk management*, Mathematics, **8** (12), 1-22, 2020.

[13] J.A. Rodriguez Lallena, *Estudio de la compatibilidad y diseño de nuevas familias en la teoría de cópulas aplicaciones*, PhD thesis, Universidad de Granada, 1992.

[14] S. Saekaow and S. Tasena, *Sobolev convergence of empirical Bernstein copulas*, Hacet. J. Math. Stat. **48** (6), 1845-1858, 2019.

[15] A. Sklar, *Fonctions de repartition a n dimensions et leurs marges*, Publ. Inst. Stat. Univ. Paris **8**, 229-231, 1959.

[16] L.H. Sun, X.W. Huang, M.S. Alqawba, J.M. Kim and T. Emura, *Copula-based Markov Models for Time Series-Parametric Inference and Process Control*, JSS Research Series in Statistics, Springer, 2020.

[17] S.O. Susam, *Parameter estimation of some Archimedean copulas based on minimum Cramér-von-Mises distance*, J. Iran. Stat. Soc. (JIRSS) **19** (1), 163-183, 2020.

[18] S.O. Susam and M.S. Erdogan, *Plug-in estimation of dependence characteristics of Archimedean copula via Bézier curve*, REVSTAT, 1-17, In Press.

[19] S.O. Susam and B.H. Ucer, *Testing independence for Archimedean copula based on Bernstein estimate of Kendall distribution function*, J. Stat. Comput. Simul. **88** (13), 2589-2599, 2018.

[20] S.O. Susam and B.H. Ucer, *A goodness-of-fit test based on Bézier curve estimation of Kendall distribution*, J. Stat. Comput. Simul. **90** (7), 1194-1215, 2020.

[21] S.O. Susam and B.H. Ucer, *On construction of Bernstein-Bézier type bivariate Archimedean copula*, REVSTAT, 1-17, In Press.

[22] M. Úbeda Flores, *Introducción a la teoria de cópulas. Aplicaciones*, Predoctoral Research Dissertation, Universidad de Almeria, 1998.

[23] G. Weiß, *Parameter estimation by maximum-likelihood and minimum-distance estimators: a simulation study*, Comput. Statist. **26** (1), 31-54, 2011.