Information Sciences Letters

Volume 11
Issue 4 *Jul. 2022*

Article 9

2022

New Bivariate MS Copula via Ru¨schendorf method

M. M. Seyam
Faculty of Science and Arts, Jouf University, Sakakah, Saudi Arabia\\ Faculty of Science, Tanta University, Tanta, Egypt, mmseyam@ju.edu.sa

S. M. Elsobky
Faculty of Science, Tanta University, Tanta, Egypt, mmseyam@ju.edu.sa

Inf. Sci. Lett. **11**, No. 4, 1087-1092 (2022) 1087

Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/110409

New Bivariate MS Copula via Rüschendorf method

M. M. Seyam[1,2,*] *and S. M. Elsobky*[2]

[1]Faculty of Science and Arts, Jouf University, Sakakah, Saudi Arabia
[2]Faculty of Science, Tanta University, Tanta, Egypt

Received: 26 Dec. 2021, Revised: 8 April 2022, Accepted: 29 Apr. 2022
Published online: 1 Jul. 2022

Abstract: Copula study is growing field which used for constructing families of bivariate and multivariate distributions also a measure of dependence structure. There are several methods for constructing copulas, one of these methods is a Rüschendorf method. A new bivariate copula called MS copula is introduced based on Rüschendorf method. Several properties concerning dependence concepts and concordance ordering are studied and also fitting a copula using maximum likelihood. Finally, an application is presented to show applicability of the proposed MS copula.

Keywords: Copula; Concordance ordering; Dependence concepts; Rüschendorf method.

1 Introduction

Copulas begin with [1], but in 1959, Sklar introduce the concept and the name of copula, and proving the theorem that now bears his name. Sklar [2] proposed copula as the functions that join a multivariate distribution function to its one-dimensional marginal distribution function. A variety families of copulas can be found in [3] as Product, Fréchet's, Farlie-Gumbel-Morgenstern's, Galambos, Marshall-Olkin's, Cuadras-Augé and Archimedean copula. There are several studies and research presented bivariate copula with different distributions as [4] proposed bivariate inverse Weibull distribution and its application in complementary risks model. [5] introduced a bayesian-model-averaging copula method for bivariate hydrologic correlation analysis. A new class of bivariate Farlie-Gumbel-Morgenstern copula presented by [6]. A new family of bivariate copulas constructed using a unit Weibull distortion by [7]. Another new family of bivariate copula proposed by [8]. The relationships between several families of copula presented by [9]. Copulas have performed in many important fields as biomedical studies, economics, civil engineering and social science which approves its importance, see [10]. A bivariate copula is a function $C : I^2 \to I(=[0,1])$ which satisfies the following properties:

a) Boundary conditions:

$$C(u,0) = C(0,v) = 0, C(u,1) = u, C(1,v) = v, C(1,1) = 1 \quad u,v \in [0,1] \tag{1}$$

b) 2-increasing property:

For every u_1,u_2,v_1,v_2 in I such that $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0 \tag{2}$$

If $C(u,v)$ is twice differentiable, then the 2-increasing property is equivalent to

$$\frac{\partial^2 C(u,v)}{\partial u \partial v} \geq 0$$

For bivariate copula, a random variables X and Y with joint distribution function $H(x,y)$ and marginal distribution functions $F(x)$ and $G(y)$, respectively, then there exists a copula C (which is uniquely determined on Range$F \times$ Range G) such that:

$$H(x,y) = C(F(x),G(y)), \forall (x,y) \in R \tag{3}$$

One of the methods for constructing copulas is a Rüchendorf method introduced by [11] using uniform marginal, which depends on choosing an arbitrary bivariate function that is integrable on a unit square $[0,1]$. Suppose $f^1(x,y)$ has integral zero on the unit square and its two marginals integrate to zero, i.e.

* Corresponding author e-mail: mmseyam@ju.edu.sa

$$\int_0^1\int_0^1 f^1(x,y)dxdy=0, \int_0^1 f^1(x,y)dy=0$$
$$\text{and } \int_0^1 f^1(x,y)dx=0 \quad (4)$$

The function $f^1(x,y)$can be found as follows:

1- Choose a function $f(x,y)$ which is an arbitrary real integrable function on the unit square [0,1] and then compute the following :

$$f_1(x)=\int_0^1 f(x,y)dy, f_2(y)=\int_0^1 f(x,y)dx$$
$$\text{and A=}\int_0^1\int_0^1 f(x,y)dxdy$$

2- Then, $f^1(x,y)=f(x,y)-f_1(x)-f_2(y)+A$

Thus$1+f^1(x,y)$ is the density of a copula which is non-negative.

2 New bivariate MS copula

A new bivariate copula called MS copula will be demonstrated using Rüchendorf method by choosing an arbitrary bivariate function $f(x,y)$ that satisfies the steps given in Section 1. The following steps proposed the construction of the copula.

Firstly: We choose the function

$$f(x,y)=\sqrt{xy}, \ \ where \ \ x,y\in[0,1] \quad (5)$$

Then calculate$f_1(x), f_2(y)$ and A as follows:

$$f_1(x)=\int_0^1\sqrt{xy}dy=\frac{2}{3}\sqrt{x}$$
$$f_2(y)=\int_0^1\sqrt{xy}dx=\frac{2}{3}\sqrt{y}$$

and

$$A=\int_0^1\int_0^1\sqrt{xy}dxdy=\frac{4}{9}$$

Secondly: Compute$f^1(x,y)$ that is

$$f^1(x,y)=f(x,y)-f_1(x)-f_2(y)+A$$

In our case, that is $f(x,y)=\sqrt{xy}$, then

$$f^1(x,y)=\sqrt{xy}-\frac{2}{3}\sqrt{x}-\frac{2}{3}\sqrt{y}+\frac{4}{9}$$

Check that equations 4 satisfies:

$$\int_0^1 f^1(x,y)dx=\int_0^1\left(\sqrt{xy}-\frac{2}{3}\sqrt{x}-\frac{2}{3}\sqrt{y}+\frac{4}{9}\right)dx=0$$
$$\int_0^1 f^1(x,y)dy=\int_0^1\left(\sqrt{xy}-\frac{2}{3}\sqrt{x}-\frac{2}{3}\sqrt{y}+\frac{4}{9}\right)dy=0$$

and

$$\int_0^1\int_0^1 f^1(x,y)dxdy=\int_0^1\int_0^1\left(\sqrt{xy}-\frac{2}{3}\sqrt{x}-\frac{2}{3}\sqrt{y}+\frac{4}{9}\right)dxdy=0$$

Let $g(x,y)=1+\theta\left[\sqrt{xy}-\frac{2}{3}\sqrt{x}-\frac{2}{3}\sqrt{y}+\frac{4}{9}\right], \theta\in[-1,1]$, is the density of a copula, so it's necessary to show that $g(x,y)>0$.

Find and classify the critical points of $g(x,y)$.

Step (1): Find $g_x(x,y)$ and $g_y(x,y)$

$$g_x(x,y)=\frac{\partial g(x,y)}{\partial x}=\theta\left[\frac{y}{2\sqrt{xy}}-\frac{1}{3\sqrt{x}}\right] \quad (6)$$
$$g_y(x,y)=\frac{\partial g(x,y)}{\partial y}=\theta\left[\frac{x}{2\sqrt{xy}}-\frac{1}{3\sqrt{y}}\right] \quad (7)$$

Step (2): Find $g_{xx}(x,y)$ and $g_{yy}(x,y)$

$$g_{xx}(x,y)=\frac{\partial^2 g(x,y)}{\partial x\partial x}=\theta\left[\frac{-y^2}{4xy\sqrt{xy}}+\frac{1}{6x\sqrt{x}}\right] \quad (8)$$
$$g_{yy}(x,y)=\frac{\partial^2 g(x,y)}{\partial y\partial y}=\theta\left[\frac{-x^2}{4xy\sqrt{xy}}+\frac{1}{6y\sqrt{y}}\right] \quad (9)$$

Step (3): Equations 6 and 7 will equal to zero to find the critical points, this gives us a simultaneous equation to solve for x and y, and they are found to be $(x,y)=(4/9,4/9)$.

Step (4): Find the discrimination of $g(x,y)$at critical point $(4/9,4/9)$

where $D(x,y)=g_{xx}(x,y)g_{yy}(x,y)-[g_{xy}(x,y)]^2$,then we get $D(x,y)=-\frac{9}{16}\theta$

The MS copula will be constructed by obtain the double integral of $g(x,y)$, namely,

$$C(u,v)=\int_0^v\int_0^u g(x,y)dxdy$$
$$=\int_0^v\int_0^u\left(1+\theta\left[\sqrt{xy}-\frac{2}{3}\sqrt{x}-\frac{2}{3}\sqrt{y}+\frac{4}{9}\right]\right)dudv$$
$$C(u,v)=uv+\frac{4}{9}uv\theta-\frac{4}{9}u^{3/2}v\theta-\frac{4}{9}uv^{3/2}\theta+\frac{4}{9}u^{3/2}v^{3/2}\theta \quad (10)$$

For $\theta\in[-1,1]$ and $u,v\in[0,1]$. It is important to ensure that the MS copula in 10 have met properties of the bivariate copulas boundary conditions and 2-increasing property in Section 1. MS copula check boundary conditions which follows immediately by

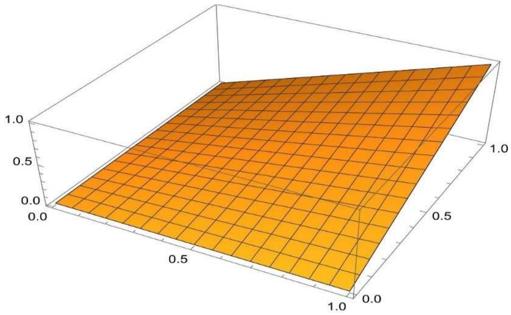

Fig. 1: MS copula density

substituting in (1). To confirm that the MS copula satisfies the 2-increasing property we will show that

$$\frac{\partial^2 C(u,v)}{\partial u \partial v} \geq 0$$

$$\frac{\partial^2 C(u,v)}{\partial u \partial v} = 1 + \frac{4}{9}\theta - \frac{12}{18}u^{1/2}\theta - \frac{12}{18}v^{1/2}\theta + u^{1/2}v^{1/2}\theta \tag{11}$$

Since $\frac{\partial^2 C(u,v)}{\partial u \partial v}$ given by11 is the density of a copula, clearly it is non-negative, so, the MS copula satisfies the 2-increasing properties.

3 MS copula properties

There are several measures of association can be study, in this section, the Kendall's tau, Spearman's rho, Blomqvist, Gini gamma, Spearman footrule coefficient and tail dependence properties of the MS copula are presented.

3.1 Kendall's tau

The Kendall's tau can be expressed as:

$$\tau = 4\int_0^1\int_0^1 C(u,v)\,dC(u,v) - 1 \tag{12}$$

or equivalently,

$$\tau = 1 - 4\int_0^1\int_0^1 \frac{\partial C}{\partial u}\frac{\partial C}{\partial v}dudv \tag{13}$$

by substituting

$$\tau = 1 - 4\int_0^1\int_0^1 \left(v + \frac{4}{9}v\theta - \frac{12}{18}u^{1/2}v\theta - \frac{4}{9}v^{3/2}\theta + \frac{12}{18}u^{1/2}v^{3/2}\theta\right)\left(u + \frac{4}{9}u\theta - \frac{4}{9}u^{3/2}\theta - \frac{12}{18}uv^{1/2}\theta + \frac{12}{18}u^{3/2}v^{1/2}\theta\right)dudv \tag{14}$$

after some computation the Kendall's τ is given by

$$\tau = \frac{8}{255}\theta \tag{15}$$

where $\theta \in [-1,1]$is the copula parameter.

3.2 Spearman's rho

Spearman's rho is the population form of the measure of association and it is based on concordance and discordance. Spearman's rho can be expressed as

$$\rho = 12\int_0^1\int_0^1 uv\,dC(u,v) - 3 \tag{16}$$

or equivalently,

$$\rho = 12\int_0^1\int_0^1 C(u,v)\,dudv - 3 \tag{17}$$

by substituting

$$\rho = 12\int_0^1\int_0^1 \left(uv + \frac{4}{9}uv\theta - \frac{4}{9}u^{3/2}v\theta - \frac{4}{9}uv^{3/2}\theta + \frac{4}{9}u^{3/2}v^{3/2}\theta\right)dudv - 3$$

after some computation ρ is given by:

$$\rho = \frac{4}{75}\theta,\ \theta \in [-1,1] \tag{18}$$

3.3 Blomqvist medial correlation coefficient

Blomqvist also called the medial correlation such a measure using population medians, it can be expressed as;

$$\beta = 4C\left(\frac{1}{2},\frac{1}{2}\right) - 1 \tag{19}$$

after some calculations we get

$$\beta = \frac{4}{50}\theta, \theta \in [-1,1] \tag{20}$$

3.4 Gini gamma

This measure rest on the distance between the lower bound M and upper bound W of copula (Fréchet-Hoeffding bounds). Gini gamma can be given as:

$$\gamma_C = 4\left(\int_0^1 C(u,1-u)\,du - \int_0^1 (u - C(u,u))\,du\right) \tag{21}$$

by substituting

$$\gamma_C = 4(\int_0^1 u(1-u) + \frac{4}{9}u(1-u)\theta - \frac{4}{9}u^{3/2}(1-u)\theta - \frac{4}{9}u(1-u)^{3/2}\theta + \frac{4}{9}u^{3/2}(1-u)^{3/2}\theta)du$$

after some calculations we get

$$\gamma_C = -\frac{4}{45}\theta \tag{22}$$

3.5 Spearman footrule coefficient

Another measure of association between two variates is Spearman's foot-rule which expressed by:

$$\varphi_C = 6\int_0^1 C(u,u)\,du - 2 \tag{23}$$

by substituting

$$\varphi_C = 6\int_0^1 \left(u^2 + \frac{4}{9}u^2\theta - \frac{8}{9}u^{5/2}\theta + \frac{4}{9}u^3\theta\right)du - 2$$

after some computations

$$\varphi_C = \frac{2}{63}\theta$$

3.6 Tail Dependence

Tail dependence measures which is the dependence between the variables in the upper-right quadrant and in the lower left quadrant of unit square I^2. If a bivariate copula $C(u,v)$ is such that:

$$\lambda_U = \lim_{u\to 1}\frac{1-2u+C(u,u)}{1-u} \tag{24}$$

exist, then $C(u,u)$ has an upper tail dependence if $\lambda_U \in (0,1]$ and no upper tail dependence if $\lambda_U = 0$.
By substituting in 24 we get

$$\lambda_U = \lim_{u\to 1}\frac{1-2u+u^2+\frac{4}{9}u^2\theta-\frac{8}{9}u^{5/2}\theta+\frac{4}{9}u^3\theta}{1-u}$$

after some calculations we find that $\lambda_U = 0$ then no upper tail dependence. Similarly, if

$$\lambda_L = \lim_{u\to 0}\frac{C(u,u)}{u} \tag{25}$$

exist, then $C(u,v)$ has a lower tail dependence if $\lambda \in (0,1]$ and no lower tail dependence if $\lambda_L = 0$.
By substituting in (25) we get

$$\lambda_L = \lim_{u\to 0}\frac{u^2+\frac{4}{9}u^2\theta-\frac{8}{9}u^{5/2}\theta+\frac{4}{9}u^3\theta}{u}$$

after some calculations we find that $\lambda_L = 0$ then no lower tail dependence.

4 Fitting a Copula Using Maximum Likelihood

Parameters of a statistical model are often estimated using maximum likelihood techniques, this means that copula parameters and parameters of the marginal distributions are estimated by maximizing. Let the following partial derivatives given as

$$C_1 = \frac{\partial g(x,y)}{\partial x}, C_2 = \frac{\partial g(x,y)}{\partial y} \text{ and } C_{12} = \frac{\partial^2 g(x,y)}{\partial x \partial y}$$

Consisting of likelihood function is

$$C_{12} = \int_0^1 g(x,y)dy \int_0^1 g(x,y)dx \frac{\partial^2 g(x,y)}{\partial x \partial y} \tag{26}$$

By substituting

$$g(x) = \int_0^1 g(x,y)dy = \int_0^1 \left(1+\theta\left(\sqrt{xy}-\frac{2}{3}\sqrt{x}-\frac{2}{3}\sqrt{y}+\frac{4}{9}\right)\right)dy$$

$$g(x) = 1$$

$$g(y) = \int_0^1 g(x,y)dx = \int_0^1 \left(1+\theta\left(\sqrt{xy}-\frac{2}{3}\sqrt{x}-\frac{2}{3}\sqrt{y}+\frac{4}{9}\right)\right)dx$$

$$g(y) = 1$$

then we get

$$C_{12} = \frac{\partial^2 g(x,y)}{\partial x \partial y} = \frac{76}{16\sqrt{xy}} \tag{27}$$

$$g(y) - C_2 = g(y)(1-C_2) \tag{28}$$

by substituting

$$g(y) - C_2 = 1 - \theta\left(\frac{x}{2\sqrt{xy}} - \frac{1}{3\sqrt{y}}\right) \tag{29}$$

$$\log L(x,y,\theta) = (1-\theta)\log\frac{\partial^2 g(x,y)}{\partial x\partial y} + \theta(\log g(x) + \log(1-C_2)) \tag{30}$$

By substituting we get

$$\log L(x,y,\theta) = (1-\theta)\log\left(\frac{76}{16\sqrt{xy}}\right) + \theta\left(\log 1 + \log\left(1-\theta\left(\frac{x}{2\sqrt{xy}}-\frac{1}{3\sqrt{y}}\right)\right)\right) \tag{31}$$

Parameters estimate are determined by

$$\sum_{i=1}^{n}\log L(x,y,\theta) \tag{32}$$

5 Application

To examine the proposed copula, a real data of 332 women of Pima Indian heritage. They were tested for diabetes according to WHO criteria. The data from [1] were collected by the US National Institute of Diabetes and Digestive and Kidney Diseases. Two variables (Body Mass Index "BMI" and Diabetes Pedigree Function "Ped") were utilized in this application according to the dependence between them. The dependence of the two variables can be tested using Kendall's ρ and Spearman's τ as follows:

$$H_0^\tau : \tau = 0 \ \& \ H_0^\rho : \rho = 0$$

Now, τ= 0.064237 with p-value = 0.0815 and ρ = 0.096973 with p-value = 0.0777, which prove the variables dependency.

The BMI and Ped can be fitted marginally using Gamma (20.831, 1.5957) and Gamma (2.1156, 0.24976) distributions respectively as shown in fig. 2.

The data was fitted using the proposed copula against [12,13,14] copulas. The results in table (1) show that the proposed copula fits the data better than the other copulas.

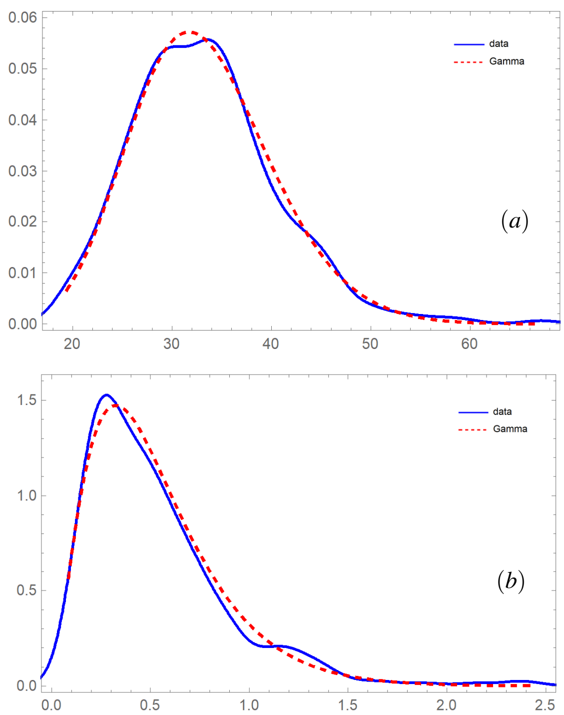

Fig. 2: (a) is the BMI and (b) is the Ped vs Estimated Gamma distribution.

Table 1: Information criterion for different copulas.

Copula	-Log	AIC	BIC	AICC	HQIC	τ/ρ
MS	1160.7659	2331.532	2350.557	2331.716	2339.119	1.5
Frank	1166.4055	2342.811	2361.837	2342.995	2350.398	1.498
Clayton	1166.3415	2342.683	2361.709	2342.867	2350.270	1.497
Cubic	1187.2714	2384.543	2403.569	2384.727	2392.130	1.001
Sine	1166.0625	2342.125	2361.151	2342.309	2349.712	1.5

6 Conclusion

In this paper, a new bivariate MS copula using an important method called Rüschendorf method is introduced. Several measures of association and dependence structure for this copula is discussed. The importance of MS copula showed in the application comparison with other copulas.

Conflict of Interest

All authors declare that there is no conflict regarding the publication of this paper.

References

[1] J. Smith, J. Everhart, W. Dickson, W. Knowler, and R. Johannes, *Using the ADAP learning algorithm to forecast the onset of diabetes mellitus.* In Proceedings of the Symposium on Computer Applications in Medical Care (Washington, 1988), ed. R. A. Greenes, 261–265. Los Alamitos, CA: IEEE Computer Society Press (1988).

[2] A. Sklar, Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231 (1959).

[3] B. Schweizer and A. Sklar, Espaces metriques aleatoires, *C. R. Acad. Sci. Paris*, **247**, 2092–2094 (1958).

[4] S. Mondal and D. Kundu, A bivariate inverse Weibull distribution and its application in complementary risks model, *Journal of Applied Statistics*, **47**, 1084-1108 (2020).

[5] Y. Wen, A. Yang, X. Kong, and Y. Su, A bayesian-model-averaging copula method for bivariate hydrologic correlation analysis, *Frontiers in Environment Science*, **9**, 1-13 (2022).

[6] H. Bekrizadeh and B. Jamshidi, A new class of bivariate copulas: Dependence measures and properties, *Metron*, **75** 31-50 (2017).

[7] F. Aldhufairi and J. Sepanski, New families of bivariate copulas via unit weibull distortion, *Journal of statistical distributions and applications*, **7**, 1-20 (2020).

[8] M. Seyam, Construct a new family of bivariate copula, *Technology Reports of Kansai University*, **62**, 6035-6039 (2020).

[9] M. Abd El-Monsef, S. Elsobky and M. Seyam, The relations between some families of copulas, *Journal of Advances in Mathematics and Computer Science*, **36**, 31-49 (2021).

[10] R. Nelsen, *An Introduction to Copula, Second Edition*, Springer, New York, USA (2006).

[11] P. Mah and M. Shitan, Construction of a new bivariate copula based on Rüschedorf method, *Applied Mathematical Sciences*, **8,** 7645 – 7658 (2014).
[12] M. Frank, On the simultaneous associativity of F(x,y) and x+y-F(x,y), *Aequationes Mathematicae*, **21**, 194-226 (1979)
[13] D. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, **65**, 141-151 (1978).
[14] X. Li and R. Fang, A new family of bivariate copulas generated by univariate distributions, *Journal of Data Science*, **10**, 1-17 (2012).