

Cumhuriyet Science Journal

e-ISSN: 2587-246X ISSN: 2587-2680 Cumhuriyet Sci. J., 42(1) (2021) 201-208 http://dx.doi.org/10.17776/csj.753556



Construction of a bivariate copula by Rüschendorf's method

Mehmet YILMAZ^{1,*}, Muhammet BEKÇİ²

¹Ankara University, Faculty of Science, Department of Statistics, Ankara/TURKEY

²Sivas Cumhuriyet University, Faculty of Science, Department of Statistics and Computer Sciences, Sivas/TURKEY

Abstract

In this paper, a new copula model with given unit marginals is proposed, based on Rüschendorf's Method. A new bivariate copula family is introduced by adding a proper term to independence copula. Thus, we avoid the complexity of the proposed copula model. By choosing a baseline copula from the same marginal, we derive a new copula that can approach from above towards the independence copula. Furthermore, it is established that a bivariate copula constructed by this method allows some flexibility in the dependence measure according to Spearman's correlation coefficient. Additionally, tail dependence measures are investigated. Illustrative examples are given taking into account the specific choices of a baseline copula.

Article info

History: Received:16.06.2020 Accepted:18.01.2021

Keywords: Dependence, Rüschendorf's method, Bivariate copula, Fréchet bounds, Spearman's rho.

1. Introduction

When creating new bivariate copula models, researchers generally try to obtain models that can express the high correlation. Lai and Xie [1] studied on construction of continuous bivariate positive quadrant dependent distributions. Bairamov et al. [2] provided an extension of the maximal correlation coefficient for the Farlie-Gumbel-Morgenstern (FGM) family. Firstly, we discuss the necessary conditions to construct a new copula. Secondly, we obtain bounds for dependence measure in accordance with Spearman's rank correlation coefficient. At last, the usefulness of this family is discussed by considering illustrative examples.

The genesis of the proposal is based on both works of [1] and [3]. Accordingly, a function g(u, v) can be defined on the unit square as g(u, v) = uv + k(u, v), where k(u, v) = 0 at the endpoints of the unit square, with $\frac{-\partial^2 k(u,v)}{\partial u \partial v} \le 1$.

Let C(u, v) denote the bivariate copula function. Then, according to the eq. (2) and the condition (3) of [1], and Rüschendorf's method, the function k(u, v) is chosen as $k(u, v) = uv \overline{C}(u, v)$, where \overline{C} denotes survival copula.

The following theorem shows that g(u, v) meets the conditions (3)-(5) given in [1].

Theorem 1. $C_H(u, v) = uv(1 + \overline{C}(u, v))$ is a well-defined copula function.

Proof. According to Definition 2.2.2 of [4], a bivariate copula must satisfy the following properties:

(P1) For every u, v in [0,1], it is obvious that $\lim_{u \to 0} C_H(u, v) = \lim_{u \to 0} uv (1 + \bar{C}(u, v)) = 0, \qquad \lim_{v \to 0} C_H(u, v) = \lim_{v \to 0} uv (1 + \bar{C}(u, v)) = 0$ and $\lim_{u \to 1} C_H(u, v) = \lim_{u \to 1} uv (1 + \bar{C}(u, v)) = v, \qquad \lim_{v \to 1} C_H(u, v) = \lim_{v \to 1} uv (1 + \bar{C}(u, v)) = u,$ $\lim_{u \to v \to 1} C_H(u, v) = \lim_{u \to v \to 1} uv (1 + \bar{C}(u, v)) = 1.$

(P2) We prove the continuous case only.

^{*}Corresponding author. email address: mehmetyilmaz@ankara.edu.tr

http://dergipark.gov.tr/csj ©2021 Faculty of Science, Sivas Cumhuriyet University

$$c_H(u,v) = \frac{\partial^2 C_H(u,v)}{\partial u \partial v} = 1 + \bar{C}(u,v) + v \frac{\partial \bar{C}(u,v)}{\partial v} + u \frac{\partial \bar{C}(u,v)}{\partial u} + u v \frac{\partial^2 \bar{C}(u,v)}{\partial u \partial v}.$$
 (1)

Now, by noting that $\bar{C}(u,v) = 1 - u - v + C(u,v)$, $\frac{\partial \bar{C}(u,v)}{\partial u} = -1 + \frac{\partial C(u,v)}{\partial u}$ and $\frac{\partial^2 \bar{C}(u,v)}{\partial u \partial v} = c(u,v) \ge 0$. Then eq. (1) can be rewritten as

$$c_H(u,v) = 1 - u - v + \bar{\mathcal{C}}(u,v) + v \frac{\partial \mathcal{C}(u,v)}{\partial v} + u \frac{\partial \mathcal{C}(u,v)}{\partial u} + uvc(u,v).$$
(2)

Assume that the copula C(u, v) belongs to the class of negative dependent copulas, and then it is easy to conclude that from corollary 5.2.6 of [4], each of the inequalities below holds:

$$u\frac{\partial C(u,v)}{\partial u} \ge C(u,v),$$

and

$$v\frac{\partial C(u,v)}{\partial v} \geq C(u,v).$$

Using both inequalities, a lower bound can be achieved as

$$c_{H}(u,v) \ge 1 - u - v + C(u,v) + \bar{C}(u,v) + C(u,v) + uvc(u,v)$$

= $2\bar{C}(u,v) + C(u,v) + uvc(u,v)$
 $\ge 0.$

Hence, under the assumption of the negative dependence, it is showed that (P2) holds.

Now, assume that the copula C(u, v) belongs to the class of positive dependent copulas. Thus, the eq. (1) can be rewritten as follows:

$$c_{H}(u,v) = \left(1 + u\frac{\partial\bar{C}(u,v)}{\partial u}\right) \left(1 + v\frac{\partial\bar{C}(u,v)}{\partial v}\right) - uv\frac{\partial\bar{C}(u,v)}{\partial u}\frac{\partial\bar{C}(u,v)}{\partial v} + \bar{C}(u,v) + uvc(u,v).$$
(3)

According to Definition 1.2 of [5], and Theorem 5.2.10 and Theorem 5.2.15 of [4], the positive dependence implies

$$\frac{\partial \bar{\mathcal{C}}(u,v)}{\partial u} \frac{\partial \bar{\mathcal{C}}(u,v)}{\partial v} \leq c(u,v) \bar{\mathcal{C}}(u,v).$$

Then by applying the latter inequality in eq. (3), we have the following lower bound

$$c_{H}(u,v) \ge \left(1 + u \frac{\partial \bar{C}(u,v)}{\partial u}\right) \left(1 + v \frac{\partial \bar{C}(u,v)}{\partial v}\right) + \bar{C}(u,v) + uvc(u,v) \left(1 - \bar{C}(u,v)\right).$$
(4)

Note that each summand in eq. (4) is nonnegative thus $c_H(u, v) \ge 0$ is obtained. This completes the proof.

Remark 1. Copula family $C_H(u, v) = uv(1 + \overline{C}(u, v))$ belongs to the class of positive dependent copulas since $uv \le uv(1 + \overline{C}(u, v))$ for all $(u, v) \in [0, 1]^2$.

2. Lower and Upper Bounds on Spearman's Rho Measure for C_H

This section is about obtaining the lower and upper bounds of the measure of dependency for the proposed bivariate copula family. According to [6] and [7], for any bivariate copula defined on the unit square contains Fréchet lower and upper bounds, which respectively are defined by

$$C^{-}(x,y) = \max\{u + v - 1,0\}$$
(5)

$$C^{+}(x, y) = \min\{u, v\}.$$
 (6)

For C(u, v), Spearman's rho can be expressed as

$$\rho = 12 \int_{0}^{1} \int_{0}^{1} \{C(u,v) - uv\} dv du$$
(7)

(see, [8]). The coefficient of Spearman's rho for the new family can be obtained by

$$\rho_{C_{H}} = 12 \int_{0}^{1} \int_{0}^{1} uv \bar{C}(u, v) dv du$$

$$= 12 \int_{0}^{1} \int_{0}^{1} uv [1 - u - v + C(u, v)] dv du.$$
(8)

Hence, by using eq. (5), the lower bound is as follows

$$\rho_{C_H} \ge -1 + 12 \iint_{u+v-1>0} uv[u+v-1]dvdu = \frac{1}{10}.$$

As we expect from Corollary 1, this lower bound must be positive. We can say that C_H achieves weakly positive dependence from its lower bound. To obtain the upper bound, we use eq. (6) then

$$\rho_{C_H} \le -1 + 12 \iint_{v > u} u^2 v dv du + \iint_{u > v} uv^2 dv du = \frac{3}{5}$$

As a result, ρ_{C_H} lies in the interval $\left[\frac{1}{10}, \frac{3}{5}\right]$. The proposed copula C_H can only detect positive dependence. Furthermore, C_H can achieve a high correlation through positive dependence.

3. Tail Dependence Measures of C_H

The tail dependence measures can detect how likely both components jointly exceed extreme quantiles. It is useful in determining the common behavior of random variables in the upper and lower quadrants. Upper and lower tail dependence coefficients are defined by

$$\lambda_u = 2 - \lim_{t \to 1} \frac{1 - \mathcal{C}(t, t)}{1 - t}, \lambda_L = \lim_{t \to 0} \frac{\mathcal{C}(t, t)}{t}$$

(see [4], subsection 5.4). Accordingly, after some algebraic calculation, the upper and the lower dependence coefficients for $C_H(u, v)$ are respectively equal to $\lambda_u = \lim_{t \to 1} \frac{-d\bar{c}(t,t)}{dt}$ and $\lambda_L = 0$. Next, an example is given to illustrate this family.

4. Illustrative Examples

In this section, we compare in terms of Spearman's rho correlation coefficient the proposed copula with respect to its baseline copula chosen from the most used copulas in modeling real data.

Example 1. (FGM copula) The Farlie-Gumbel-Morgenstern (FGM) family of a bivariate copula is given by $C(u, v) = uv[1 + \theta (1 - u)(1 - v)]$, for $\theta \in [-1, 1]$ (see [9] and [10]). Then, the copula $C_H(u, v)$ is given by

$$C_H(u, v) = uv[1 + (1 - u)(1 - v)(1 + \theta uv)].$$

Hence, $\rho_{C_H} = \frac{1}{12}(\theta + 4)$. Since $\theta \in [-1,1], \frac{1}{4} \le \rho_{C_H} \le \frac{5}{12}$. One can conclude that this family can model weakly positive dependence as FGM does. Furthermore, tail dependence coefficients of $C_H(u, v)$ are equal to zero as FGM has. Hence, $C_H(u, v)$ has no tail dependence in both directions.

Example 2. (Gumbel-Hougaard copula) The bivariate version of the Gumbel-Hougaard family copula is given by

$$C(u,v) = e^{-[(-\log(u))^{\theta} + (-\log(v))^{\theta}]^{\frac{1}{\theta}}},$$

where θ lies in the interval $[1, \infty)$ (see, [10-13]). Since the Gumbel–Hougaard copula can detect positive dependence, it is used in modeling bivariate flood frequency, storm, drought, etc. (see, [14]). [15] claimed that no closed-form expression exists for Spearman's rho. [16] obtained an integral form for Spearman's rho as

$$\rho_{GH} = \frac{12}{\theta} \int_0^1 \frac{[t(1-t)]^{\frac{1}{\theta}-1}}{\left[1+t^{\frac{1}{\theta}}+(1-t)^{\frac{1}{\theta}}\right]^2} dt - 3$$

Calculated values of Spearman's rho for the base model (Gumbel–Hougaard copula) and $C_H(u, v)$ are given in Table 1.

Table 1. Calculated values of Spearman's rho of Gumbel-Hougaard and C_H copulas for some values of θ

θ	1.1	1.3	1.5	1.7	1.9	2.1	3	5	15
ρ_{GH}	0.1353	0.3368	0.4767	0.5773	0.6520	0.7088	0.8488	0.9432	0.9935
$ ho_{C_{H}}$	0.3729	0.4300	0.4683	0.4951	0.5147	0.5293	0.5644	0.5870	0.5986

As can be seen from Table 1, when θ is close to 1, C_H seems slightly more dominant than Gumbel-Hougaard in terms of Spearman's rho. However, Gumbel-Hougaard is more dominant in the large values of the θ .

Gumbel-Hougaard copula family has upper tail dependence measured as $\lambda_U = 2 - 2^{\frac{1}{\theta}}$ (see, [4] p.215). After some algebraic calculation, C_H is similarly found to have the same upper tail dependency measure.

Example 3. (Ali-Mikhail-Haq copula) Ali-Mikhail-Haq (AMH) copula is given by

$$C(u,v) = \frac{uv}{1-\theta(1-u)(1-v)},$$

where $\theta \in [-1,1]$ (see, [17, 18]). Spearman's rho coefficient for this family is given below:

$$\rho_{AMH} = \frac{12(1+\theta)dilog(1-\theta) - 24(1-\theta)\log(1-\theta)}{\theta^2} - \frac{3(\theta+12)}{\theta},$$

where $dilog(\cdot)$ represents the dilogarithm function defined as $dilog(x) = \int_{1}^{x} \frac{\log(t)}{1-t} dt$ (see, [4] p. 172 and [19]). They also reported that $\rho_{AMH} \in [-0.2711, 0.4784]$. We use the geometric series representation for $1/(1 - \theta(1 - u)(1 - v))$ to calculate the correlation coefficient numerically for $C_H(u, v)$. Then, we have

$$\rho_{C_H} = 48 \sum_{j=0}^{\infty} \frac{\theta^j}{((j+3)(j+2)(j+1))^2} - 1.$$

Calculated approximate values of Spearman's rho for the base model AMH copula and $C_H(u, v)$ are tabulated as in Table 2.

Table 2. Calculated values of Spearman's rho of AMH and C_H copulas for some values of θ

θ	-1	7	4	-0.1	0.1	0.4	0.7	1
ρ _{ΑΜΗ}	-0.2711	-0.2004	-0.1216	-0.0325	0.0342	0.1490	0.2896	0.4784
$\rho_{C_{H}}$	0.2608	0.2806	0.3019	0.32513	0.3418	0.3691	0.3997	0.4353

As can be seen from Table 2, for nonnegative small values of θ , C_H seems more dominant than the AMH copula in terms of Spearman's rho. However, the AMH copula is more dominant in the large positive values of the θ . For negative values of θ , C_H produces a positive correlation, unlike AMH copula does.

AMH copula family has lower tail dependence as $\lambda_L = \frac{1}{2}$ for $\theta = 1$ reported by [19]. After some algebraic calculation, C_H has no tail dependency measure.

Example 4. (Clayton copula) The Clayton copula is first introduced by [20]. The bivariate version of the Clayton copula is given by

 $C_{Clayton}(u,v) = \left(u^{-\frac{1}{\theta}} + v^{-\frac{1}{\theta}} - 1\right)^{-\theta},$

where $\theta > 0$. This copula can only model positively associated variables. Spearman's rho coefficient for the Clayton copula is more complex. We calculate numerically to obtain these coefficients for both Clayton and C_H copulas. Table 3 tabulates Spearman's rho coefficients for various values of θ .

θ	0.1	0.3	0.6	0.9	3	5	8	10
$\rho_{Clayton}$	0.9583	0.8100	0.6300	0.5095	0.2124	0.1356	0.0881	0.0714
$\rho_{C_{H}}$	0.5798	0.5254	0.4732	0.4427	0.3763	0.3605	0.3509	0.3475

Table 3. Calculated values of Spearman's rho of Clayton and C_H copulas for some values of θ

As can be seen from Table 3, while θ increases, the correlation coefficient of both copula decreases. In terms of Spearman's rho, while Clayton copula is dominant for the small values of θ , the C_H is dominant for the large values.

Clayton copula family has lower tail dependence measured as $\lambda_L = 2^{\frac{-1}{\theta}}$ (see, [4], p.215). After some algebraic calculation, C_H has no tail dependency measure.

Example 5. (Frank copula) The Frank copula is given by

$$C_{Frank}(u,v) = \frac{-1}{\theta} \log\left(1 + \frac{\left(e^{-\theta u} - 1\right)\left(e^{-\theta v} - 1\right)}{\left(e^{-\theta} - 1\right)}\right),$$

where $\theta \in (-\infty, +\infty)$ (see, [21]). This copula can symmetrically model both directions of dependence. Spearman's rho coefficient for the Frank copula is given as

$$\rho_{Frank} = 1 - \frac{12}{\theta} [D_1(\theta) - D_2(\theta)],$$

where $D_k(x)$ is the Debye function which is defined by $\frac{k}{x^k} \int_0^x \frac{z^k}{e^z - 1} dz$ for any positive integer k (see [4], p. 171). To obtain this coefficient for both Frank and C_H copulas, we compute numerically. Table 4 gives tabulated values of Spearman's rho coefficients for various values of θ .

Table 4. Calculated values of Spearman's rho of Frank and C_H copulas for some values of θ

θ	-10	-8	-5	-0.8	-0.3	0.3	0.8	5	8	10
ρ_{Frank}	8602	8035	6435	1322	0499	.0499	.1322	.6435	.8035	.8602
$\rho_{C_{H}}$.1283	.1407	.1772	.3005	.3209	.3458	.3666	.4990	.5425	.5584

As can be seen from Table 4, while the C_H copula is dominant for the small values of the $\theta \ge 0$, the Frank copula is dominant for the positively large values of θ in terms of Spearman's rho.

Frank copula has no tail dependence measure (see, [4], p.215). After some algebraic calculation, C_H has also no tail dependence measure.

Example 6. (**Bivariate Gumbel-Exponential (BGE) copula**) The bivariate version of the Gumbel-Exponential (BGE) copula is given by

 $C(u,v) = u + v - 1 + (1-u)(1-v)e^{-\theta \log(1-u)\log(1-v)},$

for $\theta \in [0,1]$. This copula is also known as Gumbel-Barnett copula (see, [10, 22] and [4], p. 23). BGE copula can model only negative dependence. According to [23], the Spearman's rho coefficient of BGE copula is

$$\rho_{BGE} = 12 \left[-\frac{e^{\frac{4}{\theta}}}{\theta} Ei\left(-\frac{4}{\theta}\right) - \frac{1}{4} \right],$$

where $Ei(\cdot)$ is the exponential integral function. After some algebraic manipulation, ρ_{C_H} can be obtained as

$$\rho_{C_H} = 12 \frac{e^{\frac{4}{\theta}}}{\theta} \Big[-Ei\left(-\frac{4}{\theta}\right) + 2e^{\frac{2}{\theta}}Ei\left(-\frac{6}{\theta}\right) - e^{\frac{5}{\theta}}Ei\left(-\frac{9}{\theta}\right) \Big].$$

Calculated values of Spearman's rho for the base model BGE copula and $C_H(u, v)$ are tabulated as in Table 5.

θ	0.1	0.3	0.5	0.7	0.9	1
ρ_{BGE}	-0.0715	-0.1972	-0.3053	-0.4002	-0.4848	-0.5239
$ ho_{C_{H}}$	0.3119	0.2769	0.2494	0.2270	0.2083	0.2000

Table 5. Calculated values of Spearman's rho of BGE and C_H copulas for some values of θ

As can be seen from Table 5, both BGE and C_H decreases in θ in terms of Spearman's rho. Note that, when θ goes to zero, ρ_{C_H} approaches to 1/3. Note also that both BGE and C_H have no tail dependence. In this case, we can say that both positive and negative associated random variables can be modeled by a mixture of copulas C_H and C_{BGE} . According to Subsection 3.2.4 of [4], for $\delta \in [0,1]$, we can write a convex combination of C_{BGE} and C_H as follows

 $C^*(u, v) = \delta C_H(u, v) + (1 - \delta)C_{BGE}(u, v).$ Hence, ρ^* lies in interval [-.52, 0.33].

4. Conclusion

In this study, based on Rüschendorf's Method, we proposed a new bivariate copula distorting independence copula by adding the baseline copula. The baseline copula with a negatively correlated structure certainly transforms into a positively correlated structure in the proposed copula according to the different copula families. However, there may be differences in the direction of change of the correlation coefficients of the proposed copula constructed from baseline copulas with positively correlated structure. This situation is influenced by the presence or absence of lower and upper tail dependencies of the copulas. If we pay attention to the upper and lower bounds of ρ_{C_H} , the values of the Spearman's rho correlation coefficient for this copula family lie in the interval [0.1, .60]. Besides, as a result of illustrative examples, it can be said that except for the extreme value copulas, generated copulas using this method achieve reasonable correlations considering some baseline copulas of which correlation coefficients less than 0.6.

In particular, as illustrated in Example 6, if mixing the proposed copula with a copula that can only model the negative dependence is made, a mixture copula that can model both positive and negative dependencies can be created.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

- [1] Lai C. D., Xie M., A New Family of Positive Quadrant Dependent Bivariate Distributions, *Statistics and Probability Letters*, 46 (4) (2000) 359-364.
- [2] Bairamov I., Kotz S., Bekçi M., New generalized Farlie-Gumbel-Morgenstern distributions and concomitants of order statistics. *Journal of Applied Statistics*, 28 (5) (2001) 521-536.
- [3] Rüschendorf L., Construction of multivariate distributions with given marginals. Annals of the Institute of Statistical Mathematics, 37 (1985) 225-233.
- [4] Nelsen R. B., An Introduction to Copulas. Second Edition. Springer, New York (2006).
- [5] Asadian N., Amini M., Bozorgnia A. Some concepts of negative dependence for bivariate distributions with applications. *Journal of Mathematical Extension*, 4 (1) (2009) 43-59.
- [6] Hoeffding W., Masstabinvariante Korrelationstheorie, Schriften des Mathematischen Instituts und Instituts fur Angewandte Mathematik der Universitat Berlin, 5 (1940) 181-233.
- [7] Fréchet M., Sur Les Tableaux de Corrélation Dont Les marges Sont Donnees, Annales de l'Université de Lyon, Sciences 4 (1951) 13-84.
- [8] Schweizer B., Wolff E., On Nonparametric Measures of Dependence for Random Variables, *The Annals of Statistics*, 9(4) (1981) 879-885.
- [9] Farlie D., The Performance of Some Correlation Coefficients for a General Bivariate Distribution, *Biometrika*, 47 (3/4) (1960) 307-323.
- [10] Gumbel E. J., Bivariate Exponential Distributions, *Journal of the American Statistical Association*, 55 (292) (1960a) 698-707.

- [11] Gumbel E. J., Bivariate Logistic Distributions, Journal of the American Statistical Association, 56 (1961) 335-349.
- [12] Hougaard P., A Class of Multivariate Failure Time Distributions, *Biometrika*, 73 (1986) 671-678.
- [13] Joe H., Multivariate Models and Dependence Concepts. Chapman and Hall, London (1997).
- [14] Singh V.P., Zhang L. Copula–Entropy Theory for Multivariate Stochastic Modeling in Water Engineering, *Geosci. Lett.*, 5(6) (2018).
- [15] Frees E.W., Valdez E.A., Understanding Relationships Using Copulas, North American Actuarial Journal, 2 (1) (1998) 1-25.
- [16] Hürlimann W., Properties and Measures of Dependence for the Archimax Copula, *Advances and Applications in Statistis*, 5 (2005) 125-143.
- [17] Ali M.M., Mikhail N.N., Haq M.S., A Class of Bivariate Distributions Including the Bivariate Logistic, *Journal of Multivariate Analysis*, 8 (1978) 405-412.
- [18] Balakrishnan N., Lai C. D., Continuous Bivariate Distributions. Springer Science & Business Media, (2009).
- [19] Kumar P., Probability Distributions and Estimation of Ali-Mikhail-Haq Copula, *Applied Mathematical Sciences*, 4 (14) (2010) 657-666.
- [20] Clayton D., A Model for Association in Bivariate Life Tables and its Application in Epidemiological Studies of Family Tendency in Chronic Disease Incidence, *Biometrika*, 65 (1978) 141-151.
- [21] Frank M.J., On the Simultaneous Associativity of F(x,y) and x + y F(x,y), *Aequationes Math*, 19 (1979) 194-226.

- [22] Barnett V., Some Bivariate Uniform Distributions, *Communications in Statistics: Theory and Methods*, 9 (1980) 453-461.
- [23] Fredricks G.A., Nelsen R.B., On the Relationship Between Spearman's Rho and Kendall's Tau for Pairs of Continuous Random Variables, *Journal* of Statistical Planning and Inference, 137 (7) (2007) 2143-2150.