Cumhuriyet Science Journal

e-ISSN: 2587-246X
ISSN: 2587-2680

Cumhuriyet Sci. J., 41(4) (2020) 938-943
http://dx.doi.org/10.17776/csj.742159

A new approach to bivariate transmutation: construction of continuous bivariate distribution under negative dependency

Hüseyin ÜNÖZKAN [1,*], *Mehmet YILMAZ* [2]

[1]*Ankara University, Graduate School of Natural and Applied Science, Department of Statistics, Ankara / TURKEY*

[2]*Ankara University, Faculty of Science, Department of Statistics, 06100, Ankara / TURKEY*

Abstract

In this study, a new approach to transmutation theory is developed by using negative dependency basement. Once choosing a distribution that has negative dependency with the same marginal, a new bivariate distribution is derived. In this study, we examined a new transmutation technique in which a negative dependency offers a big success in modeling rather than most known and used statistical distributions. This approach clash with classical transmutation methods. In this study at the beginning, the classical transmutation is defined. Later, we introduce the new technique and obtain lower and upper bounds of distribution to show that this approach gives us a distribution. Gaining new bivariate continuous distributions with this technique may be more appropriate in theory, and modeling of some data sets in terms of this approach may be more efficient.

Article info

History:
Received: 24.05.2020
Accepted: 19.08.2020

Keywords:
Transmuted bivariate distribution, Dependence, Bivariate distribution, Negative dependency, Fréchet bounds.

1. Introduction

In both theory and practice in statistics, univariate distributions are generally inadequate to model random phenomena. Besides, bivariate statistical distributions are crucial in modeling in many different data sets. For generating new distribution, there needs to construct a joint distribution by using marginal distributions with higher or lower correlations. Dolati and Ubeda-Flores [1] studied a method by considering pairs of order statistics. Lai and Xie [2] examined continuous bivariate distributions which possess the Positive Quadrant Dependency. According to similar works [3], [4] some conditions for negative dependency decided.

Though many different methods in generating new distribution, studies on transmuted distributions have become popular. Although Shaw and Buckley [5] offered transmutation as alternative technique for copulas in generating new distribution, today transmutation becomes the first approach in generating distribution.

There are also studies for some other methods for generating distribution. In a study, Shaw and Buckley [5] joined an inverse of a statistical distribution with another distribution. In that study authors offered for generating new distributions by a new formulation. Quadratic rank order transmutation is studied for alternating to copulas.

After these studies, some basic distributions were used in generating distributions [6,7]. Later transmutation with many distributions was studied [8-10]. To derive transmuted distributions the transformation below is used.

$$F_{QRT} = (1+\lambda)G - \lambda G^2, \quad |\lambda| \le 1 \qquad (1)$$

*Corresponding author. *e-mail address: hunozkan@gmail.com*

In that studies by using the Eq. (1), the new distributions were obtained and probability density functions, moments, parameter estimations and hazard rates were examined [11-15].

Different from these studies in [16], by using a conditional Farlie-Gumbel-Morgenstern copula with exponential marginals is studied. In this study, a new method for transmutation is introduced. Inspired by these studies, we desire to propose a simpler but useful model.

Theorem1. Assume $H(x,y)$ is a bivariate distribution function belonging to the distribution family $\mathcal{F}(F,G)$, $F(x)$ and $G(y)$ are marginal distributions, and $H(x,y)$ is differentiable on $\mathbb{R}^2$, and $h(x,y)=\frac{\partial^2 H(x,y)}{\partial x \partial y}$ denotes the joint probability density function. Then $H_1^*(x,y)=H(x,y)(1+\bar{H}(x,y))$ is a distribution function if $H(x,y)\leq F(x)G(y)$, for all $(x,y)\in\mathbb{R}^2$, $\bar{H}(x,y)$ is survival function of this bivariate distribution.

Proof. According to Barlow and Proschan [17], any bivariate distribution function holds the following properties (see, Chapter 5):

(P1)

$$\lim_{x\to\infty} H(x,y)(1+\bar{H}(x,y)) = G(y),$$

$$\lim_{y\to\infty} H(x,y)(1+\bar{H}(x,y)) = F(x),$$

$$\lim_{x\wedge y\to\infty} H(x,y)(1+\bar{H}(x,y)) = 1.$$

(P2) $\frac{\partial H_1^*(x,y)}{\partial x}\geq 0$ and $\frac{\partial H_1^*(x,y)}{\partial y}\geq 0$. Then

$$\frac{\partial H_1^*(x,y)}{\partial x}=\frac{\partial H(x,y)}{\partial x}\big(1+\bar{H}(x,y)\big)+H(x,y)\frac{\partial \bar{H}(x,y)}{\partial x}$$

$$=\frac{\partial H(x,y)}{\partial x}\big(1+\bar{H}(x,y)\big)+H(x,y)\left(-h(x)+\frac{\partial H(x,y)}{\partial x}\right)$$

$$=\frac{\partial H(x,y)}{\partial x}\big(1+\bar{H}(x,y)+H(x,y)\big)-h(x)H(x,y)$$

Now, by noting that negative dependence implies $\frac{\partial H(x,y)}{\partial x}H(x)\geq h(x)H(x,y)$. Using this inequality, we have

$$\frac{\partial H_1^*(x,y)}{\partial x}\geq \frac{\partial H(x,y)}{\partial x}\big(1-H(x)+\bar{H}(x,y)+H(x,y)\big)$$

$\geq 0.$

Obviously, $\frac{\partial H_1^*(x,y)}{\partial y}\geq 0.$

(P3) $\frac{\partial^2 H_1^*(x,y)}{\partial x\partial y}\geq 0$. Then

$$\frac{\partial^2 H_1^*(x,y)}{\partial x\partial y}=h(x,y)\big(1+H(x,y)+\bar{H}(x,y)\big)+\frac{\partial H(x,y)}{\partial x}\frac{\partial \bar{H}(x,y)}{\partial y}+\frac{\partial H(x,y)}{\partial y}\frac{\partial \bar{H}(x,y)}{\partial x}. \quad (2)$$

According to Domma (2011) [18], and Kimeldorf and Sampson (1989) [3], by noting that negative dependence implies $P_H(I_1,J_1)P_H(I_2,J_2)\leq P_H(I_2,J_1)P_H(I_1,J_2)$ for all $I_1<I_2$ and $J_1<J_2$. Under the

specially chosen numbers, $\epsilon > 0$ and $\delta > 0$, $I_1 = (x, x+\epsilon]$, $I_2 = (x+\epsilon, \infty)$ and for $J_1 = (-\infty, y]$, $J_2 = (y, y+\delta]$, we have

$$P_H(X \in (x, x+\epsilon], Y \in (-\infty, y])P_H(X \in (x+\epsilon, \infty), Y \in (y, y+\delta])$$

$$\leq P_H(X \in (x+\epsilon, \infty), Y \in (-\infty, y])P_H(X \in (x, x+\epsilon], Y \in (y, y+\delta]).$$

When both sides divided with $\epsilon\delta$ and taking double limiting both sides while $\epsilon \to 0$ and $\delta \to 0$ the inequality turns below:

$$\frac{\partial H(x,y)}{\partial x}\left(\frac{-\partial \overline{H}(x,y)}{\partial y}\right) \leq P_H(X > x, Y \leq y)h(x,y). \quad (3)$$

Easily seen that if we specially choose of $I_1 = (-\infty, x], I_2 = (x, x+\epsilon]$ and $J_1 = (y, y+\epsilon], J_2 = (y+\epsilon, \infty)$, the inequality which is below is also valid.

$$\frac{\partial H(x,y)}{\partial y}\left(\frac{-\partial \overline{H}(x,y)}{\partial x}\right) \leq P_H(X \leq x, Y > y)h(x,y) \quad (4)$$

On the other hand, since $P(X > x, Y \leq y) + P(X \leq x, Y > y) + H(x,y) + \overline{H}(x,y) = 1$, the inequality is obtained as $P(X > x, Y \leq y) + P(X \leq x, Y > y) \leq 1$. By using this latter inequality and combining with the Eq. (3) and (4), we have

$$\frac{\partial H(x,y)}{\partial x}\left(\frac{-\partial \overline{H}(x,y)}{\partial y}\right) + \frac{\partial H(x,y)}{\partial y}\left(\frac{-\partial \overline{H}(x,y)}{\partial x}\right) \leq h(x,y). \quad (5)$$

By considering the Eq. (5), the lower bound is as follows:

$$\frac{\partial^2 H_1^*(x,y)}{\partial x \partial y} \geq h(x,y)\big(H(x,y) + \overline{H}(x,y)\big)$$

$$\geq 0.$$

The proof is completed.

Inspired by Dolati and Úbeda-Flores [1], an alternative distribution where we can write the convex combination with H_1^* is $H_1 = H(x,y)(1 - \overline{H}(x,y))$. In this way, two different distributions belonging to the distribution family $\mathcal{F}(F,G)$ are obtained. Then we may obtain a new distribution under their convex combinations for $\delta \in [0,1]$ as follows:

$$H_2(x,y) = \delta H_1^*(x,y) + (1-\delta)H_1(x,y)$$

$$= \delta H(x,y)\big(1 + \overline{H}(x,y)\big) + (1-\delta)H(x,y)\big(1 - \overline{H}(x,y)\big)$$

$$= H(x,y) + (2\delta - 1)H(x,y)\overline{H}(x,y).$$

By letting $2\delta - 1 = \lambda$, the new distribution proposal is

$$H_2(x,y) = H(x,y) + \lambda H(x,y)\overline{H}(x,y), \quad \lambda \in [-1,1]. \quad (6)$$

Note that, $H(x,y)$ belongs to negatively dependent subclass or independent subclass of the family $\mathcal{F}(F,G)$. Respectively, the joint probability density function and the survival function are

$$h_2(x,y) = h(x,y)\left(1 + \lambda\big(H(x,y) + \overline{H}(x,y)\big)\right)$$

$$+\lambda\left[\frac{\partial H(x,y)}{\partial x}\frac{\partial \overline{H}(x,y)}{\partial y}+\frac{\partial H(x,y)}{\partial y}\frac{\partial \overline{H}(x,y)}{\partial x}\right], \tag{7}$$

$$\overline{H}_2(x,y)=\overline{H}(x,y)+\lambda\, H(x,y)\overline{H}(x,y), \quad \lambda\in[-1,1]. \tag{8}$$

Therefore, we finally propose distribution family with simple structures like Farlie-Gumbel-Morgenstern (FGM) distribution family (see, Farlie [19] and Gumbel [20]).

2. Hazard Rate and Reversed Hazard Rate Functions of H_2

The bivariate reversed hazard rate function is defined as $rv(x,y)=f(x,y)/F(x,y)$ [21]. Furthermore, Basu [22] defines the bivariate hazard rate as $r(x,y)=f(x,y)/\overline{F}(x,y)$. By using Eq. (6-8) we may gain a close form for both r_{H_2} and rv_{H_2}.

$$r_{H_2}(x,y)=\frac{h(x,y)\left(1+\lambda\left(H(x,y)+\overline{H}(x,y)\right)\right)+\lambda\left[\frac{\partial H(x,y)}{\partial x}\frac{\partial \overline{H}(x,y)}{\partial y}+\frac{\partial H(x,y)}{\partial y}\frac{\partial \overline{H}(x,y)}{\partial x}\right]}{\overline{H}(x,y)+\lambda\, H(x,y)\overline{H}(x,y)}$$

$$=r_H(x,y)\left(1+\frac{\lambda\overline{H}(x,y)}{1+\lambda H(x,y)}\right)-\left(1-\frac{1}{1+\lambda H(x,y)}\right)\left(r_H(x|y)rv_H(y|x)+r_H(y|x)rv_H(x|y)\right),$$

where $r_H(y|x)=\frac{\frac{-\partial\overline{H}(x,y)}{\partial x}}{\overline{H}(x,y)}$ and $rv_H(y|x)=\frac{\frac{\partial H(x,y)}{\partial x}}{H(x,y)}$ are respectively conditional hazard rate and conditional reversed hazard rate functions of H with given $X=x$.

$$rv_{H_2}(x,y)=\frac{h(x,y)\left(1+\lambda\left(H(x,y)+\overline{H}(x,y)\right)\right)+\lambda\left[\frac{\partial H(x,y)}{\partial x}\frac{\partial \overline{H}(x,y)}{\partial y}+\frac{\partial H(x,y)}{\partial y}\frac{\partial \overline{H}(x,y)}{\partial x}\right]}{H(x,y)+\lambda H(x,y)\overline{H}(x,y)}$$

$$=rv_H(x,y)\left(1+\frac{\lambda H(x,y)}{1+\lambda\overline{H}(x,y)}\right)-\left(1-\frac{1}{1+\lambda\overline{H}(x,y)}\right)\left(r_H(x|y)rv_H(y|x)+r_H(y|x)rv_H(x|y)\right).$$

3. Spearman's Rho Measure Bounds for H_2

In this section, we consider obtaining lower and upper bounds of dependence measure for H_2 given by the Eq. (6). According to Hoeffding [23] and Fréchet [24], any bivariate distribution $F(x,y)$ belonging to $\mathcal{F}(F,G)$ must contain Fréchet lower and upper bounds. These bounds are respectively defined as

$$F^-(x,y)=\max\{F(x)+G(y)-1,0\} \tag{9}$$

$$F^+(x,y)=\min\{F(x),G(y)\}. \tag{10}$$

For $F\in\mathcal{F}(F,G)$, Spearman's rho measure can be expressed as

$$\rho_s(X,Y)=12\int_{\mathbb{R}}\int_{\mathbb{R}}\{F(x,y)-F(x)G(y)\}dG(y)dF(x) \tag{11}$$

(Schweizer and Wolff [25]). The Spearman's rho correlation coefficient for H_2 can be obtained by

$$\rho_s = 12 \int_{\mathbb{R}} \int_{\mathbb{R}} \big(H(x,y) + \lambda H(x,y)\bar{H}(x,y) - F(x)G(y)\big) dG(y) dF(x).$$

Hence, by using the fact that $H(x,y) \leq F(x)G(y)$, for $\lambda > 0$, the upper bound can be obtained as $\rho_s \leq \frac{\lambda}{3}$. For $\lambda > 0$, by using the Eq. (9), to obtain the lower bound, then the lower bound is $\rho_s \geq 0$. For $\lambda < 0$, by similar algebraic manipulations, ρ_s lies in the interval $\left[-\frac{\lambda}{3}, 0\right]$. Thus, according to sign of λ, bounds can be written as below:

$$\rho_s \in \begin{cases} \left[-\dfrac{\lambda}{3}, 0\right], & \lambda < 0 \\ 0 \quad , & \lambda = 0 \\ \left[0, \dfrac{\lambda}{3}\right] \quad , & \lambda > 0. \end{cases}$$

We have an illustrative example to see the success of dependence modeling for this family.

Example1. The FGM distribution is defined by $H(x,y) = F(x)G(y)[1 + \theta\, \bar{F}(x)\bar{G}(y)]$, for $\theta \in [-1,1]$, (see, Farlie [19] and Gumbel [20]). Because of the assumption of negative dependency, taking $\theta \in [-1,0]$, the distribution H_2 is given by

$$H_2(x,y) = F(x)G(y)[1 + \theta\, \bar{F}(x)\bar{G}(y)]\big[1 + \lambda \bar{F}(x)\bar{G}(y)[1 + \theta\, F(x)G(y)]\big]$$

where $\lambda \in [-1, 1]$ and $\theta \in [-1,0]$. Hence, $\rho_s = \frac{\theta}{3} + \lambda\left(\frac{1}{3} + \frac{\theta}{6} + \frac{\theta^2}{75}\right)$. Since $\theta \in [-1,0]$, ρ_s attains a minimum value as $\frac{-77}{150} \cong -0.513$ at $(\theta, \lambda) = (-1,1)$, and ρ_s attains maximum value as $\frac{1}{3}$ at $(\theta, \lambda) = (0,1)$.

We conclude that this family can detect a weakly positive dependence as much as FGM can. However, even if the base distribution is negative dependent, H_2 can detect both positive dependence and negative dependence. It has a wider correlation coefficient in the negative values of λ than FGM has.

4. Conclusion

In this study, by using transmutation method in bivariate case, we offered a new approach in generating a bivariate continuous distribution using a baseline distribution from the subclass consisting of negatively dependent distributions of $\mathcal{F}(F,G)$. This transmutation restricts to negative dependency and other situations than negative dependence, in theory transmutation does not provide distribution species totally in new generations.

With derivating this new distribution via orders statistics and examining coefficient of Spearman's rho for dependency, we conclude that this new distribution may be more capable than most known and most used bivariate distributions in modeling negative dependency. The new distribution which is suggested in this study gives the same coefficient value of Spearman's rho in positive dependency. Thereby distributions generated with negative dependency conditions are more reliable.

Under negative dependency condition of this study indicates that better modeling than FGM in the same class which is in $\mathcal{F}(F,G)$ is possible.

Conflict of interest

No conflict of interest was declared by the author.

References

[1] Dolati A., Ubeda-Flores M., Constructing Copulas by Means of Pairs of Order Statistics, *Kybernetika*, 45(6) (2009) 992-1002.

[2] Lai C. D., Xie M., A New Family of Positive Quadrant Dependent Bivariate Distributions, *Statistics and Probability Letters,* 46(4) (2000) 359-364.

[3] Kimeldorf G., Sampson A.R., A Framework for Positive Dependence, *Ann. Inst. Statist. Math.,* 41(1) (1989) 31-45.

[4] Han Kwang-Hee., A New Family of Negative Quadrant Dependent Bivariate Distributions with Continuous Marginal, *Journal of the Chungcheong Mathematical Society*, 24 (4) (2011) 795-805.

[5] Shaw W.T., Buckley I.R.C., The Alchemy of Probability Distributions:Beyond Gram-Charlier & Cornish-Fisher Expansions,and Skew-Normal or Kurtotic-Normal Distributions, *Financial Mathematics Group, King's College*, 1 (28) (2007).

[6] Aryal G.R., Tsokos C.P., On the transmuted extreme value distribution with application. Nonlinear Analysis, *Theory, Methods & Applications,* 71 (12) (2009) 1401-1407.

[7] Aryal G.R., Tsokos C.P., Transmuted Weibull Distribution, *European Journal of Pure and Applied Mathematics,* 4 (2) (2011) 89-102.

[8] Merovci F., Transmuted Lindley Distribution, *Int. J. Open Problems Compt. Math*., 6 (2) (2013) 63-72.

[9] Merovci F., Transmuted Exponentiated Exponential Distribution, *Mathematical Sciences and Applications E-Notes*, 1 (2) (2013) 112-122.

[10] Merovci F., Transmuted Rayleigh Distribution, *Austrian Journal of Statistics*, 42 (1) (2013) 21-31.

[11] Aryal G.R., Transmuted Log-Logistic Distribution, *Journal of Statistics Applications & Probability*, 2 (1) (2013) 11-20.

[12] Elbatal I., Diab L.S. and AbdulAlim N.A., Transmuted Generalized Linear Exponential Distribution, *International Journal of Computer Applications*, 83 (17) 82013) 29-37.

[13] Merovci F., Puka L., Transmuted Pareto distribution, *ProbStat Forum*, 7 (2014) 1-11.

[14] Mohamed H.A., Transmuted Exponentiated Gamma Distribution: A Generalization of the Exponentiated Gamma Probability Distribution, *Applied Mathematical Sciences*, 8 (27) (2014) 1297-1310.

[15] Oguntunde E.P., Adejumo O.A., The Transmuted Inverse Exponential Distribution, *International Journal of Advanced Statistics and Probability*, 3 (1) 82015) 1-7.

[16] Ünözkan H., Yılmaz M., A New Method for Generating Distributions: An Application to Flow Data, *International Journal of Statistics and Applications*, 9 (3) (2019) 92-99

[17] Barlow R. E., Proschan, F., Statistical Theory of Reliability and Life Testing: Probability Models, Technical Report, New York: Holt, Rinehart and Winston, 1975.

[18] Domma F., Bivariate Reversed Hazard Rate, Notions, and Measures of Dependence and their Relationships, *Communications in Statistics - Theory and Methods*, 40(6) (2011) 989-999, DOI: 10.1080/03610920903511777.

[19] Farlie D., The Performance of Some Correlation Coefficients for a General Bivariate Distribution, *Biometrika*, 47(3/4) (1960) 307-323.

[20] Gumbel E. J., Bivariate Exponential Distributions, *Journal of the American Statistical Association*, 55-292 (1960) 698-707.

[21] Bismi G., Bivariate Burr Distributions, unpublished PhD Thesis, Cochin University of Science and Technology, 2005.

[22] Basu A.P., Bivariate Failure Rate, *Journal of the American Statistical Association*, 66 (1971) 103–104.
[23] Hoeffding W., Masstabinvariante Korrelationstheorie, *Schriften des Mathematischen Instituts und Instituts fur Angewandte Mathematik der Universitat Berlin*, 5 (1940) 181-233.
[24] Fréchet M., Sur Les Tableaux de Corrélation Dont Les marges Sont Donnees, *Annales de l'Université de Lyon, Sciences*, 4 (1951) 13–84.
[25] Schweizer B. and Wolff E., On Nonparametric Measures of Dependence for Random Variables, *The Annals of Statistics,* 9(4) (1981) 879-885.