Austrian Journal of Statistics
July 2023, Volume 52, 1–14.
http://www.ajs.or.at/
doi:10.17713/ajs.v52i4.1446

Multivariate Asymmetric Distributions of Copula Related Random Variables

Ayyub Sheikhi, Freshteh Arad
Shahid Bahonar University of Kerman,
Kerman, Iran

Radko Mesiar
STU Bratislava, Slovakia,
IRAFM, Ostrava, Czech Republic

Abstract

It is known that normal distribution plays an important role in analysing symmetric data. However, this symmetric assumption may not hold in many real word and in such cases, asymmetric distribution, including skew normal distribution, are known as the best alternative. Constructing asymmetric distributions is carried out using the conditional/selection approach of several independent variable conditioning on other set of variables and this approach does not work well when the independence between variables violated. In this work we construct an asymmetric distribution when variables are dependent using a copula. Specifically, we consider the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ are connected using a copula function $C_{\boldsymbol{X},\boldsymbol{Y}}$ and we study the selection distribution $\boldsymbol{Z} \stackrel{\text{d}}{=} (\boldsymbol{X}|\boldsymbol{Y} \in T)$. We present some special cases of our proposed distribution, among them, multivariate skew-normal distribution. Some properties such as moments and moment generating function are investigated. Also, numerical analysis including simulation study as well as a real data set analysis are presented for illustration.

Keywords: copula, selection distribution, skew-normal, *Gaussian* copula.

1. Introduction and preliminaries

Although many of symmetric distribution such as normal, t, logistic, etc., are common distributions to analyze research data, still data from many applied fields exhibit asymmetric behavior and thus there has been a sustained interest in the literature in constructing asymmetric distributions to analyze data which are far from being symmetric. Most of the existing asymmetric distributions for modeling non-symmetric data, consists of modifying a symmetric probability density function by introducing skewness, among them skew-normal distribution, which has been introduced independently by Roberts (1966), Aigner, Lovell, and Schmidt (1977), Anděl, Netuka, and Zvára (1984) and Azzalini (1985).

It has been said that the variable Z follows a skew normal distribution with parameter λ, is denoted by $Z_{sn} \sim SN(\lambda)$, if its density can be written as

$$f(z) = 2\varphi(z)\Phi(\lambda z), \qquad z \in \mathbb{R}, \tag{1}$$

where $\varphi(.)$ is the standard normal density function, $\Phi(.)$ denotes the standard normal distribution function and λ is the skewness parameter. A general univariate form of these distributions, which is called skew-symmetric family, can be represented as a product of a cumulative distribution function and a density distribution function.

Including the location and scale parameters μ and σ respectively, it is said that the univariate random variable Z_{sn} has a skew-normal distribution with mean μ, variance σ^2 and the skewness parameter λ, if its density can be written as

$$f_\lambda(z) = 2\varphi(z; \mu, \sigma^2)\Phi(\lambda\frac{z-\mu}{\sigma}), \quad z \in \mathbb{R} \tag{2}$$

where $\varphi(.; \mu, \sigma^2)$ is the normal density function with mean μ and variance σ^2, $\Phi(.)$ and λ are as in (1). We adopt the notation $Z_{sn} \sim SN(\mu, \sigma^2; \lambda)$ and it reduces to $Z_{sn} \sim SN(\lambda)$ if $\mu = 0, \sigma = 1$. Also it yields normal distribution $N(\mu, \sigma^2)$ if $\lambda = 0$. One approach for constructing skew distribution functions is based on conditioning random variables. Arnold (2002) Arnold, Beaver, Azzalini, Balakrishnan, Bhaumik, Dey, Cuadras, and Sarabia (2002) has used this approach to obtain a skew normal distribution. Suppose that (X, Y) has a bivariate normal distribution with standardized marginals and correlation ρ, where $\rho \in (-1, 1)$. The conditional density of X given $Y > 0$ follows a skew-normal $SN(\lambda)$ with $\lambda = \rho/\sqrt{(1-\rho^2)}$. See Arnold *et al.* (2002) and references therein for more details and the other scenarios. Moreover substituiting $Y \in T$ instead of $Y > 0$ in the Arnold's proposal, where T is a Borel measurable subset of $\mathbb{R}$, the selection distributions has been introduced Arellano-Valle, Branco, and Genton (2006). More specific, for two random variables $X, Y \in \mathbb{R}$ the selection distribution is defined as the conditional distribution of X given $Y \in T$, i.e., it is said that a random variable $Z \in \mathbb{R}$ has a selection distribution if $Z \overset{\text{d}}{=} (X|Y \in T)$, denoted by $Z \sim SLCT(\theta)$ with parameter(s) θ depending on the characteristics of X, Y, and T. Evidently, a well known special case of these distributions is the skew-normal distribution.

In this context, a multivariate form of these distribution has been investigated in the literature. Due to Arellano-Valle and Azzalini (2006), we say that a d-dimensional random vector $\boldsymbol{Y}$ has a unified multivariate skew-normal distribution, ($\boldsymbol{Y} \sim SUN_{d,m}(\boldsymbol{\xi},\ \boldsymbol{\delta},\ \boldsymbol{\Omega},\ \boldsymbol{\Gamma},\ \boldsymbol{\Lambda})$), if it has a density function of the form

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \phi_d(\boldsymbol{y},\ \boldsymbol{\xi};\ \boldsymbol{\Omega}) \frac{\Phi_m\left(\boldsymbol{\delta} + \boldsymbol{\Lambda}^T\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{\xi});\ \boldsymbol{\Gamma} - \boldsymbol{\Lambda}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}\right)}{\Phi_m(\boldsymbol{\delta};\ \boldsymbol{\Gamma})} \quad y \in \mathbb{R}^d, \tag{3}$$

where $\varphi_d(.,\ \boldsymbol{\xi},\ \boldsymbol{\Omega})$ is the density function of a multivariate normal and $\Phi_m(.;\ \boldsymbol{\Sigma})$ is the multivariate normal cumulative function with the covariance matrix $\boldsymbol{\Sigma}$. Based on the definition (1) of Arellano -Valle et al. (2006), let $\boldsymbol{V} \in \mathbb{R}^p$ and $\boldsymbol{U} \in \mathbb{R}^q$ be two random vectors, and denote by T a measurable subset of $\mathbb{R}^q$. The selection distribution is defined as the conditional distribution of $\boldsymbol{V}$ given $\boldsymbol{U} \in T$, i.e., it is said that a random variable $\boldsymbol{Z} \in \mathbb{R}^p$ has a selection distribution if $Z \overset{\text{d}}{=} (\boldsymbol{V}|\boldsymbol{U} \in T)$, denoted by $\boldsymbol{Z} \sim SLCT_{p,q}(\boldsymbol{\theta})$ with parameter(s) $\boldsymbol{\theta}$ depending on the characteristics of $\boldsymbol{U}, \boldsymbol{V}$ and T with its joint distribution as

$$F_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{\int_T f_{\boldsymbol{U},\boldsymbol{V}}(\boldsymbol{u}, \boldsymbol{v})d\boldsymbol{u}}{\int_T f_{\boldsymbol{U}}(\boldsymbol{u})d\boldsymbol{u}}, \quad \boldsymbol{z} \in \mathbb{R}^p. \tag{4}$$

For more information about univariate and multivariate skew-normal distributions and their updates, we refer to Azzalini and Valle (1996), Teimouri (2021), and Wei, Conlon, and Wang (2021) as well as the books edited by Genton (2004) and Azzalini (2013).

The aim of this work is the study of selection distribution of copula related random vectors. Based on Sklar's theorem Sklar (1959), for a random vector $\boldsymbol{X} = (X_1, X_2, ..., X_d)$, the related multivariate cumulative distribution function $F_{\boldsymbol{X}}(\boldsymbol{x}) : \mathbb{R}^d \to [0, 1]$ can be expressed in terms of its marginals $F_{X_i}(x_i)$, $i = 1, 2, ..., d$, and some d-dimensional copula function C (shortly, $d-$copula) such that

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = C(F_{X_1}(x_1), F_{X_2}(x_2), ..., F_{X_d}(x_d)), \tag{5}$$

where $C : [0, 1]^d \to [0, 1]$ satisfies the following three properties:

1. grounded: $C(u_1, u_2, ..., u_d) = 0$ if, for some $k \leq d$, $u_k = 0$.
2. uniformly marginal: $C(1, ..., 1, u_k, 1, ..., 1) = u_k$ for any $k \leq d$ and all $u_k \in [0, 1]$.
3. $d-$increasing: For all $0 \leq u_{i1} \leq u_{i2} \leq 1$ and $i = 1, \ldots, d$

$$V_K \equiv \sum_{i_1=1}^{2} \ldots \sum_{i_d=1}^{2} (-1)^{i_1+\ldots+i_d} C\left(u_{1i_1}, \ldots, u_{di_d}\right) \geq 0.$$

We assume that all components of $\boldsymbol{X}$ are continuous random variables which guaranties that copula C is unique and hence its density, if exists, is just function $c(.)$ such that

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = c(F_{X_1}(x_1), F_{X_2}(x_2), ..., F_{X_d}(x_d))\Pi_{i=1}^{d} f_{X_i}(x_i), \tag{6}$$

where $f_{\boldsymbol{X}}(.) : \mathbb{R}^d \to \mathbb{R}^+$ is the joint density function of $\boldsymbol{X}$ and $f_{X_i} : \mathbb{R} \to \mathbb{R}^+,\ i = 1, 2, ..., d$ are respectively density functions of $X_i,\ i = 1, 2, ..., d$. For more details, refer to Durante and Sempi (2015); Joe (1997); Nelsen (2006); Mesiar, Sheikhi, and Komorníková (2019).

Denoting the copula coupling random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ as $C_{\boldsymbol{X},\boldsymbol{Y}}$, we study the distribution of $\boldsymbol{Z} \stackrel{\text{d}}{=} (\boldsymbol{X}|\boldsymbol{Y} \in T)$. We consider the *Gaussian* copula. Denoting Φ the standard normal cumulative distribution and $\boldsymbol{\Phi}_d(., \boldsymbol{R})$ the bivariate standard multivariate normal distribution function with correlation matrix $\boldsymbol{R}$, it is said that random vector $\boldsymbol{X}$ is associated to a *Gaussian* copula with correlation matrix $\boldsymbol{R}$ if

$$C^{Ga}(u_1, u_2, ..., u_d) = \boldsymbol{\Phi} d\big(\Phi^{-1}(u1), \Phi^{-1}(u_2), ..., \Phi^{-1}(u_d),\ \boldsymbol{R}\big) \tag{7}$$

where $\Phi^{-1}(.)$ is the inverse function of the standard normal distribution function and $\boldsymbol{R} \in [-1, 1]^{d\times d}$ is a correlation matrix $\boldsymbol{R}$ In the sequel, we assume that the correlation matrix $\boldsymbol{R}$ is invertible and all random variables are continues, so, the derivatives of their distribution functions are exist. We assume this copula as a connection function of our variables in this work. The rest of this paper is organized as follows. The main results are given in the next section. In Section 3, we apply our results to do a simulation analysis. Also an application of our theoretical results in a real dataset is given. Finally, some concluding remarks are presented in section 4.

2. The main results

Let $\boldsymbol{X} = (X_1, X_2, ..., X_d)$ and $\boldsymbol{Y} = (Y_1, Y_2, ..., Y_s)$ and the random vector $(\boldsymbol{X}, \boldsymbol{Y})$ is associated with a $(d + s)-$ copula function $C_{\boldsymbol{X},\boldsymbol{Y}}$, then from the classical probability, we have the distribution of $\boldsymbol{Z} \stackrel{\text{d}}{=} \boldsymbol{X}|\boldsymbol{Y} > \boldsymbol{\mu_y}$ as

$$\begin{aligned} F_{\boldsymbol{Z}}(\boldsymbol{z}) &= \frac{\oint_{-\infty}^{\boldsymbol{z}} \oint_{\boldsymbol{Ay}} f\boldsymbol{X}, \boldsymbol{Y}(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{y} d\boldsymbol{x}}{\oint_{\boldsymbol{Ay}} f\boldsymbol{Y}(\boldsymbol{y}) d\boldsymbol{y}} \\ &= m_y \oint_{-\infty}^{\boldsymbol{z}} \oint_{\boldsymbol{Ay}} c(FX_1(x_1), ..., F_{X_d}(x_d), F_{Y_1}(y_1), ..., F_{Y_s}(y_s))\Pi_{i=1}^{d} f_{X_i}(x_i)\Pi_{i=1}^{s} f_{Y_i}(y_i) d\boldsymbol{y} d\boldsymbol{x} \end{aligned}$$

where $\boldsymbol{Ay} = \{\boldsymbol{y} \ \in \mathbb{R}^s | \boldsymbol{y} > \boldsymbol{\mu y}\}$ and $\boldsymbol{\mu y}$ is the mean vector of $\boldsymbol{Y}$, $\oint$ denotes a corresponding multi-integral and $my = \oint_{\boldsymbol{Ay}} f\boldsymbol{Y}(\boldsymbol{y}) d\boldsymbol{y}$. Differentiating with respect to $\boldsymbol{z}$ we readily obtain the density of $\boldsymbol{Z}$ as

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = m_y \oint_{\boldsymbol{Ay}} c(FX_1(z_1), ..., F_{X_p}(z_p), F_{Y_1}(y_1), ..., F_{Y_q}(y_q))\Pi_{i=1}^{p} f_{X_i}(z_i)\Pi_{i=1}^{q} f_{Y_i}(y_i) d\boldsymbol{y} \tag{8}$$

where $c(.)$ is the density copula of $(\boldsymbol{X}, \boldsymbol{Y})$. The following theorem states a more specific version of (8) in which $s = 1$ and the selection is made on positive values of $Y - \mu_y$.

Theorem 1. *Assume that $C_{\boldsymbol{X},Y}$ is the copula function of random vector $\boldsymbol{X}$ and the random variable Y with marginal distribution functions F_{X_1}, F_{X_2}, ..., F_{X_d} and F_Y, then the distribution of $\boldsymbol{Z} \stackrel{\mathrm{d}}{=} \boldsymbol{X}|Y > \mu_y$ is given by*

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = m_y[f_{\boldsymbol{X}}(\boldsymbol{z}) - \Pi_{i=1}^{d} f_{X_i}(z_i) D_{12...d} C_{\boldsymbol{X},Y}], \qquad \boldsymbol{z} \in \mathbb{R}^d. \tag{9}$$

where $D_{12...d}C_{\boldsymbol{X},Y} = \frac{\partial^d C_{\boldsymbol{X},Y}}{\partial F_{X_1}(z_1)\partial F_{X_2}(z_2),...,F_{X_d}(z_d)}$ if exist, otherwise 0.

Proof. We have

$$\begin{aligned}
F_{\boldsymbol{Z}}(\boldsymbol{z}) &= m_y P(X_1 \le z_1, X_2 \le z_2, ..., X_d \le z_d,\ Y > \mu_y) \\
&= m_y[P(X_1 \le z_1, X_2 \le z_2, ..., X_d \le z_d) - P(X_1 \le z_1, X_2 \le z_2, ..., X_d \le z_d, Y \le \mu_y)] \\
&= m_y[F_{\boldsymbol{X}}(\boldsymbol{z}) - C_{\boldsymbol{X},Y}\big(F_{X_1}(z_1), F_{X_2}(z_2), ..., F_{X_d}(z_d), F_Y(\mu_y)\big)]
\end{aligned}$$

using some chain rules we readily obtain

$$\begin{aligned}
f_{\boldsymbol{Z}}(\boldsymbol{z}) &= m_y\big[f_{\boldsymbol{X}}(\boldsymbol{z}) - \frac{\partial^d C_{\boldsymbol{X},Y}}{\partial F_{X_1}(z_1)\partial F_{X_2}(z_2)...\partial F_{X_d}(z_d)} \frac{\partial F_{X_1}(z_1)}{\partial z_1}\frac{\partial F_{X_2}(z_2)}{\partial z_2}\frac{\partial F_{X_d}(z_d)}{\partial z_d}\big] \\
&= m_y[f_{\boldsymbol{X}}(\boldsymbol{z}) - \Pi_{i=1}^{d} f_{X_i}(z_i) D_{12...d} C_{\boldsymbol{X},Y}]
\end{aligned}$$

which proves the assertion. □

Corollary 1. *Let $\boldsymbol{X}$ be a vector of random variables that have multivariate standard normal and Y has a standard normal so that they are connected via the Gaussian copula with correlation matrix $\boldsymbol{R}$, then the distribution of $\boldsymbol{Z} \stackrel{\mathrm{d}}{=} \boldsymbol{X}|Y > \mu_y$ is*

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = 2\varphi(\boldsymbol{z}, \boldsymbol{\Omega})\Phi(\frac{\alpha^T \boldsymbol{z}}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}}), \qquad \boldsymbol{z} \in \mathbb{R}^d, \tag{10}$$

where $\boldsymbol{\Omega}$ is the correlation matrix of $\boldsymbol{X}$ and $\alpha^T = (-|\boldsymbol{R}|a_{12} \ .... - |\boldsymbol{R}|a_{1d})$ where a_{1i}, $i = 2, 3, ..., d$, are the elements of the first row of matrix $\boldsymbol{R^{-1}}$.

Proof. Since $m_y = 2$, similar to the proof of Theorem 1 we have

$$F_{\boldsymbol{Z}}(\boldsymbol{z}) = 2[\Phi(\boldsymbol{z}) - C_{\boldsymbol{X},Y}(\Phi(z_1), ..., \Phi(z_d), \Phi(0))].$$

and again using some chain rules we have

$$\begin{aligned}
f_{\boldsymbol{Z}}(\boldsymbol{z}) &= 2\big[\varphi(\boldsymbol{z}, \boldsymbol{\Omega}) - \Pi_{i=1}^{d}\varphi_{X_i}(z_i) D_{12...d} C^{Ga}_{\boldsymbol{X},y}\big] \\
&= 2\big[\varphi(\boldsymbol{z}, \boldsymbol{\Omega}) - \Pi_{i=1}^{d}\varphi_{X_i}(z_i)\frac{\varphi(\boldsymbol{z}, \boldsymbol{\Omega})}{\Pi_{i=1}^{d}\varphi_{X_i}(z_i)}\Phi(-\frac{\alpha^T \boldsymbol{z}}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}})\big] \\
&= 2\varphi(\boldsymbol{z}, \boldsymbol{\Omega})\Phi(\frac{\alpha^T \boldsymbol{z}}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}}), \qquad \boldsymbol{z} \in \mathbb{R}^d,
\end{aligned}$$

which is 10. □

Remark 1. *Putting $d = 2$ in (10), after some algebras, the density function $(Z_1, Z_2)^T \stackrel{\mathrm{d}}{=} (X_1, X_2)^T|Y > 0$ will be*

$$f_{Z_1,Z_2}(z_1, z_2) = 2\varphi(z_1, z_2, \boldsymbol{\Omega})\Phi(\alpha_1 z_1 + \alpha_2 z_2), \quad z_1,\ z_2 \in \mathbb{R}, \tag{11}$$

where

$$\begin{aligned}
\alpha_1 &= \frac{\rho_1 - \rho_{12}\rho_2}{\sqrt{(1-\rho_{12}^2)(1-\rho_{12}^2-\rho_1^2-\rho_2^2+2\rho_{12}\rho_1\rho_2)}}, \\
\alpha_2 &= \frac{\rho_2 - \rho_{12}\rho_1}{\sqrt{(1-\rho_{12}^2)(1-\rho_{12}^2-\rho_1^2-\rho_2^2+2\rho_{12}\rho_1\rho_2)}}
\end{aligned}$$

with $\rho_{12} = \rho_{X_1,X_2}$ *and* $\rho_1 = \rho_{X_1,Y}$*,* $\rho_2 = \rho_{X_2,Y}$ *and* $\boldsymbol{\Omega} = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}$*, which has been presented by Azzalini and Valle (1996).*

We can generalize or simplify this distribution by considering any related copula or marginals. For example, if the marginal distributions of the random variable $\boldsymbol{X}$ are standard normal, while Y follows an arbitrary distribution, say $F_Y(y)$, with mean μ_y, then we may present the following results.

Corollary 2. *let* $\boldsymbol{X}$ *be a vector of random variables that have multivariate standard normal. If* Y *has a distribution function* $F_Y(y)$ *and* $\boldsymbol{X}$ *and* Y *are associated with a Gaussian copula with correlation matrix* $\boldsymbol{R}$*, then the distribution of* $\boldsymbol{Z} \stackrel{\mathrm{d}}{=} \boldsymbol{X}|Y > \mu_y$ *is*

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = m_y \varphi(\boldsymbol{z}, \boldsymbol{\Omega}) \Phi(\frac{\alpha^T \boldsymbol{z}}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}} - \frac{\Phi^{-1}(F_Y(\mu_y))\sqrt{|\boldsymbol{\Omega}|}}{\sqrt{|\boldsymbol{R}|}}), \quad \boldsymbol{z} \in \mathbb{R}^d \tag{12}$$

Proof. The proof is similar to the proof of Corollary 1 and is omitted. □

Corollary 3. *let* X_1 *and* X_2 *have a normal bivariate distribution with mean* $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ *and correlation matrix* $\boldsymbol{\Omega} = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}$*. If* Y *has a distribution function* $F_Y(y)$ *and* $\boldsymbol{X}$ *and* Y *are associated with a Gaussian copula with correlation matrix* $\boldsymbol{R}$*, then the distribution of* $\boldsymbol{Z} \stackrel{\mathrm{d}}{=} \boldsymbol{X}|Y > \mu_y$ *is*

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = m_y \varphi(\boldsymbol{z}, \boldsymbol{\Omega}) \Phi(\frac{(\rho_1 - \rho_{12}\rho_2)z_1 + (\rho_2 - \rho_{12}\rho_1)z_2}{\sqrt{(1-\rho_{12}^2)(1-\rho_{12}^2-\rho_1^2-\rho_2^2+2\rho_{12}\rho_1\rho_2)}} - \frac{\Phi^{-1}(F_Y(\mu_y))\sqrt{1-\rho_{12}^2}}{\sqrt{1-\rho_{12}^2-\rho_1^2-\rho_2^2+2\rho_{12}\rho_1\rho_2}}) \tag{13}$$

Proof. The proof is similar to the proof of Corollary 1 and is omitted. □

Corollary 4. *Under the assumptions of corollary 2 the moment generating function (mgf) of the random variable* $\boldsymbol{Z}$ *is*

$$M_{\boldsymbol{Z}}(\boldsymbol{t}) = m_y \; exp(\frac{\boldsymbol{t}^T \boldsymbol{\Omega} \boldsymbol{t}}{2}) \; \Phi(\frac{\alpha^T \boldsymbol{\Omega} \boldsymbol{t} - \Phi^{-1}(F_Y(\mu_y))|\boldsymbol{\Omega}|}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}| + \alpha^T \boldsymbol{\Omega} \alpha}}). \tag{14}$$

Proof. From the definition of the mgf function we have

$$\begin{aligned}
M_{\boldsymbol{Z}}(\boldsymbol{t}) &= \int_{\mathbb{R}^k} e^{\boldsymbol{t}^T \boldsymbol{z}} f_{\boldsymbol{Z}}(\boldsymbol{z}) d\boldsymbol{z} \\
&= \frac{m_y}{(2\pi)^{\frac{k}{2}} |\boldsymbol{\Omega}|^{\frac{1}{2}}} \int_{\mathbb{R}^k} exp(\frac{-1}{2}(\boldsymbol{z}^T |\boldsymbol{\Omega}|^{-1} \boldsymbol{z} - 2\boldsymbol{t}^T \boldsymbol{z}) \Phi(\frac{\alpha^T \boldsymbol{z}}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}} - \frac{\Phi^{-1}(F_Y(\mu_y))\sqrt{|\boldsymbol{\Omega}|}}{\sqrt{|\boldsymbol{R}|}}) d\boldsymbol{z} \\
&= m_y \; exp(\frac{\boldsymbol{t}^T \boldsymbol{\Omega} \boldsymbol{t}}{2}) \; \Phi(\frac{\alpha^T \boldsymbol{\Omega} \boldsymbol{t} - \Phi^{-1}(F_Y(\mu_y))|\boldsymbol{\Omega}|}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}| + \alpha^T \boldsymbol{\Omega} \alpha}}). \\
&= m_y \; exp(\frac{\boldsymbol{t}^T \boldsymbol{\Omega} \boldsymbol{t}}{2}) \; \Phi(\rho_1 t_1 + \rho_2 t_2 + \cdots + \rho_d t_d - \frac{\Phi^{-1}(F_Y(\mu_y))|\boldsymbol{\Omega}|}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}| + \alpha^T \boldsymbol{\Omega} \alpha}}).
\end{aligned}$$

where $\rho_i = \rho_{X_i,Y}$, $i = 1, 2, ..., d$, which is 14. □

As a quick result of the previous corollary, derivation of mgf and getting the torques, we have can obtain mean, variance and covariances of these random variables, respectively as

$$\mu_i = m_y \rho_i \; \phi(-h_{\mu_y}), \quad i = 1, 2, ..., d$$

$$\sigma_i^2 = m_y \Phi\big(-h_{\mu_y}\big) + m_y \rho_i^2 h_{\mu_y} \phi(-h_{\mu_y}) - (m_y \rho_i \; \phi(-h_{\mu_y}))^2, \quad i = 1, 2, ..., d$$

$$\sigma_{ij} = m_y\rho_{ij}\Phi(-h_{\mu_y}) + m_y\rho_i\rho_j h_{\mu_y}\phi(-h_{\mu_y}) - m_y^2\ \rho_i\rho_j\ \phi(-h_{\mu_y})^2, \quad i \neq j = 1,2,...,d$$

where $\rho_{X_i,Y} = \rho_i,\ \rho_{X_j,Y} = \rho_j, ..., \rho_{X_i,X_j} = \rho_{ij} \quad i,j = 1,2,...,d.$

The following remark brings a special case of the above results when the components of random vector $\boldsymbol{X}$ are independent and generalizes the results of Arnold *et al.* (2002) and those obtained by Azzalini and Valle (1996).

Remark 2. *Let $\boldsymbol{X}$ be a standard normal random vector and let Y has a distribution function $F_Y(y)$ and they are associated with the Gaussian copula with correlation matrix $\boldsymbol{R}$ with $\rho_{X_i,X_j} = 0, \quad i \neq j = 1,2,...,d$, then the distribution of $\boldsymbol{Z} \stackrel{d}{=} \boldsymbol{X}|Y > \mu_y$ is*

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = m_y\Pi_{i=1}^d\varphi(z_i)\Phi\big(\frac{\Sigma_{i=1}^d\rho izi - \Phi^{-1}(F_Y(\mu_y))}{\sqrt{1-\Sigma_{i=1}^d\rho_i^2}}\big), \quad \boldsymbol{z} \in \mathbb{R}^d. \tag{15}$$

Also, denoting with $\lambda_i = \dfrac{\rho_i}{\sqrt{1-\Sigma_{i=1}^d\rho_i^2}}, \quad i = 1,...,d$, the distribution (15) simplifies to

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = m_y\Pi_{i=1}^d\varphi(z_i)\Phi\big(\Sigma_{i=1}^d\lambda_i\Phi^{-1}(F_{X_i}(x_i)) - \sqrt{1+\Sigma_{i=1}^d\lambda_i^2}\Phi^{-1}(F_Y(\mu_y))\big), \quad \boldsymbol{z} \in \mathbb{R}^d,$$

if Y follows the normal distribution as well.

It is known that, besides of the aforementioned "conditional method", there is another "transformation method" that can be employed to construct an asymmetric distribution, see e.g., Azzalini and Capitanio (2003). In this regards, under the assumptions of corollary 2 and if $\rho_i = \rho_{X_i,Y} \in (-1,1),\ i = 1,2,...,d$ we may define

$$Z_i = \rho_i H(Y) + \sqrt{(1-\rho_i^2)}X_i, \quad i = 1,2,3,...,d, \tag{16}$$

where $H(.)$ is a measurable function, then the distribution function $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_d)^T$ will be

$$F_{\boldsymbol{Z}}(\boldsymbol{z}) = \int D_{d+1}C^{Ga}(u_1,u_2,...,u_d,v)\ dF_y(y), \tag{17}$$

where $D_{d+1}C^{Ga}(u_1,u_2,...,u_k,v) = \dfrac{\partial}{\partial v}C^{Ga}(u_1,u_2,...,u_d,v)$, and $u_i = \Phi(\dfrac{z_i - \rho_i h(y)}{\sqrt{(1-\rho_i^2)}}), i = 1,2,...,d$ and hence, the density function $f_{\boldsymbol{Z}}(\boldsymbol{z})$ will be obtained accordingly.

As a special case, If Y has a standard normal distribution and $H(.) = |.|$, we will reach to the results of Azzalini and Dalla Valle Azzalini and Valle (1996). In this case, the joint PDF of $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_d)$ is as

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = 2\varphi_d(\boldsymbol{z}, \boldsymbol{\Omega})\Phi(\alpha^T\boldsymbol{z}), \quad \boldsymbol{z} \in \mathbb{R}^d, \tag{18}$$

where

$$\alpha^T = \frac{\lambda^T\Psi^{-1}\Delta^{-1}}{(1+\lambda^T\Psi^{-1}\lambda)^{\frac{1}{2}}},$$
$$\Delta = diag((1-\rho_1^2)^{\frac{1}{2}}, ..., (1-\rho_d^2)^{\frac{1}{2}}),$$
$$\lambda = (\tfrac{\rho_1}{\sqrt{(1-\rho_1^2)}}, ..., \tfrac{\rho_d}{\sqrt{(1-\rho_d^2)}})^T,$$
$$\boldsymbol{\Omega} = \Delta(\Psi + \lambda\lambda^T)\Delta,$$

where $\varphi_d(\boldsymbol{z}, \boldsymbol{\Omega})$ denotes the density function of the d-dimensional multivariate normal distribution with standardized marginals and correlation matrix $\boldsymbol{\Omega}$.

3. Parameter estimation

In order to estimate the parameters of the proposed distribution, we apply the maximum likelihood (ML) method. First, focusing on the distribution (15) we find ML estimation of its parameters. Taking in mind that the $\Phi^{-1}(F_Y(\mu_y))$ part in this distribution may include parameter(s), in order to find a closed form of estimators we assume that the distribution of Y dose not contain any parameters and so, $\Phi^{-1}(F_Y(\mu_y))$ will be constant. Let $\boldsymbol{z_j} = (z_{1j}, z_{2j}, \cdots, z_{dj})^T$, $j = 1, 2, ..., n$, and $\boldsymbol{z}1, \boldsymbol{z}2, ..., \boldsymbol{z}n$ be a sample of size n from distribution (15) in which $\Phi^{-1}(FY(\mu_y))$ is free of parameter, then likelihood function is

$$\begin{aligned} L_n(\lambda_1, \lambda_2, \cdots, \lambda_d) &= f_n(\boldsymbol{z_1}, ..., \boldsymbol{z_n}; \lambda_1, \lambda_2, \cdots, \lambda_d) \\ &= m_y^n \Pi_{j=1}^n \Pi_{i=1}^d \varphi(z_{ij}) \Pi_{j=1}^n \Phi(\Sigma_{i=1}^d \lambda_i z_{ij} - \sqrt{1 + \Sigma_{i=1}^d \lambda_i^2} \Phi^{-1}(F_Y(\mu_y))) \end{aligned}$$

and its log-likelihood will be

$$\ell(\lambda_1, \lambda_2; \boldsymbol{z_1}, \boldsymbol{z_2}) = nlnm_y + \Sigma_{j=1}^n \Sigma_{i=1}^d ln\varphi(z_{ij}) + \Sigma_{j=1}^n ln\Phi(\Sigma_{i=1}^d \lambda_i z_{ij} - \sqrt{1 + \Sigma_{i=1}^d \lambda_i^2} \Phi^{-1}(F_Y(\mu_y)))$$

Then, derivatives with respect to λ_1 and $\lambda_2, \cdots, \lambda_d$ yield the following simultaneous equations:

$$\begin{cases} z_{1j} = \frac{\lambda_1 \Phi^{-1}(F_y(\mu_y))}{\sqrt{1+\Sigma_{i=1}^d \lambda_i^2}} \quad or \quad \phi(\Sigma_{i=1}^d \lambda_i z_{ij} - \sqrt{1 + \Sigma_{i=1}^d \lambda_i^2} \Phi^{-1}(F_Y(\mu_y))) = 0, \quad j = 1, ..., n \\ z_{2j} = \frac{\lambda_2 \Phi^{-1}(F_y(\mu_y))}{\sqrt{1+\Sigma_{i=1}^d \lambda_i^2}} \quad or \quad \phi(\Sigma_{i=1}^d \lambda_i z_{ij} - \sqrt{1 + \Sigma_{i=1}^d \lambda_i^2} \Phi^{-1}(F_Y(\mu_y))) = 0, \quad j = 1, ..., n \\ \vdots \\ z_{dj} = \frac{\lambda_d \Phi^{-1}(F_y(\mu_y))}{\sqrt{1+\Sigma_{i=1}^d \lambda_i^2}} \quad or \quad \phi(\Sigma_{i=1}^d \lambda_i z_{ij} - \sqrt{1 + \Sigma_{i=1}^d \lambda_i^2} \Phi^{-1}(F_Y(\mu_y))) = 0, \quad j = 1, ..., n \end{cases}.$$

So, from the first equation we have

$$\hat{\lambda}1 = \frac{z1j}{\sqrt{\Phi^{-1}(F_y(\mu_y))^2 - \Sigma_{i=1}^d z_{ij}^2}}, \quad j = 1, 2, ..., n, \quad or \quad \hat{\lambda}_1 = \pm\infty.$$

In order to assess the local concavity of $\ell(\lambda_1, \lambda_2, \cdots, \lambda_d; \boldsymbol{z_1}, \boldsymbol{z_2}, \cdots, \boldsymbol{z_n})$ we observe that it's Hessian matrix is negative semi-definite at $\hat{\lambda}1$ and $\hat{\lambda}2, \cdots, \hat{\lambda}_d$. Hence, we have

$$\hat{\rho}1 = \frac{z1j}{\Phi^{-1}(F_y(\mu_y))}, \quad j = 1, 2, ..., n \; or \; \hat{\rho}_1 = \pm 1$$

but, $\hat{\rho}1 = \pm 1$ is unacceptable because $\rho 1 \; \in (-1, 1)$. So

$$\hat{\rho}1 = \frac{z1j}{\Phi^{-1}(F_y(\mu_y))}, \quad j = 1, 2, ..., n.$$

Using a similar manner it can be obtained:

$$\hat{\rho}2 = \frac{z2j}{\Phi^{-1}(F_y(\mu_y))}, \cdots, \hat{\rho}d = \frac{zdj}{\Phi^{-1}(F_y(\mu_y))}, \quad j = 1, 2, ..., n.$$

As another and a general case, consider $\boldsymbol{z_j} = (z_{1j}, z_{2j}, \cdots, z_{dj})^T$, $j = 1, 2, ..., n$, let $\boldsymbol{z}1, \boldsymbol{z}2, ..., \boldsymbol{z}n$ be a sample of size n from distribution (12) in which $\Phi^{-1}(FY(\mu_y))$ is free of parameter, then likelihood function is

$$\begin{aligned} L_n(\rho_i, \rho_{kl}) &= f_n(\boldsymbol{z_1}, ..., \boldsymbol{z_n}; \rho_i, \rho_{kl}) \\ &= m_y^n \Pi_{j=1}^n \varphi(\boldsymbol{z}j, \boldsymbol{\Omega}) \Pi j = 1^n \Phi(\frac{\alpha^T \boldsymbol{z}j}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}} - \frac{\Phi^{-1}(FY(\mu_y))\sqrt{|\boldsymbol{\Omega}|}}{\sqrt{|\boldsymbol{R}|}}) \end{aligned}$$

So that $i = 1, 2, \cdots, d$ and $k = 1, 2, \cdots, d-1$, $l = 2, \cdots, d$ and its log-likelihood will be

$$\ell(\rho_i, \rho_{kl}) = nlnm_y + \Sigma_{j=1}^{n} ln\varphi(\boldsymbol{z}j, \boldsymbol{\Omega}) + \Sigma j = 1^n ln\Phi(\frac{\alpha^T \boldsymbol{z}j}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}} - \frac{\Phi^{-1}(FY(\mu_y))\sqrt{|\boldsymbol{\Omega}|}}{\sqrt{|\boldsymbol{R}|}}).$$

By deriving with respect to ρ_i, $i = 1, 2, ..., d$ we have

$$\frac{\partial}{\partial \rho_i}\ell(\rho_i, \rho_{kl}) = \frac{\alpha_i' \boldsymbol{z}j\sqrt{|\boldsymbol{R}||\boldsymbol{\Omega}|} - Mi'\alpha^{\boldsymbol{T}}\boldsymbol{z}j + ri'|\boldsymbol{\Omega}|^{\frac{3}{2}}\Phi^{-1}(F_Y(\mu_y))}{|\boldsymbol{R}||\boldsymbol{\Omega}|} \times \frac{\phi(\frac{\alpha^T \boldsymbol{z}j}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}} - \frac{\Phi^{-1}(FY(\mu_y))\sqrt{|\boldsymbol{\Omega}|}}{\sqrt{|\boldsymbol{R}|}})}{\Phi(\frac{\alpha^T \boldsymbol{z}j}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}} - \frac{\Phi^{-1}(FY(\mu_y))\sqrt{|\boldsymbol{\Omega}|}}{\sqrt{|\boldsymbol{R}|}})},$$

where $\alpha_i' = \frac{\partial}{\partial \rho_i}\alpha^{\boldsymbol{T}}$, $M_i' = \frac{\partial}{\partial \rho_i}\sqrt{|\boldsymbol{R}||\boldsymbol{\Omega}|}$ and $r_i' = \frac{\partial}{\partial \rho_i}\sqrt{|\boldsymbol{R}|}$.
and the derivation with respect to ρ_{kl}, $k = 1, 2, \cdots, d-1$, $l = 2, \cdots, d$ yields:

$$\begin{aligned}\frac{\partial}{\partial \rho_{kl}}\ell(\rho_i, \rho_{kl}) &= \frac{\varphi'(\boldsymbol{z}j, \boldsymbol{\Omega})}{\varphi(\boldsymbol{z}j, \boldsymbol{\Omega})} \\ &+ \frac{\alpha_{kl}' \boldsymbol{z}j\sqrt{|\boldsymbol{R}||\boldsymbol{\Omega}|} - Mkl'\alpha^{\boldsymbol{T}}\boldsymbol{z}j - (\sqrt{|\boldsymbol{\Omega}|})'\sqrt{|\boldsymbol{R}|}|\boldsymbol{\Omega}|\Phi^{-1}(FY(\mu_y)) + r_{kl}'|\boldsymbol{\Omega}|^{\frac{3}{2}}\Phi^{-1}(F_Y(\mu_y))}{|\boldsymbol{R}||\boldsymbol{\Omega}|} \\ &\times \frac{\phi(\frac{\alpha^T \boldsymbol{z}j}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}} - \frac{\Phi^{-1}(FY(\mu_y))\sqrt{|\boldsymbol{\Omega}|}}{\sqrt{|\boldsymbol{R}|}})}{\Phi(\frac{\alpha^T \boldsymbol{z}j}{\sqrt{|\boldsymbol{\Omega}||\boldsymbol{R}|}} - \frac{\Phi^{-1}(FY(\mu_y))\sqrt{|\boldsymbol{\Omega}|}}{\sqrt{|\boldsymbol{R}|}})},\end{aligned}$$

where $\varphi'(\boldsymbol{z}j, \boldsymbol{\Omega}) = \frac{\partial}{\partial \rho kl}\varphi(\boldsymbol{z}j, \boldsymbol{\Omega})$ and $\alpha kl' = \frac{\partial}{\partial \rho_{kl}}\alpha^{\boldsymbol{T}}$ and $M_{kl}' = \frac{\partial}{\partial \rho_{kl}}\sqrt{|\boldsymbol{R}||\boldsymbol{\Omega}|}$ and $r_{kl}' = \frac{\partial}{\partial \rho_{kl}}\sqrt{|\boldsymbol{R}|}$. Hence, the numerical solution of the simultaneous equations $\frac{\partial}{\partial \rho_i}\ell(\rho_i, \rho_{kl}) = 0$ and $\frac{\partial}{\partial \rho_{kl}}\ell(\rho_i, \rho_{kl}) = 0$ will give the estimated values of the parameters, see the next section for illustration.

4. Numerical analysis

In order to assess the performance of our proposed distribution and compare its goodness of fit with the traditional skew-normal distribution, we used a simulation study.

4.1. Simulation study

Applying a Monte Carlo simulation study, we consider the following three cases.

Case 1. We generated $n = 1000$ sample from a four dimensional distribution of random variables X_1, X_2, X_3 and Y which are connected using a *Gaussian* copula with $\rho_{X_1,Y} = 0.5$, $\rho_{X_2,Y} = 0.6$, $\rho_{X_3,Y} = 0.7$ and $\rho_{X_i,X_j} = 0, 1 \leq i \leq j \leq 3$. Also, we assumed that X_1, X_2, X_3 come from trivariate standard normal distribution with mean $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and correlation matrix $\boldsymbol{\Omega} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and Y follows an exponential distribution with $\lambda = 10$. Table 1 summarizes AIC and BIC and HQC values for the goodness of fit for our proposed copula-normal distribution and skew-normal one. As seen from this table, the copula-skew distribution behave superior to skew-normal.

Table 1: AIC, BIC and HQC of copula-skew and skew-normal for Case 1. of simulation

Estimation	skew-normal	copula-skew
AIC	1148	1101
BIC	1192	1145
HQC	1162	1116

Case 2. We generated $n = 1000$ sample from a three dimensional distribution of random variables X_1, X_2 and Y which are connected using a *Gaussian* copula with $\rho_{X_1,Y} = 0.8$, $\rho_{X_2,Y} = 0.5$, $\rho_{X_1,X_2} = 0.6$. Also, we assumed that X_1, X_2, a normal bivariate distribution with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and correlation matrix $\boldsymbol{\Omega} = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}$ and Y follows an chi-squared distribution with 5 degrees of freedom. Table 2 summarizes AIC, BIC and HQC values for the goodness of fit for our proposed copula-normal distribution and skew-normal one. As seen from this table, the copula-skew distribution behave superior to the skew-normal.

Table 2: AIC, BIC and HQC of copula-skew and skew-normal for Case 2. of simulation

Estimation	skew-normal	copula-skew
AIC	1920	1917
BIC	1947	1945
HQC	1931	1928

Case 3. We generated 1000 random samples for X_1, X_2 similar to the case 1 and using $Y = 0.1X_1 + 0.3X_2 + \gamma$ where $\gamma \sim Beta(2, 3)$, we generated 1000 cases for Y. We found that they are related with the correlation matrix $\boldsymbol{R} = \begin{pmatrix} 1 & 0.6 & 0.84 \\ 0.6 & 1 & 0.5 \\ 0.84 & 0.5 & 1 \end{pmatrix}$ through the Gaussian copula. Table 3 summarizes AIC, BIC and HQC values for the goodness of fit for our proposed copula-normal distribution and skew-normal one. As seen from this table, the copula-skew distribution behaves superior to skew-normal.

Table 3: AIC, BIC and HQC of copula-skew and skew-normal for Case 3 of simulation

Estimation	skew-normal	copula-skew
AIC	2287	2285
BIC	2315	2313
HQC	2398	2397

4.2. Real data

Regarding to apply our previous material in a real data set analysis, we considered the well known data set collected by Australian Institute of sports (AIC)Cook and Weisberg (2009). This date set is compressing of biological characteristics of 202 Australian athletes in both sexes (102 males and 100 females) and was studied as an example of skew-normal data in the literature Azzalini and Valle (1996); Azzalini and Capitanio (1999). We examined the variables height (Ht) and body weight (Wt) and hematocrit (HCT) and observed that Ht and Wt followed normal distribution with mean 180 and variance 95 and mean 75 and variance 194, respectively. HCT followed exponential distribution with mean 43.
Moreover we found that these three variables were connected via a *Gaussian*-copula with values of correlations $\rho_{1y} = 0.37$ and $\rho_{2y} = 0.42$ and $\rho_{12} = 0.78$. By conditioning the

distribution of $\begin{pmatrix} Ht \\ Wt \end{pmatrix}$ given HCT greater than 43, from (15), the distribution of

$$\boldsymbol{Z} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \stackrel{\text{d}}{=} \begin{pmatrix} Ht \\ Wt \end{pmatrix} |HCT > 43$$

was fitted as

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{exp(1)}{P(y > 43)} \varphi(z_1, z_2, \boldsymbol{\Omega}) \Phi\big(\frac{0.042 z_1 + 0.131 z_2}{\sqrt{0.12}} - \frac{\Phi^{-1}(F_Y(43))\sqrt{0.39}}{\sqrt{0.32}}\big), \quad \mathbf{z} \in \mathbb{R}.$$

where $\boldsymbol{\Omega} = \begin{pmatrix} 1 & 0.78 \\ 0.78 & 1 \end{pmatrix}$.

In order to compare the goodness of fit of this distribution, we compare its AIC, BIC and Hannan-Quinn criterion (HQC) with those of skew-normal's one. A graphical representation of these criterions for these two distributions is presented in figure 1 and their AIC, BIC and HQC are calculated in table 4. Comparing the AIC, BIC and HQC values of this table and seeing figure 1, we conclude that implementing the dependence between variables Ht and Wt and HCT and constructing the copula-skew distribution yields an improvement to the fitting the distribution of $\boldsymbol{Z} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \stackrel{\text{d}}{=} \begin{pmatrix} Ht \\ Wt \end{pmatrix} |HCT > 43$, in contrast to the traditional skew-normal distribution.

Table 4: AIC, BIC and HQC of copula-skew and skew-normal for $(Ht, Wt)|HCT > 43$

Estimation	skew-normal	copula-skew
AIC	490	489
BIC	509	508
HQC	498	497

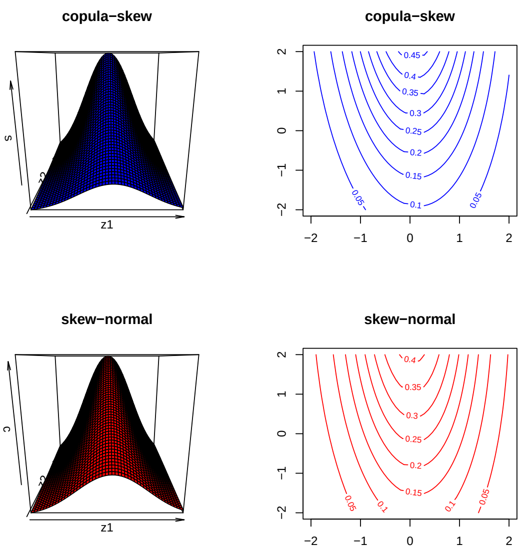

Figure 1: Performance of skew-normal and copula-skew $(Ht, Wt)|HCT > 43$

Moreover, we examined the variables height (Ht) and white blood cell (WBC) and Protocol-buffer Binary Format (PBF) observed that they are associated to a *Gaussian*-copula with values of correlations $\rho_{1y} = 0.11$ and $\rho_{2y} = -0.19$ and $\rho_{12} = 0$. Ht followed normal distribution with mean 180 and variance 95 and WBC followed a normal distribution with mean

7.1 and variance 3.2, respectively. Also, PBF followed generalized Pareto distributions with location parameter 5.38 and scale parameter 11.01 and shape parameter -0.35. By calculating the mean of PBF and conditioning the distribution of $\begin{pmatrix} Ht \\ WBC \end{pmatrix}$ given PBF greater than $\mu_{PBF} = 14$, from (15), the distribution of

$$\boldsymbol{Z} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \stackrel{\mathrm{d}}{=} \begin{pmatrix} Ht \\ WBC \end{pmatrix} |PBF > 14$$

was fitted as

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{P(y > 14)} \varphi(z_1)\varphi(z_2)\Phi\Big(\frac{0.11z_1 - 0.19z_2 - \Phi^{-1}(F_Y(14))}{\sqrt{0.951}}\Big), \quad \mathbf{z} \in \mathbb{R}.$$

A graphical representation of these two distributions is presented in figure 1 and their AIC and BIC and HQC are calculated in table 5. Comparing the AIC and BIC and HQC values of the table 5, we conclude that implementing the correlation between two variables Ht and WBC and PBF yields an improvement to the fitting the distribution of $\boldsymbol{Z} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \stackrel{\mathrm{d}}{=} \begin{pmatrix} Ht \\ WBC \end{pmatrix} |PBF > 14$, in contrast to the traditional skew-normal distribution. See also figures 1 and 2 for more details.

Table 5: AIC, BIC and HQC of copula-skew and skew-normal for $(Ht, WBC)|PBF > 14$

Estimation	skew-normal	copula-skew
AIC	439	439
BIC	458	457
HQC	447	446

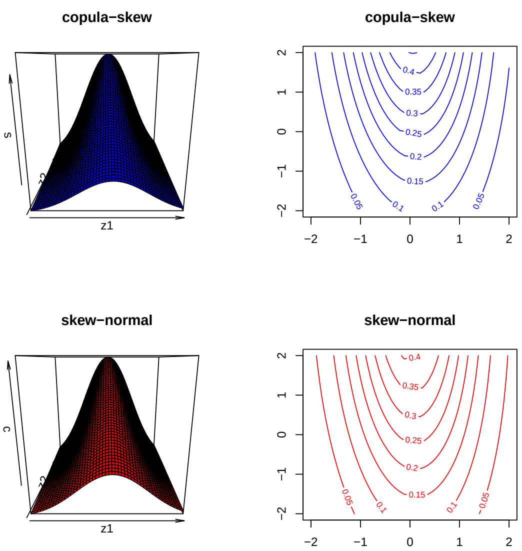

Figure 2: Performance of skew-normal and copula-skew for $(Ht, WBC)|PBF > 14$

5. Conclusion

Assuming that random vectors are connected using a *Gaussian* copula, we used a selection method to construct a copula based asymmetric distribution. Our proposed distribution includes some well-known multivariate and bivariate skew normal distributions which have been introduced in the literature. We carried out a simulation analysis and observed that taking into account the relationship between variables will improve goodness of fit of constructed distribution. Also, we have applied this distribution to analyze a real data sets of athletic sport data.

The subject of this work may be extended in many scenarios. Although we have considered the *Gaussian* copula as coupling function, one may assume some other type of copulas. As a generalization of this work, we address the family of elliptical copulas in our next work. Since the copula functions have flexibility to reflect the non-linear relationships between variables, so constructing an asymmetric distribution from nonlinear dependent variable would be of interest. Moreover, the approach of this work can be extended for modeling high-dimensional asymmetric dependence patterns, see, e.g., Wei, Kim, Choi, and Kim (2019); Kurowicka and Cooke (2006). Finally, it is well known that the asymetric distributions are frequently used as an approach to model the behavior of order statistics, see e.g., Loperfido (2008) and Sheikhi and Tata (2013). Our copula skew distribution will be used in this subject as well as in concomitants of order statistics.

Acknowledgment

The work of the third author was supported by the grant APVV-18-0052.

References

Aigner D, Lovell CAK, Schmidt P (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics*, **6**(1), 21–377. doi:10.1016/0304-4076(77)90052-5.

Anděl J, Netuka I, Zvára K (1984). "On Threshold Autoregressive Processes." *Kybernetika*, **20**(2), 89–106.

Arellano-Valle RB, Branco MD, Genton MG (2006). "A Unified View on Skewed Distributions Arising from Selections." *Canadian Journal of Statistics*, **34**(4), 581–601.

Arnold BC, Beaver RJ, Azzalini A, Balakrishnan N, Bhaumik A, Dey DK, Cuadras CM, Sarabia JM (2002). "Skewed Multivariate Models Related to Hidden Truncation and/or Selective Reporting." *Test*, **11**(1), 7–54.

Azzalini A (1985). "A Class of Distributions which Includes the Normal Ones." *Scandinavian Journal of Statistics*, pp. 171–178.

Azzalini A (2013). *The Skew-normal and Related Families*, volume 3. Cambridge University Press.

Azzalini A, Capitanio A (1999). "Statistical Applications of the Multivariate Skew-normal Distribution." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **61**(3), 579–602.

Azzalini A, Capitanio A (2003). "Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t-Distribution." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **65**(2), 367–389.

Azzalini A, Valle AD (1996). "The Multivariate Skew-normal Distribution." *Biometrika*, **83**(4), 715–726.

Cook RD, Weisberg S (2009). *An Introduction to Regression Graphics*, volume 405. John Wiley & Sons.

Durante F, Sempi C (2015). *Principles of Copula Theory*. CRC press.

Genton MG (2004). *Skew-elliptical Distributions and Their Applications: A Journey beyond Normality*. CRC Press.

Joe H (1997). *Multivariate Models and Multivariate Dependence Concepts*. CRC Press.

Kurowicka D, Cooke RM (2006). *Uncertainty Analysis with High Dimensional Dependence Modeling*. John Wiley & Sons.

Loperfido N (2008). "Modeling Maxima of Longitudinal Contralateral Observations." *Test*, **17**(2), 370–380.

Mesiar R, Sheikhi A, Komorníková M (2019). "Random Noise and Perturbation of Copulas." *Kybernetika*, **55**(2), 422–434.

Nelsen RB (2006). *An Introduction to Copulas*. Springer Science & Business Media.

Roberts C (1966). "A Correlation Model Useful in the Study of Twins." *Journal of the American Statistical Association*, **61**(316), 1184–1190.

Sheikhi A, Tata M (2013). "The Exact Joint Distribution of Concomitants of Order Statistics and Their Order Statistics under Normality." *REVSTAT–Statistical Journal*, **11**(2), 121–134.

Sklar A (1959). "Fonctions de Répartition an Dimensions et Leurs Marges." *Publications de l'Institut de statistique de l'Université de Paris*, **8**, 229–231. doi:123.

Teimouri M (2021). "EM Algorithm for Mixture of Skew-normal Distributions Fitted to Grouped Data." *Journal of Applied Statistics*, **48**(7), 1154–1179.

Wei Z, Conlon EM, Wang T (2021). "Asymmetric Dependence in the Stochastic Frontier Model Using Skew Normal Copula." *International Journal of Approximate Reasoning*, **128**, 56–68.

Wei Z, Kim S, Choi B, Kim D (2019). "Multivariate Skew Normal Copula for Asymmetric Sependence: Estimation and Application." *International Journal of Information Technology & Decision Making*, **18**(01), 365–387.

Affiliation:

Ayyub Sheikhi
Department of Statistics, Faculty of Mathematics and Computer,
Shahid Bahonar University of Kerman, Kerman, Iran
E-mail: sheikhy.a@uk.ac.ir

Freshteh Arad
Department of Statistics, Faculty of Mathematics and Computer,
Shahid Bahonar University of Kerman, Kerman, Iran

Radko Mesiar
Department of Mathematics and Descriptive Geometry,
Faculty of Civil Engineering, STU Bratislava, Slovakia
IRAFM, Ostrava, 30. dubna 22, 701 03 Ostrava, Czech Republic

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 52 *Submitted:* 2021-11-12
July 2023 *Accepted:* 2022-08-26