

On relative skewness for multivariate distributions

Félix Belzunce¹ · Julio Mulero² ·
José María Ruíz¹ · Alfonso Suárez-Llorens³

Received: 8 October 2014 / Accepted: 10 March 2015 / Published online: 20 March 2015
© Sociedad de Estadística e Investigación Operativa 2015

Abstract In this paper, we provide a new concept of relative skewness among multivariate distributions, extending to the multivariate case a similar concept in the univariate case. In this case, a random variable Y is said to be more right skewed than a random variable X if there exists an increasing convex transformation which maps X onto Y . Given two random vectors \mathbf{X} and \mathbf{Y} and an appropriate transformation which maps \mathbf{X} onto \mathbf{Y} , we define a new concept of relative skewness assuming the convexity of this transformation. Properties and applications of this concept are given.

Keywords Relative skewness · Standard construction · Multivariate quantile transform · Multivariate convex order · Copula

Mathematics Subject Classification 60E05 · 60E15 · 60E10 · 62H05

✉ Alfonso Suárez-Llorens
alfonso.suarez@uca.es

Félix Belzunce
belzunce@um.es

Julio Mulero
julio.mulero@ua.es

José María Ruíz
jmruiizo@um.es

¹ Dpto. Estadística e Investigación Operativa, Facultad de Matemáticas, Universidad de Murcia, Campus de Espinardo, 30100 Espinardo, Murcia, Spain

² Dpto. Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Alicante, Apartado de correos 99, 03080 Alicante, Spain

³ Dpto. Estadística e Investigación Operativa, Facultad de Ciencias Económicas y Empresariales, Universidad de Cádiz, C/ Duque de Nájera, 8, 11002 Cádiz, Spain

1 Introduction

In probability theory, the concept of skewness means asymmetry or departure from symmetry. The curve of the density function appears distorted or skewed either to the left or to the right and this fact is interpreted as the tail on the curve's right-hand side (left-hand side) being longer than the tail on the left-hand side (right-hand side). The study of skewness is of great interest in diverse areas of statistics, economics and related fields. For example, it is well known in statistics that tails have a direct effect on the efficiency of the estimators of location parameters. In finance, for example, the departure from symmetry of financial returns is crucial for the assessment of financial risk. However, all these situations in a real context are affected by more than one random variable and require the analysis of the whole problem from a multivariate point of view (see, e.g., [Serfling 2004, 2006](#)). The study and comparison of skewness in the multivariate case is an important topic not only in finance or risk theory, but also in environmental sciences and other research fields. In hydrology, for instance, extreme events are of great interest and usually involve multivariate skewed distributions. Unfortunately, the study of skewness becomes even more critical in the multivariate case, where the need to provide skewed multivariate distributions to fit multivariate data sets has been the origin of several proposals of multivariate skewed distributions. The papers by [Arellano-Valle and Azzalini \(2006\)](#), [Azzalini \(1985, 2005\)](#), [Azzalini and Capitanio \(1999\)](#) and [Azzalini and Dalla-Valle \(1996\)](#) are examples of multivariate skew-normal distributions and related distributions. Most proposals in the multivariate case consider multivariate extensions of "skewing mechanisms" developed in the univariate case. We describe two different approaches.

First, [Ferreira and Steel \(2006\)](#) developed a skewing mechanism in which a symmetric distribution F is skewed through a distortion. More precisely, given a symmetric distribution F and a distribution P with support on the interval $[0, 1]$, the distorted version of F through P , that is, $P \circ F$, is the skewed version of F . Recently, [Abtahi and Towhidi \(2013\)](#) gave a unified representation of multivariate skewed distributions extending to the multivariate case the proposal of [Ferreira and Steel \(2006\)](#) for the representation of univariate skewed distributions. They used the Rossenblatt construction (see, e.g., [Rosenblatt 1952](#)) to provide such representation of multivariate skewed distributions.

Another approach is the one developed by [Ley and Painedaveine \(2010\)](#). Given a random variable X with symmetric distribution F , they considered an increasing transformation $\Phi(X)$ of X which provides a skewed version of X . When Φ is increasing, it can be easily seen that $\Phi = G^{-1} \circ F$ (see [Marshall and Olkin 2007](#), Proposition C. 6.), where G is the distribution function of $\Phi(X)$ and G^{-1} is the quantile function associated with G , that is, $G^{-1}(p) = \inf\{x : G(x) \geq p\}$. This skewing mechanism is then extended to the multivariate case considering an appropriate transformation which maps an n -dimensional random vector with a symmetric multivariate distribution onto a skewed version. This particular transformation is a diffeomorphism, i.e., a one-to-one mapping $H : S \subset \mathbb{R}^n \mapsto T \subset \mathbb{R}^n$ such that both H and its inverse H^{-1} are continuously differentiable functions. It is worth mentioning that authors restrict to the case when H has a lower triangular Jacobian matrix with strictly positive diagonal elements.

In both cases, the point of departure is a symmetric distribution, but [Ferreira and Steel \(2006\)](#) considered a transformation (a distortion) of the quantile space, while [Ley and Paindaveine \(2010\)](#) considered a transformation of the sample space.

Traditionally, skewness has been studied through different measures that intends to represent the amount and direction of departure from horizontal symmetry. However, in the univariate case, when dealing with asymmetry or skewness, [van Zwet \(1964\)](#) introduced the concept of relative skewness. Let X be a random variable with interval support and distribution function F . Let us consider another random variable Y with distribution function G ; [van Zwet \(1964\)](#) said that the distribution function G (or the random variable Y) is more right-skewed than the distribution F (or the random variable X) if $G^{-1} \circ F$ is a convex function on the support of X .

The approaches by [van Zwet \(1964\)](#) and [Ley and Paindaveine \(2010\)](#) have in common the transformation of the sample space, but [van Zwet \(1964\)](#) did not consider the symmetry of the random variable X or general increasing transformations of X . In fact, they considered increasing convex transformations of X . That is the reason why the random variable Y is more right-skewed than the random variable X .

This idea arises in a reasonable way when trying to find a formal definition of what it means that one distribution G is more skewed to the right than a distribution F . ([Marshall and Olkin 2007](#), p. 70) provide an explanation of this fact. This idea provides a partial ordering among the set of distributions. In particular, a random variable X with distribution F is said to be less in the convex transform order than a random variable Y with distribution function G , denoted by $X \leq_c Y$, if $G^{-1} \circ F$ is a convex function (see [Shaked and Shanthikumar 2007](#)). Equivalently, it can be seen that $X \leq_c Y$ if, and only if, there exists an increasing and convex transformation Φ , which maps X onto Y , that is, $Y =_{st} \Phi(X)$. It is clear also that, in this case, there exists an increasing and concave transformation Ψ , which maps Y onto X , that is, $X =_{st} \Psi(Y)$.

Next, we describe some situations where this comparison of skewness arises in a natural way.

The first example is the case of increasing convex transformations of some parametric models of random variables. For instance, if we consider a random variable X with normal distribution, with mean equal to 0 and standard deviation equal to 1, then the random variable $Y = \exp(\sigma X + \mu)$, where σ is a positive real number and μ is a real number, is an increasing convex transformation of X and, therefore, more skewed to the right than X , or following the previous notation $X \leq_c Y$. In this case the random variable Y follows a lognormal distribution and, therefore, lognormal distributions are more skewed to the right than normal distributions.

Another example is when we consider a random variable $\exp(\lambda)$ exponentially distributed. In this case, increasing concave (convex) transformations of $\exp(\lambda)$ lead to random variables with an increasing [decreasing] failure rate, denoted by IFR [DFR]. Given a random variable X with absolutely continuous distribution function F and density function f , the hazard or failure rate is defined as $r(x) = f(x)/(1 - F(x))$ for all x , such that $F(x) < 1$. This function is one of the basic functions in the context of reliability and survival analysis, where a random variable X represents the random lifetime of a unit or a mechanism. The hazard rate describes the process of aging and can be considered as the rate at which a unit fails to survive up to a fixed time

x (see Barlow and Proschan 1975; Lai and Xie 2006). Namely, if we denote by E an exponential distribution with parameter 1, i.e., $E \sim \exp(1)$, the IFR [DFR] aging class can be characterized via the univariate convex transform order, that is, given a random variable X (or its distribution), then

$$\begin{aligned} X \text{ is IFR [DFR]} &\Leftrightarrow X \leq_c [\geq_c] E, \\ &\Leftrightarrow X \leq_c [\geq_c] aE + b, \quad \forall a > 0, b \in \mathbb{R}. \end{aligned} \tag{1}$$

As we have mentioned before in the univariate case, when Φ is increasing, then $\Phi = G^{-1} \circ F$. In the multivariate case, it is possible to find such function Φ which maps a random vector \mathbf{X} onto a random vector $\Phi(X)$ with the same distribution of \mathbf{Y} . Next, we describe the construction of such function. Throughout this paper, “increasing” means “nondecreasing” and “decreasing” means “nonincreasing”. We will denote by $=_{st}$ the equality in law, and by $\leq_{a.s.}$ the almost sure inequality. For any random vector \mathbf{X} , or random variable, we will denote by $(\mathbf{X}|A)$ a random vector, or random variable, whose distribution is the conditional distribution of \mathbf{X} given A .

Now, let us consider two n -dimensional random vectors $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ with absolutely continuous distribution.

First, we consider the multivariate quantile transform introduced by Arjas and Lehtonen (1978), O’Brien (1975), Rosenblatt (1952) and Rüschendorf (1981). Essentially, this transformation is also discussed in Ley and Paindaveine (2010). Let us consider the random vector \mathbf{Y} , the **multivariate quantile transform**, also called the **standard construction**, associated with \mathbf{Y} , which is defined recursively as

$$\begin{aligned} Q_{\mathbf{Y},1}(u_1) &= G_{Y_1}^{-1}(u_1), \\ Q_{\mathbf{Y},2}(u_1, u_2) &= G_{(Y_2|Y_1=Q_{\mathbf{Y},1}(u_1))}^{-1}(u_2), \\ &\vdots \\ Q_{\mathbf{Y},n}(u_1, \dots, u_n) &= G_{(Y_n|\bigcap_{j=1}^{n-1} Y_j=Q_{\mathbf{Y},j}(u_1, \dots, u_j))}^{-1}(u_n), \end{aligned} \tag{2}$$

for every $(u_1, u_2, \dots, u_n) \in (0, 1)^n$, where $G_{Y_1}^{-1}(\cdot)$ is the quantile function of Y_1 and for $i = 2, \dots, n$, $G_{(Y_i|\bigcap_{j=1}^{i-1} Y_j=Q_{\mathbf{Y},j}(u_1, \dots, u_j))}^{-1}(\cdot)$ is the quantile function of the univariate conditional random variable given by

$$\left(Y_i \left| \bigcap_{j=1}^{i-1} Y_j = Q_{\mathbf{Y},j}(u_1, \dots, u_j) \right. \right).$$

This known transform is widely used in simulation theory and plays the role of the quantile in the multivariate case. It can be seen that given U_1, \dots, U_n independent and identically distributed random variables uniformly distributed on the interval $(0, 1)$, then, denoting

$$\mathbf{Q}_{\mathbf{Y}}(u_1, \dots, u_n) = (Q_{\mathbf{Y},1}(u_1), Q_{\mathbf{Y},2}(u_1, u_2), \dots, Q_{\mathbf{Y},n}(u_1, \dots, u_n)),$$

we have that

$$(Y_1, \dots, Y_n) =_{st} \mathbf{Q}_Y(U_1, \dots, U_n). \tag{3}$$

Next, we recall the multivariate distributional transform. Let us consider the random vector \mathbf{X} ; the **multivariate distributional transform** is defined recursively as

$$\begin{aligned} D_{\mathbf{X},1}(x_1) &= F_{X_1}(x_1), \\ D_{\mathbf{X},2}(x_1, x_2) &= F_{(X_2|X_1=x_1)}(x_2), \\ &\vdots \\ D_{\mathbf{X},n}(x_1, \dots, x_n) &= F_{(X_n | \bigcap_{j=1}^{n-1} X_j = x_j)}(x_n), \end{aligned} \tag{4}$$

for every (x_1, \dots, x_n) in the support of \mathbf{X} , where $F_{X_1}(\cdot)$ is the distribution function of X_1 and for $i = 2, \dots, n$, $F_{(X_i | \bigcap_{j=1}^{i-1} X_j = x_j)}(\cdot)$ is the distribution function of the conditional distribution $(X_i | \bigcap_{j=1}^{i-1} X_j = x_j)$.

Denoting

$$\mathbf{D}_X(x_1, \dots, x_n) = (D_{\mathbf{X},1}(x_1), D_{\mathbf{X},2}(x_1, x_2), \dots, D_{\mathbf{X},n}(x_1, \dots, x_n)),$$

it can be seen that

$$(U_1, \dots, U_n) =_{st} \mathbf{D}_X(X_1, \dots, X_n). \tag{5}$$

Therefore, if we consider the transform

$$\Phi \equiv \mathbf{Q}_Y \circ \mathbf{D}_X, \tag{6}$$

defined for every (x_1, \dots, x_n) in the support of \mathbf{X} , we have, from (3) and (5) that

$$\mathbf{Y} =_{st} \mathbf{Q}_Y(\mathbf{D}_X(\mathbf{X})),$$

and, hence, the function Φ maps the random vector \mathbf{X} onto \mathbf{Y} .

Remark 1 From (2) and (4), the i -th component of Φ depends only on (x_1, \dots, x_i) and is given by

$$\Phi_1(x_1) = G_{Y_1}^{-1}(F_{X_1}(x_1)), \tag{7}$$

$$\Phi_i(x_1, \dots, x_i) = G_{\left\{ Y_i \mid \bigcap_{j=1}^{i-1} Y_j = \Phi_j(x_1, \dots, x_j) \right\}}^{-1} \left(F_{\left\{ X_i \mid \bigcap_{j=1}^{i-1} X_j = x_j \right\}}(x_i) \right). \tag{8}$$

From the increasingness of both the distribution function and its inverse, it is apparent that

$$\Phi_i(x_1, \dots, x_i),$$

is increasing in x_i , for all $i = 1, \dots, n$. Hence, the Jacobian matrix of Φ is always a lower triangular matrix with strictly positive diagonal elements.

Remark 2 In addition, as a clear extension of the univariate case, [Fernández-Ponce and Suárez-Llorens \(2003\)](#) proved in their Theorem 3.1 that if we take a function $\mathbf{k} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $\mathbf{Y} =_{st} \mathbf{k}(\mathbf{X})$ and \mathbf{k} has a lower triangular Jacobian matrix with strictly positive diagonal elements, then \mathbf{k} has necessarily the form of the function Φ given in (6).

The purpose of this paper is to provide a new concept of relative skewness for multivariate distributions assuming some convexity properties for the function Φ . The organization of the paper is the following. In Sect. 3, we define and study a criteria of relative skewness based on convexity properties for the function Φ . We provide some properties and examples. In Sect. 4, we study the case of random vectors with the same copula and provide several examples for this case. Along the paper, we assume absolute continuity of the multivariate distributions and convex supports for random vectors and random variables.

2 Relative skewness of multivariate distributions

In this section, we consider a new multivariate convex order based on the generalization of the convexity of $G^{-1} \circ F$ to the multivariate case for the function $\Phi = \mathbf{Q}_Y \circ \mathbf{D}_X$. This generalization is clearly inspired on the multivariate dispersive order proposed by [Fernández-Ponce and Suárez-Llorens \(2003\)](#) and also in the skewing mechanism by [Ley and Painsaveine \(2010\)](#). Along this section, we will assume that the random variables or vectors, upon which we consider convex transformations, have a convex support and, analogously to the previous section, also absolutely continuous distribution functions. Finally, we will restrict our study to the case when the function Φ , defined in (6), is differentiable.

We start by recalling the definition of a multivariate convex function (see [Marshall et al. 2011](#), for more details).

Let $\mathbf{f} : \mathbf{S} \rightarrow \mathbb{R}$ be a real-valued function defined on a convex set $\mathbf{S} \subseteq \mathbb{R}^k$, $k \geq 1$. Then, \mathbf{f} is convex on the set S if for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{S}$ and for all $\lambda \in (0, 1)$ we have

$$\mathbf{f}(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda \mathbf{f}(\mathbf{x}_1) + (1 - \lambda) \mathbf{f}(\mathbf{x}_2).$$

In literature, there exist many interesting characterizations of convex functions. Next, we recall some of these characterizations. Note that when we use them, we will assume the regularity conditions that make them possible. From now on, given a vector (or a matrix) \mathbf{v} , we denote as \mathbf{v}^t the transposition of \mathbf{v} .

Characterization 1 *If the function $\mathbf{f} : \mathbf{S} \rightarrow \mathbb{R}$, $\mathbf{S} \subseteq \mathbb{R}^k$, is differentiable in the interior of its support, the convexity is equivalent to check*

$$(\mathbf{f}(\mathbf{x}_2) - \mathbf{f}(\mathbf{x}_1))^t \geq \nabla \mathbf{f}(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1)^t,$$

for all $\mathbf{x}_1, \mathbf{x}_2$ in the support of \mathbf{f} , where

$$\nabla \mathbf{f}(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1)^t = \sum_{i=1}^k \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x}_1)(x_{2i} - x_{1i}),$$

represents the classical tangent hyperplane to the hypersurface given by \mathbf{f} at \mathbf{x}_1 .

Characterization 2 *If the function $\mathbf{f} : \mathbf{S} \rightarrow \mathbb{R}, \mathbf{S} \subseteq \mathbb{R}^k$, is twice differentiable in the interior of its support, the convexity is equivalent to check if the Hessian, denoted by $\nabla^2 \mathbf{f}(\mathbf{x})$, is a semidefinite positive matrix, for every \mathbf{x} in the support of \mathbf{f} .*

We recall that by Young’s theorem, the Hessian of any function, for which all second partial derivatives are continuous, is symmetric for all values of the argument of the function. Finally, attending to the Sylvester’s criterion, $\nabla^2 \mathbf{f}(\mathbf{x})$ is semidefinite positive if, and only if, all its principal minors are non-negative.

Definition 1 Let \mathbf{X} and \mathbf{Y} be two n -dimensional random vectors. Let Φ be the function defined in (6) which maps \mathbf{X} onto \mathbf{Y} . Then, \mathbf{X} is said to be smaller than \mathbf{Y} in the **multivariate convex transform order**, for short mct order and denoted by $\mathbf{X} \leq_{mct} \mathbf{Y}$, if, and only if, the i -th component of Φ, Φ_i , is convex in its support for all $i = 1, \dots, n$.

Roughly speaking, $\mathbf{X} \leq_{mct} \mathbf{Y}$ means that the transformation Φ that maps \mathbf{X} onto \mathbf{Y} is a multivariate convex function. Hence, Definition 1 is a clear generalization of the concept of relative skewness proposed by van Zwet (1964). It is worth mentioning that the transformation Φ depends on the ordering of the marginal distributions. Note that we first obtain Φ_1 from the marginal distributions X_1 and Y_1 , and conditioned on every possible realization $\Phi_1(x_1)$ we next construct $\Phi_2(x_1, x_2)$. Continuing this procedure, we finally arrive at Φ . Far from being a disadvantage, this dependency provides a formal way of explaining that one random vector \mathbf{X} is more “directionally skewed” than a random vector \mathbf{Y} , where directionally skewed is a clear extension of the right-skewed univariate concept. We will see later that Proposition 1 reinforces this remark. For such a reason, given two random vectors ordered in the mct-order sense, we cannot expect the mct order to hold for any arbitrary permutation of the marginal distributions.

From (6) and Remark 1, we point out that the function Φ is the kind of transformations that Ley and Paindaveine (2010) use to define a skewing mechanism. Namely, the authors only require that transformations that map a multivariate symmetric distribution onto a skewed version cannot be odd functions. This is due to the fact that if we map a symmetric distribution, assuming symmetry around the null vector, using an odd mapping, we will obtain another symmetric distribution. We recall that, except affine transformations, convex functions cannot be odd functions. We will analyze in Proposition 2 the case when Φ is an affine transformation.

From Remark 2 and although it is not our main purpose, we can consider any multivariate convex transformation having a lower triangular Jacobian matrix with strictly positive diagonal elements to provide a skewing mechanism. For example, the following transformation $\Phi(x_1, \dots, x_n) = (e^{x_1}, \dots, e^{x_n})$ will map any symmetric multivariate distribution, \mathbf{X} , onto a skewed version, $\mathbf{Y} = \Phi(\mathbf{X})$, in the mct-order

sense. In the latter case, \mathbf{X} and \mathbf{Y} share a common dependence structure, copula, as we will see in Sect. 4.

Next, we provide a particular interpretation. Denoting by $J_\Phi(\mathbf{x})$ the Jacobian matrix of Φ at \mathbf{x} and using Characterization 1 for each $\Phi_i, i = 1, \dots, n$, it is apparent that Definition 1, in case of differentiability, is equivalent to check the following inequality:

$$(\Phi(\mathbf{x}_2) - \Phi(\mathbf{x}_1))^t \geq J_\Phi(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1)^t, \tag{9}$$

which contains all the information of the tangent hyperplanes given by $\nabla\Phi_i, i = 1, \dots, n$. Note that from Remark 1, the Jacobian matrix of Φ is a lower triangular matrix with strictly positive diagonal elements having the following form:

$$J_\Phi(\mathbf{x}) = \begin{pmatrix} \frac{\partial\Phi_1}{\partial x_1}(x_1) & 0 & \dots & 0 \\ \frac{\partial\Phi_2}{\partial x_1}(x_1, x_2) & \frac{\partial\Phi_2}{\partial x_2}(x_1, x_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\Phi_n}{\partial x_1}(x_1, \dots, x_n) & \frac{\partial\Phi_n}{\partial x_2}(x_1, \dots, x_n) & \dots & \frac{\partial\Phi_n}{\partial x_n}(x_1, \dots, x_n) \end{pmatrix}.$$

Due to the fact that Φ maps the multivariate quantile transform of \mathbf{X} to the corresponding of \mathbf{Y} , i.e., $\Phi(\mathbf{Q}_\mathbf{X}(\mathbf{u})) = \mathbf{Q}_\mathbf{Y}(\mathbf{u})$, for all $\mathbf{u} = (u_1, \dots, u_n), u_i \in (0, 1)$, condition (9) can be interpreted as a particular distance between multivariate quantiles:

$$(\mathbf{Q}_\mathbf{Y}(\mathbf{v}) - \mathbf{Q}_\mathbf{Y}(\mathbf{u}))^t \geq J_\Phi(\mathbf{Q}_\mathbf{X}(\mathbf{u}))(\mathbf{Q}_\mathbf{X}(\mathbf{v}) - \mathbf{Q}_\mathbf{X}(\mathbf{u}))^t,$$

for all $\mathbf{v} = (v_1, \dots, v_n)$ and $\mathbf{u} = (u_1, \dots, u_n)$.

Another interesting interpretation of the mct order is given by the following result.

Proposition 1 *Let \mathbf{X} and \mathbf{Y} be two n -dimensional random vectors. If $\mathbf{X} \leq_{mct} \mathbf{Y}$, then*

$$X_1 \leq_c Y_1, \tag{10}$$

$$\left(X_i \left| \bigcap_{j=1}^{i-1} \{X_j = Q_{\mathbf{X},j}(u_1, \dots, u_j)\} \right. \right) \leq_c \left(Y_i \left| \bigcap_{j=1}^{i-1} \{Y_j = Q_{\mathbf{Y},j}(u_1, \dots, u_j)\} \right. \right), \tag{11}$$

for $i = 2, \dots, n$ and for all u_i , such that $0 < u_i < 1, i = 1, \dots, n$.

Proof Under hypothesis assumption, $\Phi_i(x_1, \dots, x_i)$ is convex for all $i = 1, \dots, n$. Therefore, it is also convex in x_i , when x_1, \dots, x_{i-1} remains fixed. If we take into account that Φ maps the multivariate quantile transform of \mathbf{X} to the corresponding \mathbf{Y} , the proof follows directly by just observing the expressions (7) and (8) and recalling the definition of the univariate convex order. \square

Therefore, the univariate conditional distributions are ordered in the sense of van Zwet (1964), i.e., the conditional distributions of \mathbf{X} can be interpreted to be more right-skewed than those of \mathbf{Y} .

Next, we consider an example to show that (10) and (11) in Proposition 1 are not sufficient conditions for the mct order.

Example 1 Let $\mathbf{X} = (X_1, X_2)$ be a non-negative bivariate distribution and let $m_1 \geq 1$ and $m_2 \geq 1$ be two fixed constants. Let us consider the random vector $\mathbf{Y} = (Y_1, Y_2)$ given by

$$\begin{aligned} (Y_1, Y_2) &= \mathbf{k}(X_1, X_2), \\ &= (k_1(X_1), k_2(X_1, X_2)), \\ &= (X_1^{m_1}, X_1^{m_1} X_2^{m_2}). \end{aligned}$$

Since Remark 2, it is apparent that the function Φ , given in (6), which maps \mathbf{X} onto \mathbf{Y} has the form

$$\Phi(x_1, x_2) = \mathbf{Q}_Y(\mathbf{D}_X(x_1, x_2)) = (x_1^{m_1}, x_1^{m_1} x_2^{m_2}).$$

If we compute the Hessian matrix of Φ_2 , we obtain

$$\nabla^2 \Phi_2(x_1, x_2) = \begin{pmatrix} m_1(m_1 - 1)x_1^{m_1-2}x_2^{m_2} & m_1m_2x_1^{m_1-1}x_2^{m_2-1} \\ m_1m_2x_1^{m_1-1}x_2^{m_2-1} & m_2(m_2 - 1)x_2^{m_2-2}x_1^{m_1} \end{pmatrix},$$

where

$$\det(\nabla^2 \Phi_2(x_1, x_2)) = -(m_1 + m_2 - 1)m_1m_2x_1^{2m_1-2}x_2^{2m_2-2} \leq 0,$$

which obviously does not achieve positive values. Therefore, $\nabla^2 \Phi_2(x_1, x_2)$ is not semidefinite positive and $\mathbf{X} \not\prec_{mct} \mathbf{Y}$. However, it is apparent that $\Phi_1(x_1)$ is convex in x_1 , and $\Phi_2(x_1, x_2)$ is convex in x_2 when x_1 remains fixed.

Recalling the univariate convex order, given two univariate random variables X and Y , $X =_C Y$ if, and only if, $Y =_{st} aX + b$ for all $a > 0$ and real b (see Marshall and Olkin 2007, Proposition C.9). Therefore, it is natural to wonder if a similar property holds for the mct order: the answer is yes, as we will see in Proposition 2. To prove it, we recall first the inverse function theorem (see, e.g., Burkill and Burkill 2002).

Lemma 1 *Let $A \subseteq \mathbb{R}^n$ be an open set and $\Phi : A \rightarrow \mathbb{R}^n$ a differentiable and continuous function with differentiable and continuous inverse Φ^{-1} . Then,*

$$J_{\Phi^{-1}}(\mathbf{y})|_{\mathbf{y}=\Phi(\mathbf{x})} = (J_{\Phi}(\mathbf{x}))^{-1},$$

for all $\mathbf{x} \in A$.

Proposition 2 *Let $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ be two random vectors. Then,*

$$\mathbf{X} =_{mct} \mathbf{Y},$$

if, and only if,

$$\mathbf{Y}^t =_{st} \mathbf{A}\mathbf{X}^t + \mathbf{b},$$

for a lower triangular matrix $\mathbf{A} = (a_{ij})$ with diagonal elements, $a_{ii} > 0, i = 1 \dots, n$, and a column matrix \mathbf{b} .

Proof First, we will prove the sufficient condition. If $\mathbf{Y}^t =_{st} \mathbf{A}\mathbf{X}^t + \mathbf{b}$, as specified previously, using Remark 2 we obtain that $\Phi(\mathbf{x}) = Q_{\mathbf{Y}}(D_{\mathbf{X}}(\mathbf{x})) = \mathbf{A}\mathbf{x}^t + \mathbf{b}$. Hence, it is apparent that $\mathbf{X} \leq_{mct} \mathbf{Y}$. Just observing that $\mathbf{X}^t =_{st} \mathbf{A}^{-1}(\mathbf{Y}^t - \mathbf{b})$ and taking into account that \mathbf{A}^{-1} is also a lower triangular matrix with strictly positive diagonal elements, using again Remark 2, $\mathbf{Y} \leq_{mct} \mathbf{X}$ holds with a similar argument.

We will show now the necessary condition. Let us suppose that $\mathbf{X} =_{mct} \mathbf{Y}$. Note that the function Φ , defined in (6), which maps \mathbf{X} to \mathbf{Y} has a lower triangular Jacobian matrix with diagonal elements strictly positive. If we denote by Φ^* the function that follows from (6), exchanging \mathbf{X} by \mathbf{Y} , then it is not difficult to see that $\Phi^* = \Phi^{-1}$. Hence, by hypothesis assumption, the components of Φ and Φ^{-1} are convex functions. Hence,

$$J_{\Phi}(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1)^t \leq (\Phi(\mathbf{x}_2) - \Phi(\mathbf{x}_1))^t, \tag{12}$$

and, from Lemma 1, we have

$$(J_{\Phi}(\mathbf{x}_1))^{-1}(\Phi(\mathbf{x}_2) - \Phi(\mathbf{x}_1))^t \leq (\mathbf{x}_2 - \mathbf{x}_1)^t, \tag{13}$$

where $\mathbf{x}_2, \mathbf{x}_1$ are in the support of \mathbf{X} .

Let us see that these inequalities imply that $\Phi(\mathbf{x}) = \mathbf{A}\mathbf{x}^t + \mathbf{b}$. Let us proceed by induction on $i = 1, \dots, n$.

For the case $i = 1$, the result is trivial, because we have that $\Phi_1(x_1)$ and $\Phi_1^{-1}(x_1)$ are increasing and convex and, therefore, $\Phi_1(x_1) = a_{11}x_1 + b_1$ where $a_{11} > 0$.

Let us assume that this is true for $j = 1, \dots, i - 1$, that is,

$$\Phi_j(x_1, \dots, x_j) = a_{1j}x_1 + \dots + a_{jj}x_j + b_j,$$

and let us see that it is true for $j = i$. Then, we can write

$$J_{(\Phi_1, \dots, \Phi_i)}(x_1, \dots, x_i) = \left(\begin{array}{c|c} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline A_{i-1} & \\ \hline \frac{\partial \Phi_i}{\partial x_1}(x_1, \dots, x_i) \dots \frac{\partial \Phi_i}{\partial x_{i-1}}(x_1, \dots, x_i) & \frac{\partial \Phi_i}{\partial x_i}(x_1, \dots, x_i) \end{array} \right),$$

and

$$(J_{(\Phi_1, \dots, \Phi_i)}(x_1, \dots, x_i))^{-1} = \left(\begin{array}{c|c} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline A_{i-1}^{-1} & \\ \hline B(x_1, \dots, x_i) & \left(\frac{\partial \Phi_i}{\partial x_i}(x_1, \dots, x_i) \right)^{-1} \end{array} \right),$$

where A_{i-1} is a lower triangular matrix with dimension $(i - 1) \times (i - 1)$ and

$$B(x_1, \dots, x_i) = -\frac{1}{\frac{\partial \Phi_i}{\partial x_i}(x_1, \dots, x_i)} \left(\frac{\partial \Phi_i}{\partial x_1}(x_1, \dots, x_i), \dots, \frac{\partial \Phi_i}{\partial x_{i-1}}(x_1, \dots, x_i) \right) A_{i-1}^{-1}.$$

From (12), taking $\mathbf{x}_2 = (x_{21}, \dots, x_{2n})$ and $\mathbf{x}_1 = (x_{11}, \dots, x_{1n})$ in the support of \mathbf{X} , we obtain that

$$\sum_{j=1}^i \frac{\partial \Phi_i}{\partial x_j}(x_{11}, \dots, x_{1i})(x_{2j} - x_{1j}) \leq \Phi_i(x_{21}, \dots, x_{2i}) - \Phi_i(x_{11}, \dots, x_{1i}). \tag{14}$$

From (13), we also have that

$$x_{2i} - x_{1i} \geq B(x_{11}, \dots, x_{1i}) \begin{pmatrix} \Phi_1(x_{21}) - \Phi_1(x_{11}) \\ \vdots \\ \Phi_{i-1}(x_{21}, \dots, x_{2(i-1)}) - \Phi_{i-1}(x_{11}, \dots, x_{1(i-1)}) \end{pmatrix} + \frac{1}{\frac{\partial \Phi_i}{\partial x_i}(x_{11}, \dots, x_{1i})} (\Phi_i(x_{21}, \dots, x_{2i}) - \Phi_i(x_{11}, \dots, x_{1i})).$$

Now, taking into account the expression of $B(\cdot)$ and the induction hypothesis, the previous inequality is equivalent to

$$x_{2i} - x_{1i} \geq -\frac{1}{\frac{\partial \Phi_i}{\partial x_i}(x_{11}, \dots, x_{1i})} \sum_{j=1}^{i-1} \frac{\partial \Phi_i}{\partial x_j}(x_{11}, \dots, x_{1i})(x_{2j} - x_{1j}) + \frac{1}{\frac{\partial \Phi_i}{\partial x_i}(x_{11}, \dots, x_{1i})} (\Phi_i(x_{21}, \dots, x_{2i}) - \Phi_i(x_{11}, \dots, x_{1i})).$$

From this inequality, we get

$$\sum_{j=1}^i \frac{\partial \Phi_i}{\partial x_j}(x_{11}, \dots, x_{1i})(x_{2j} - x_{1j}) \geq (\Phi_i(x_{21}, \dots, x_{2i}) - \Phi_i(x_{11}, \dots, x_{1i})).$$

Therefore, from previous inequality and (14) we get

$$\begin{aligned} \sum_{j=1}^i \frac{\partial \Phi_i}{\partial x_j}(x_{11}, \dots, x_{1i})(x_{2j} - x_{1j}) &\leq \Phi_i(x_{21}, \dots, x_{2i}) - \Phi_i(x_{11}, \dots, x_{1i}) \\ &\leq \sum_{j=1}^i \frac{\partial \Phi_i}{\partial x_j}(x_{11}, \dots, x_{1i})(x_{2j} - x_{1j}). \end{aligned}$$

Hence, Φ_i is an affine transformation. □

Proposition 2 is consistent with the skewing mechanism proposed by Ley and Paindaveine (2010). This is due to the fact that an affine transformation with a column matrix $\mathbf{b} = \mathbf{0}$ is an odd function. Therefore, an affine transformation cannot be considered as a skewing mechanism.

Now, we present an example where we show that all elliptically contoured distributions sharing a common generator are equal in the mct-order sense. First, we recall that \mathbf{X} has an elliptically contoured distribution, denoted by $E_n(\mu, \Sigma, g)$, if its density function can be expressed as

$$f(\mathbf{x}) = k \cdot g((\mathbf{x} - \mu)^t \Sigma^{-1} (\mathbf{x} - \mu)),$$

where k is the scale factor, μ is the median vector (which is also the mean vector if the latter exists), Σ is a symmetric positive definite matrix which is proportional to the covariance matrix if the latter exists, and g is a function mapping from the non-negative reals to the non-negative reals giving a finite area under the curve.

Example 2 Let $\mathbf{X} \sim E_n(\mu_1, \Sigma_1, g)$ and $\mathbf{Y} \sim E_n(\mu_2, \Sigma_2, g)$ be two nondegenerate multivariate elliptically contoured distributions sharing a common generator, where $\Sigma_i, i = 1, 2$, are two non-singular symmetric positive definite matrices. According to Theorem 14.5.11 in Harville (1997), we can find two lower triangular matrices A and B such that $AA^t = \Sigma_2$ and $B^t B = \Sigma_1^{-1}$. Furthermore, we have that

$$A^t = \mathbf{D}_A^{1/2} \mathbf{U} \quad \text{and} \quad B = \mathbf{D}_B^{-1/2} (\mathbf{V}^{-1})^t,$$

with \mathbf{U} and \mathbf{V} being the unique unit upper triangular matrices and $\mathbf{D}_A = \{d_{Ai}\}$ and $\mathbf{D}_B = \{d_{Bi}\}$ being the unique diagonal matrices with positive elements such that

$$\Sigma_1 = \mathbf{V}^t \mathbf{D}_B \mathbf{V} \quad \text{and} \quad \Sigma_2 = \mathbf{U}^t \mathbf{D}_A \mathbf{U},$$

where $\mathbf{D}_A^{1/2} = \{\sqrt{d_{Ai}}\}$ and similarly for $\mathbf{D}_B^{1/2}$. The \mathbf{U} and \mathbf{V} matrices are computed using the Cholesky decomposition (Harville 1997).

According to Remark 2 and from the well-known fact that elliptically contoured distributions are preserved by affine transformations, the function Φ , given in (6), which maps \mathbf{X} onto \mathbf{Y} has the form

$$\Phi(\mathbf{x}) = \mathbf{Q}_Y(\mathbf{D}_X(\mathbf{x})) = AB(\mathbf{x} - \mu_1) + \mu_2.$$

Consequently, the Jacobian matrix of Φ satisfies $J_\Phi = AB$. Due to the fact that the Jacobian matrix is constant, it follows directly that the Hessian matrix of Φ_i is the null matrix for all $i = 1, \dots, n$, and analogously for the function that maps \mathbf{Y} to \mathbf{X} . Hence, $\mathbf{X} =_{mct} \mathbf{Y}$.

We would like to emphasize that the elliptical family contains the multivariate normal family that corresponds to the functional parameter $g(t) = \exp(-t/2)$, which is a particular case of the power exponential family, $g(z) = \exp(-z^\beta)/2$. It also contains a number of widely used subfamilies, which can be useful for robustness purpose as the

multivariate t-distribution family, the multivariate symmetric generalized hyperbolic family, etc.

It is easy to see that the mct order is closed under conjunctions and verifies a sort of marginalization closure property. We state these properties without proof.

Proposition 3 *The following properties hold:*

- (i) *Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be a set of independent random vectors where the dimension of \mathbf{X}_i is k_i , $i = 1, \dots, n$. Let $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be a set of independent random vectors where the dimension of \mathbf{Y}_i is k_i , $i = 1, \dots, n$. If $\mathbf{X}_i \leq_{mct} \mathbf{Y}_i$ for $i = 1, \dots, n$, then*

$$(\mathbf{X}_1, \dots, \mathbf{X}_n) \leq_{mct} (\mathbf{Y}_1, \dots, \mathbf{Y}_n).$$

- (ii) *Let $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ be two n -dimensional random vectors. Let $1 \leq i \leq n$, $\mathbf{X}_I = (X_1, \dots, X_i)$ and $\mathbf{Y}_I = (Y_1, \dots, Y_i)$. If $\mathbf{X} \leq_{mct} \mathbf{Y}$, then $\mathbf{X}_I \leq_{mct} \mathbf{Y}_I$.*

Due to the fact that the composition of multivariate convex functions is not always convex, the multivariate transform convex order does not satisfy the transitive property. Let us see the following example:

Example 3 Let $\mathbf{X} = (X_1, X_2)$, $\mathbf{Y} = (X_1^2, X_2^2)$ and $\mathbf{Z} = (2X_1^2, -3X_1^2 + X_2^2)$ be three bivariate random vectors on $(0, \infty)^2$. From Remark 2, a straightforward computation shows that

$$\begin{aligned} \Phi^{(1)}(x_1, x_2) &= \mathbf{Q}_Y(\mathbf{D}_X(x_1, x_2)) = (x_1^2, x_2^2), \\ \Phi^{(2)}(y_1, y_2) &= \mathbf{Q}_Z(\mathbf{D}_Y(y_1, y_2)) = (2y_1, -3y_1 + y_2), \\ \Phi(x_1, x_2) &= \mathbf{Q}_Z(\mathbf{D}_X(x_1, x_2)) = (2x_1^2, -3x_1^2 + x_2^2). \end{aligned}$$

Just computing the Hessian matrices of all component functions, it is easily obtained that $\mathbf{X} \leq_{mct} \mathbf{Y} =_{mct} \mathbf{Z}$, but $\mathbf{X} \not\leq_{mct} \mathbf{Z}$.

However, transitive property holds for some particular transformations, as we can see in the following proposition. The proof is a direct consequence of the composition of convex functions (see Marshall et al. 2011, Proposition B.7.) and it has been omitted.

Proposition 4 *Let \mathbf{X} , \mathbf{Y} and \mathbf{Z} be three n -dimensional random vectors such that $\mathbf{X} \leq_{mct} \mathbf{Y} \leq_{mct} \mathbf{Z}$, and let $\Phi^{(1)} \equiv \mathbf{Q}_Y \circ \mathbf{D}_X$ and $\Phi^{(2)} \equiv \mathbf{Q}_Z \circ \mathbf{D}_Y$ as described in (6). If $\Phi^{(2)}(\mathbf{y})$ is increasing for all $\mathbf{y} \in \mathbb{R}^n$, then $\mathbf{X} \leq_{mct} \mathbf{Z}$.*

3 On relative skewness for random vectors with the same copula

In this section, we discuss the case of random vectors with the same copula. A copula C is a cumulative distribution function with uniform margins on $[0, 1]$. The notion of copula was introduced by Sklar (1959). The main purpose of copulas is to describe

the interrelation of several random variables. Given a random vector \mathbf{X} with margins F_1, \dots, F_n , there exists a copula C such that

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)).$$

Moreover, if F_1, \dots, F_n are continuous, then C is unique (see [Nelsen 1999](#), for a complete review about copulas). On the other hand, any copula evaluated with some margins in the right way leads to a multivariate distribution function. Next, we show that for random vectors with the same copula, the mct order is equivalent to compare in the convex order the marginal distributions.

Theorem 1 *Let $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ be two random vectors sharing a common copula. Then, $\mathbf{X} \leq_{mct} \mathbf{Y}$ if, and only if, $X_i \leq_c Y_i$ for all $i = 1, \dots, n$.*

Proof [Arias-Nicolás et al. \(2005\)](#) showed that, for two random vectors with the same copula, the function Φ , defined in (6), which maps \mathbf{X} to \mathbf{Y} satisfies that

$$\Phi_i(x_1, \dots, x_i) = G_{Y_i}^{-1}(F_{X_i}(x_i)), \quad (15)$$

for all $i = 1, \dots, n$. The result follows directly from expression (15) and recalling the definition of the univariate convex order. \square

As an immediate consequence of Theorem 1, for two random vectors having independent components, independence copula, the mct order is reduced to the univariate convex order between the marginal distributions.

From Proposition 4 and Theorem 1, we obtain the following corollary.

Corollary 1 *Let $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ be two random vectors such that $\mathbf{X} \leq_{mct} \mathbf{Y}$. Then, $\mathbf{X} \leq_{mct} \mathbf{W}$ for all random vector $\mathbf{W} = (W_1, \dots, W_n)$ having the same copula than \mathbf{Y} such that $Y_i \leq_c W_i$, for all $i = 1, \dots, n$.*

Proof Using Theorem 1, $\mathbf{Y} \leq_{mct} \mathbf{W}$ holds and the transformation $\Phi \equiv \mathbf{Q}_Y \circ \mathbf{D}_W$, defined in (6), only depends on the marginal distributions, i.e., it can be expressed as

$$\Phi_i(x_1, \dots, x_n) = G_{Y_i}^{-1}(F_{W_i}(x_i)),$$

where it is apparent that Φ is trivially increasing. The proof concludes just using Proposition 4. \square

Theorem 1 can be used to provide many examples of random vectors ordered in the mct order. Just fixing a copula, many multivariate random vectors can be ordered in the mct order via the univariate comparison of the marginal distributions. We would like to mention that the fact of using copulas for providing a flexible skewing mechanism is mentioned in [Ley and Painsaveine \(2010\)](#). Due to the well-known fact that the copula is preserved by strictly increasing transformations of the marginal distributions, given a random vector $\mathbf{X} = (X_1, \dots, X_n)$ with a symmetric distribution and copula C , any multivariate increasing convex transformation of the type $\Phi(x_1, \dots, x_n) =$

$(\Phi_1(x_1), \dots, \Phi_n(x_n))$ will provide a skewed version of \mathbf{X} , say $\mathbf{Y} = \Phi(\mathbf{X})$, in the mct-order sense having the same copula C .

Next, we describe some situations where Theorem 1 can be applied.

3.1 Multivariate normal and lognormal distributions

Given a random vector $\mathbf{X} = (X_1, \dots, X_n)$ with multivariate normal distribution (see Example 2), the random vector $\mathbf{Y} = (Y_1, \dots, Y_n)$, where $Y_i = \exp X_i$ for $i = 1, \dots, n$, follows a multivariate lognormal distribution (see Aitchison and Brown 1957). Clearly, \mathbf{X} and \mathbf{Y} share the same copula and the function that maps X_i onto Y_i is convex, for all $i = 1, \dots, n$. Therefore, using Theorem 1, we obtain that $\mathbf{X} \leq_{mct} \mathbf{Y}$. Hence, the multivariate lognormal distribution is a right skewed transformation of a multivariate normal distribution. We also observe that given any other multivariate normal distribution \mathbf{X}' , from Example 2, $\mathbf{X}' =_{mct} \mathbf{X}$ holds. Then, using Proposition 4, $\mathbf{X}' \leq_{mct} \mathbf{Y}$ also holds.

3.2 Multivariate distributions with IFR [DFR] margins

The next corollary provides a situation where we can apply previous ideas.

Corollary 2 *Let $\mathbf{X} = (X_1, \dots, X_n)$ be a random vector having a copula C such that all marginal distributions, $X_i, i = 1, \dots, n$, satisfy the IFR [DFR] aging property. If we consider a random vector $\mathbf{Y} = (Y_1, \dots, Y_n)$ with the same copula C but having shifted exponential marginal distributions, that is $Y_i \sim a_i E + b_i$, where $a_i > 0, b_i \in \mathbb{R}$ and $E \sim \exp(1)$, then $\mathbf{X} \leq_{mct} [\geq_{mct}] \mathbf{Y}$.*

Proof The proof follows easily from Theorem 1 and the univariate characterization of the IFR [DFR] aging property given in (1). □

Taking into account previous results, we can construct a great bunch of examples of multivariate distributions ordered in the mct order. Let us see some examples.

Example 4 Let \mathbf{X} and \mathbf{Y} be two random vectors with the same Gumbel copula given by:

$$C(u_1, u_2) = \exp \left\{ - \left[(-\log u)^{1/\theta} + (-\log v)^{1/\theta} \right]^\theta \right\} \quad \text{for } u_1, u_2 \in (0, 1).$$

Let \mathbf{X} be a random vector with Weibull distributed margins with scale parameter equal to one, i.e., with distribution function given by:

$$F_i(x) = 1 - \exp\{-x^{\beta_i}\} \quad \text{for } i = 1, 2.$$

This bivariate distribution \mathbf{X} can be found in Lu and Bhattacharyya (1990).

Let \mathbf{Y} be a random vector with exponential distributed margins, i.e.,

$$G_i(x) = 1 - \exp\{-\alpha_i x\} \quad \text{for } i = 1, 2.$$

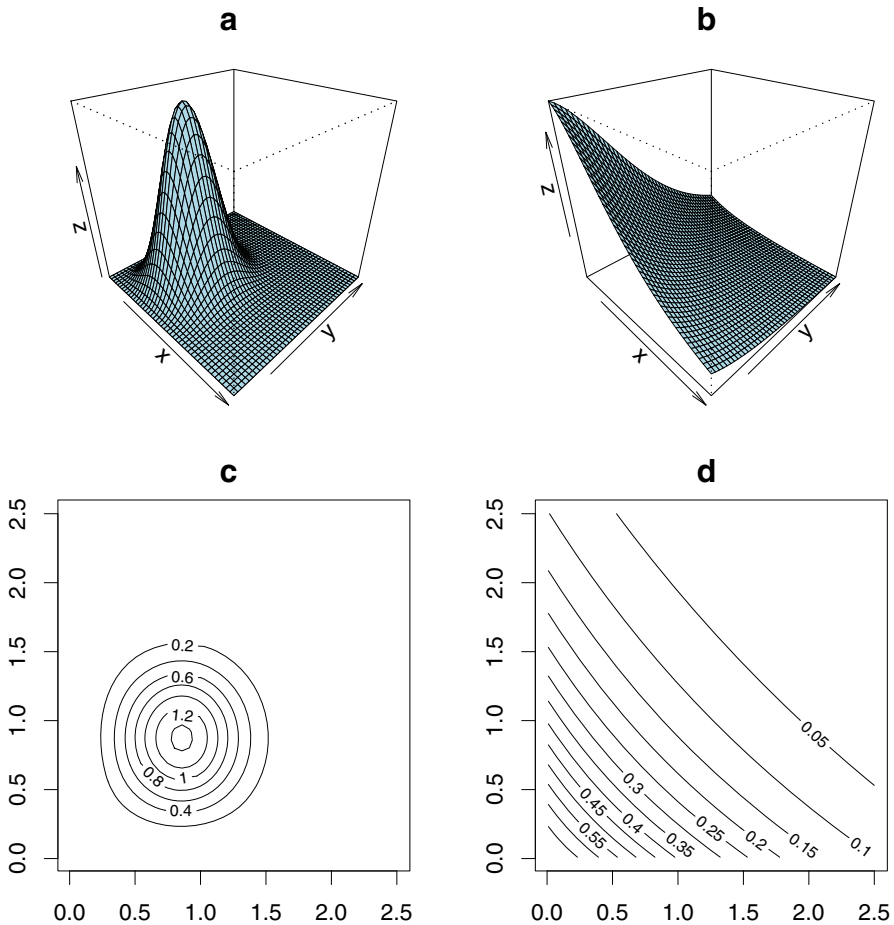


Fig. 1 Joint density function of **a, c** Weibull and **b, d** Gumbel bivariate distributions

Then, we have again that $\mathbf{X} \leq_{mct} \mathbf{Y} [\geq_{mct}]$ if $\beta_i > 1$ [< 1] for $i = 1, 2$, respectively.

In fact, in Fig. 1, we plot the joint density functions of the bivariate Weibull and Gumbel distributions for $\theta = 0.3$, $\beta_1 = \beta_2 = 3$ and $\alpha_1 = \alpha_2 = 1$. Clearly, the bivariate Gumbel distribution is more skewed than the bivariate Weibull distribution.

Example 5 Let us consider a random vector \mathbf{X} with a Clayton copula which is given by

$$C(u_1, u_2) = \left[u_1^{-1/\theta} + u_2^{-1/\theta} + 1 \right]^{-\theta} \quad \text{for } u_1, u_2 \in (0, 1),$$

where $\theta > 0$ and exponentially distributed margins. Let \mathbf{Y} be a bivariate Pareto distribution as defined in Lindley and Singpurwalla (1986). \mathbf{Y} has a Clayton copula

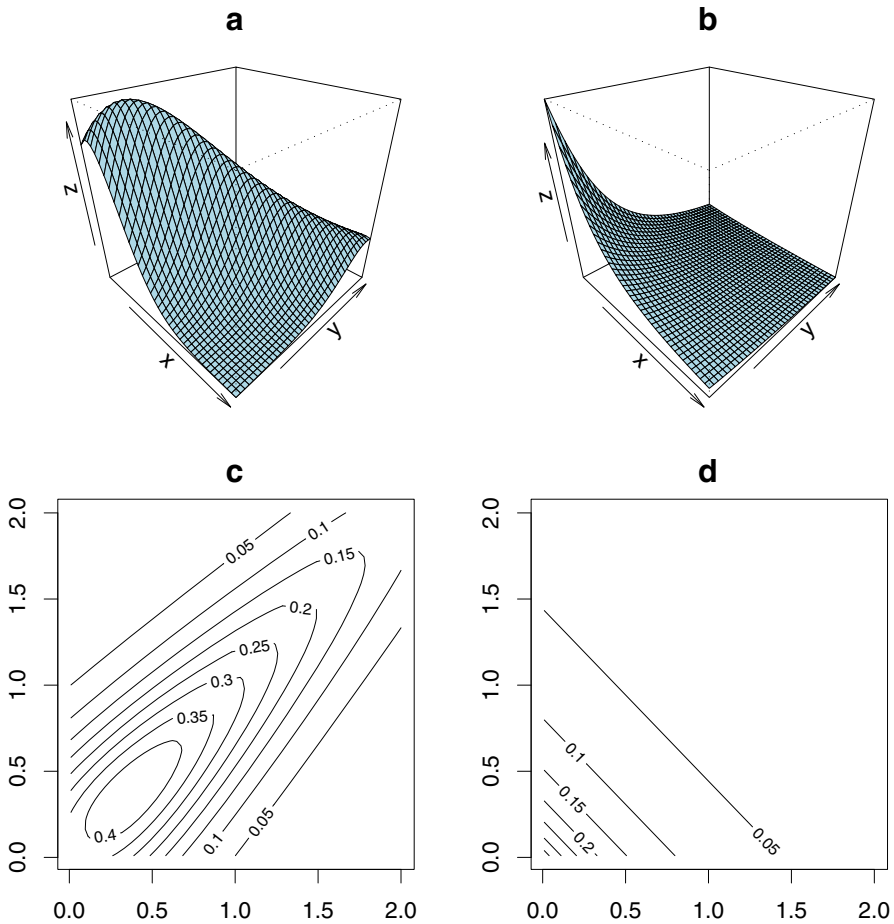


Fig. 2 Joint density function **a**, **c** of the Clayton copula with exponential margins and **b**, **d** Pareto bivariate distributions

and Pareto distributed margins (see Balakrishnan and Lai 2009). It is known that the Pareto distribution is DFR. Hence, again from Corollary 2, we have $\mathbf{X} \leq_{mct} \mathbf{Y}$.

In Fig. 2, we can see the plots of the joint density functions of a distribution with a Clayton copula and exponential margins and a Pareto bivariate distribution for $\theta = 0.3$ in both cases.

3.3 Relative skewness for ordered data

The model of a random vector with ordered components arises in a natural way when we arrange in increasing order a set of observations from a random variable. Another example is the case of epoch times of a counting process, like the case of a nonhomogeneous Poisson process. Epoch times of nonhomogeneous Poisson processes can be

introduced as record values of a proper sequence of random variables, which is another typical example of ordered data. Finally, it is worth mentioning that order statistics is also an interesting way of skewing symmetric distributions (see, e.g., Jones 2004). Given the similarity of several results for order statistics and record values, Kamps (1995a) introduced the model of generalized order statistics. This model provides a unified approach to study order statistics and record values, and several other models of ordered data. First, we recall the definition of generalized order statistics following (Kamps 1995a, b).

Definition 2 Let $n \in \mathbb{N}$, $k \geq 1$, $m_1, \dots, m_{n-1} \in \mathbb{R}$, $M_r = \sum_{j=r}^{n-1} m_j$, $1 \leq r \leq n-1$ be parameters such that $\gamma_r = k + n - r + M_r \geq 1$ for all $r \in 1, \dots, n-1$, and let $\tilde{m} = (m_1, \dots, m_{n-1})$, if $n \geq 2$ ($\tilde{m} \in \mathbb{R}$ arbitrary, if $n = 1$). We call **uniform generalized order statistics** the random vector $(U_{(1,n,\tilde{m},k)}, \dots, U_{(n,n,\tilde{m},k)})$ with joint density function

$$h(u_1, \dots, u_n) = k \left(\prod_{j=1}^{n-1} \gamma_j \right) \left(\prod_{j=1}^{n-1} (1 - u_j)^{m_j} \right) (1 - u_n)^{k-1},$$

on the cone $0 \leq u_1 \leq \dots \leq u_n \leq 1$. Now, given a distribution function F , we call generalized order statistics based on F the random vector

$$(X_{(1,n,\tilde{m},k)}, \dots, X_{(n,n,\tilde{m},k)}) \equiv (F^{-1}(U_{(1,n,\tilde{m},k)}), \dots, F^{-1}(U_{(n,n,\tilde{m},k)})).$$

If F is an absolutely continuous distribution with density f , the joint density function of $(X_{(1,n,\tilde{m},k)}, \dots, X_{(n,n,\tilde{m},k)})$ is given by

$$f(x_1, \dots, x_n) = k \left(\prod_{j=1}^{n-1} \gamma_j \right) \left(\prod_{j=1}^{n-1} \bar{F}(x_j)^{m_j} f(x_j) \right) \bar{F}(x_n)^{k-1} f(x_n),$$

on the cone $F^{-1}(0) \leq x_1 \leq \dots \leq x_n \leq F^{-1}(1)$.

Let us see now several other models that are included in this model. As we have mentioned previously, order statistics and record values are a particular case of this model (see Belzunce 2013, for a detailed review).

Taking $m_i = 0$ for all $i = 1, \dots, n-1$ and $k = 1$, we get the random vector of order statistics $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$ from a set of n independent and identically distributed (i.i.d) observations X_1, X_2, \dots, X_n with common absolutely continuous distribution F ; in particular, we get that $X_{i:n} =_{\text{st}} X_{(i,n,0,1)}$.

Taking $m_i = -1$ for all $i = 1, \dots, n-1$ and $k = 1$, we get the random vector of the first n record values (see Chandler 1952).

Some additional particular cases of GOSs are the following. Taking $m_i = -1$ for all $i = 1, \dots, n-1$ and $k \in \mathbb{N}$, we get k -records. Taking $n = m$, $m_i = R_i$ and $k = R_m + 1$, we get order statistics from Type-II censored data. Another particular case is the that of order statistics under multivariate imperfect repair (see Shaked and Shanthikumar 1986).

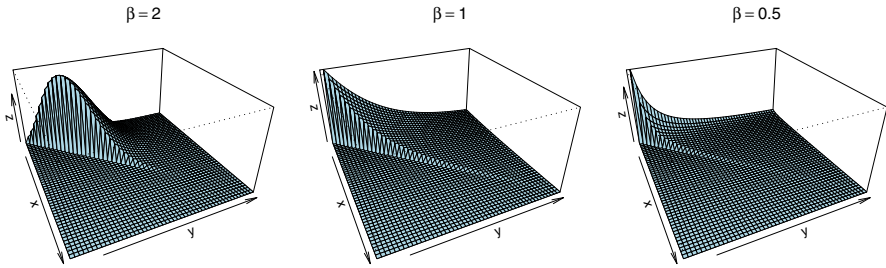


Fig. 3 Joint density function of bivariate random vectors based on the Weibull distribution for $\beta = 2, 1, 0.5$, respectively

Next, we show a property dealing with the comparison of relative skewness for generalized order statistics. Note that the dependency of the ordering of the marginal distribution for the mct order appears in a natural way for random vectors with ordered components.

Theorem 2 *Let \mathbf{X} and \mathbf{Y} be two random vectors of generalized order statistics based on distributions F and G from random variables X and Y , respectively. Then, $\mathbf{X} \leq_{mct} \mathbf{Y}$ if, and only if, $X \leq_c Y$.*

Proof Since the vectors of generalized order statistics have the same copula (see Belzunce et al. 2008), the result follows easily if $X_{(r,n,\tilde{m},k)} \leq_c Y_{(r,n,\tilde{m},k)}$, for all $r = 1, \dots, n$, if, and only if, $X \leq_c Y$. This follows observing that

$$G_{Y_{(r,n,\tilde{m},k)}}^{-1}(F_{X_{(r,n,\tilde{m},k)}}(x)) = G^{-1}(F(x)),$$

where $G_{Y_{(r,n,\tilde{m},k)}}$ and $F_{X_{(r,n,\tilde{m},k)}}$ denote the distribution functions of $Y_{(r,n,\tilde{m},k)}$ and $X_{(r,n,\tilde{m},k)}$, respectively (Cramer and Kamps 2003, see also). \square

In Fig. 3, we can see the plots of the joint density functions of three random vectors of order statistics, \mathbf{X} , \mathbf{Y} and \mathbf{Z} , based on the Weibull distribution with parameters $\lambda = 1$ and $\beta = 2, 1, 0.5$, respectively. As previously mentioned, the aging of the Weibull distribution depends on the value of β . In particular, we have that if $\beta > 1$, then the Weibull distribution is IFR [DFR] and it is smaller [bigger] than the exponential distribution, which is obtained with $\beta = 1$, in the convex order. From Corollary 4 and Theorem 2, we have $\mathbf{X} \leq_{mct} \mathbf{Y} \leq_{mct} \mathbf{Z}$.

3.4 Relative skewness with real data

In environmental sciences, there are situations where we can be interested in the analysis of skewness or concentration of our data. In this context, the study of skewness or concentration of hydrological data, such as droughts or precipitations, has attracted the attention for economic or wealth reasons. They can be described through some fundamental characteristics such as for example duration, denoted by D , and severity, denoted by S , which are usually fitted to the gamma and exponential distributions (see

Table 1 Parameters of the margins of \mathbf{X} and \mathbf{Y}

Station	Abadan	Anazli
Severity (gamma)		
$\hat{\theta}$	1.527	1.497
$\hat{\alpha}$ (shape)	0.737	0.917
Duration (exponential)		
$\hat{\beta}$ (scale)	2.796	3.277
$\hat{\lambda}$	2.125	3.129

Shiau 2006). Moreover, the Clayton copula, as introduced in Example 5, is widely used to model the relationship between drought severity and drought duration. In Shiau and Modarres (2009), a copula-based drought severity duration study is given. They analyzed rainfall data for the period 1954–2003 from two gauge stations in Iran, Abadan and Anazli. Let us consider two random vectors $\mathbf{X} = (S_1, D_1)$ and $\mathbf{Y} = (S_2, D_2)$ with severity duration of Abadan and Anazli, respectively. Shiau and Modarres (2009) showed that both bivariate distributions have a Clayton copula with estimated parameters $\hat{\theta}_{\mathbf{X}} = 1.527$ and $\hat{\theta}_{\mathbf{Y}} = 1.497$, respectively. We provide in Table 1 the marginal distributions.

Just taking into account the estimation error of the dependency parameter, θ , we can assume that \mathbf{X} and \mathbf{Y} share the same copula. On the other hand, van Zwet (1964) also studied the convex transform order between gamma distributions. In particular, if X and Y are univariate gamma distributions with shape parameters α_1 and α_2 , respectively, such that $\alpha_1 \geq \alpha_2$, then $X \leq_c Y$. In our case, we have that the estimation of the shape parameter of S_1 ($\hat{\alpha}_1 = 0.737$) is smaller than that corresponding to S_2 ($\hat{\alpha}_2 = 0.917$); therefore, there is reasonable empirical evidence to affirm that $S_1 \geq_c S_2$. On the other hand, it is apparent that all exponential distributions are equal in the univariate convex transform order. Thus, we have that $S_1 =_c S_2$. Hence, using jointly these results and Proposition 1, $\mathbf{X} \geq_{mct} \mathbf{Y}$ holds and we expect that \mathbf{X} is more right-skewed than \mathbf{Y} .

This type of comparison implies, from Proposition 1, that both margins and conditioned random variables are ordered in the convex order. It is well known that the convex transform order preserves the comparison between the main skewness coefficients, such as the Fisher's and Bowley's coefficients (see MacGillivray 1986) and the coefficient of variation and Gini indexes (see Shaked and Shanthikumar 2007) for concentration. Therefore, data in Anazli are relatively more dispersed than those of Abadan, and therefore they are more unpredictable.

4 Conclusions

The study of skewness is of interest in many areas of research such as risk theory, finance, hydrology, etc. Skewness can be studied through comparisons of many single measures, but stochastic orders and, in particular when we deal with random variables, the univariate convex transform order proposed by van Zwet (1964) provides a more complete comparison between the skewness of two random variables. Unfortunately,

in real phenomena, more than one random variable is involved and its analysis is more complicated.

In this work, we provide a new multivariate stochastic order, the so-called mct order, as a tool based on the standard construction for comparing the skewness of two random vectors. This transformation was firstly proposed by Rosenblatt (1952), and rediscovered periodically, and captures the essence of how a random vector is mapped onto another. The convexity of the standard construction is the key of the cct order which generalizes the univariate version of the convex transform order given by van Zwet (1964).

On the other hand, when one compares two random vectors in the mct order, we also show that it is important to take into account their dependence structures, because in case that they share the same copula, the comparison can be reduced to compare the margins in the univariate convex transform order.

Finally, we illustrate the results with different applications to some probabilistic models such as normal and lognormal distributions and generalized ordered statistics, as well as an example dealing with real data related to environmental sciences.

Acknowledgments The authors acknowledge the comments by the two anonymous referees and the Editor of Test which have improved significantly the presentation of this paper. Félix Belzunce, Julio Mulero and José M. Ruiz acknowledge support received from the Ministerio de Economía y Competitividad (Spain) under grant MTM2012-34023-FEDER.

References

- Abtahi A, Towhidi M (2013) The new unified representation of multivariate. *Statistics* 47:126–140
- Aitchison J, Brown J (1957) *The lognormal distribution*. Cambridge University Press, Cambridge
- Arellano-Valle R, Azzalini A (2006) On the unification of families of skew-normal distributions. *Scand J Stat* 33:561–574
- Arias-Nicolás J, Fernández-Ponce J, Luque-Calvo P, Suárez-Llorens A (2005) Multivariate dispersion order and the notion of copula applied to the multivariate t -distribution. *Probab Eng Inf Sci* 19:363–375
- Arjas E, Lehtonen T (1978) Approximating many server queues by means of single server queues. *Math Oper Res* 3:205–223
- Azzalini A (1985) A class of distribution which includes the normal ones. *Scand J Stat* 12:171–178
- Azzalini A (2005) The skew-normal distribution and related multivariate families (with discussion). *Scand J Stat* 32:159–188
- Azzalini A, Capitanio A (1999) Statistical applications of multivariate skew-normal distributions. *J R Stat Soc Ser B Stat Methodol* 61:579–602
- Azzalini A, Dalla-Valle A (1996) The multivariate skew-normal distribution. *Biometrika* 83:715–726
- Balakrishnan N, Lai C (2009) *Continuous bivariate distributions*, 2nd edn. Springer, New York
- Barlow R, Proschan F (1975) *Statistical theory of reliability and life testing*, 2nd edn. Rinehart and Winston Inc, New York
- Belzunce F (2013) Multivariate comparisons of ordered data. In: Li X (ed) *Stochastic orders in reliability and risks* (Li, H). Springer, New York, pp 83–102
- Belzunce F, Ruíz J, Suárez-Llorens A (2008) On multivariate dispersion orderings based on the standard construction. *Statist Probab Lett* 78:271–281
- Burkill J, Burkill H (2002) *A second course in mathematical analysis*. Cambridge University Press, Cambridge
- Chandler K (1952) The distribution and frequency of record values. *J R Stat Soc Ser B Stat Methodol* 14:220–228
- Cramer E, Kamps U (2003) Marginal distributions of sequential and generalized order statistics. *Metrika* 58:293–310

- Fernández-Ponce J, Suárez-Llorens A (2003) A multivariate dispersion ordering based on quantiles more widely separated. *J Multivar Anal* 85:40–53
- Ferreira J, Steel M (2006) A constructive representation of univariate skewed distributions. *J Am Statist Assoc* 101:823–829
- Harville D (1997) *Matrix algebra from a statistician's perspective*. Springer, New York
- Jones M (2004) Families of distributions arising from distributions of order statistics. *TEST* 13:1–43
- Kamps U (1995a) A concept of generalized order statistics. *J Statist Plann Inference* 48:1–23
- Kamps U (1995b) A concept of generalized order statistics. B.G. Taubner, Stuttgart
- Lai C, Xie M (2006) *Stochastic ageing and dependence for reliability*. Springer, New York
- Ley C, Paindaveine D (2010) Multivariate skewing mechanisms: a unified perspective based on the transformation approach. *Statist Probab Lett* 48:1685–1694
- Lindley D, Singpurwalla N (1986) Multivariate distribution for the life lengths of a system sharing a common environment. *J Appl Probab* 23:418–431
- Lu J, Bhattacharyya G (1990) Some new constructions of bivariate Weibull models. *Ann Inst Statist Math* 23:543–559
- MacGillivray H (1986) Skewness and asymmetry: measures and orderings. *Ann Statist* 14:994–1011
- Marshall A, Olkin I (2007) *Life distributions*. Springer series in statistics. Springer, New York
- Marshall A, Olkin I, Arnold B (2011) *Inequalities: theory of majorization and its applications*. Springer series in statistics. Springer, New York
- Nelson N (1999) An introduction to copulas, vol. 139 of *Lectures notes in statistics*. Springer, New York
- O'Brien G (1975) The comparison method for stochastic processes. *Ann Probab* 3:80–88
- Rosenblatt M (1952) Remarks on a multivariate transformation. *Ann Math Statist* 23:470–472
- Rüschendorf L (1981) Stochastically ordered distributions and monotonicity of the OC-function of sequential probability ratio tests. *Math Operat Statist Ser Stat* 12:327–338
- Serfling R (2004) Nonparametric multivariate descriptive measures based on spatial quantiles. *J Statist Plann Inf* 123:259–278
- Serfling R (2006) Multivariate symmetry and asymmetry. In: Kotz S, Balakrishnan CRN, Vidakovic B (eds) *Encyclopedia of statistical sciences*, 2nd edn. Wiley, New York, pp 5338–5345
- Shaked M, Shanthikumar J (1986) Multivariate imperfect repair. *Math Operationsforsch Statist Ser Statist* 34:437–448
- Shaked M, Shanthikumar J (2007) *Stochastic orders*. Springer series in statistics. Springer, New York
- Shiau J (2006) Fitting drought duration and severity with two-dimensional copulas. *Water Resour Manag* 20:795–815
- Shiau J, Modarres R (2009) Copula-based drought severity-duration-frequency analysis in iran. *Meteorol Appl* 16:481–489
- Sklar A (1959) Fonctions de repartition a n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* 8:229–231
- van Zwet W (1964) Convex transformations of random variables. *Mathematical Centre Tracts*, Amsterdam