

On Relative Skewness for Multivariate Distributions

January 14, 2014

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, Multivariate IFR distributions.

1 Introduction

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 Azzalini and Dalla Valle (1996), Azzalini and Capitanò (1999), Azzalini (2005) and Arellano-Valle and Azzalini (2006) are examples of multivariate skew-normal distributions and related distributions.

Other proposals in the multivariate case consider multivariate extensions of “skewing mechanisms” developed in the univariate case. We describe next two different approaches.

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 an 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 use the Rossenblatt construction (see Arnold et al., 2006) to provide such representation of multivariate skewed distributions.

Another approach is the one developed by Ley and Paindaveine (2010). Given a random variable X with symmetric distribution F , they consider an increasing transformation $\Phi(X)$ of X which provides a skewed version of X . When Φ is increasing, it can be easily seen that $\Phi \equiv G^{-1} \circ F$ (see Proposition C.6. by Marshall and Olkin, 2007), 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 increasing transformation for a random vector.

In both cases, the point of departure is a symmetric distribution, but in the approach by Ferreira and Steel (2006), they consider a transformation (distortion) of the quantile space, and in the approach by Ley and Paindaveine (2010), they consider a transformation of the sample space.

In the univariate case, when dealing with asymmetry or skewness, van Zwet (1964) intro-

duced 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) says 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 and by Ley and Paindaveine have in common the transformation of the sample space, but van Zwet does not consider that the random variable X is symmetric, and does not consider general increasing transformations of X , but 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 equals to 0 and standard deviation equals 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 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 basics 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 ageing and can be considered as the rate at which a unit fails when survive up to a fixed time x (see Barlow and Proschan, 1975, and 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, i.e., given a random variable X (or its distribution) then

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

The purpose of this paper is to extend the idea of van Zwet to the multivariate case. That is, to compare the degree of skewness from one random vector \mathbf{X} to another random vector \mathbf{Y} if there exists some transformation Φ such that maps \mathbf{X} onto \mathbf{Y} , that satisfies some convexity property.

As we have mentioned before in the univariate case, when Φ is increasing then $\Phi = G^{-1} \circ F$. In the multivariate case is possible to find such function Φ which maps a random vector \mathbf{X} onto a random vector $\Phi(\mathbf{X})$ with the same distribution that \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 surely 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 .

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 O'Brien (1975), Arjas and Lehtonen (1978) and Ruschendorf (1981). Let us consider the random vector \mathbf{Y} , the **multivariate quantile transform**, also called the **standard construction**, associated with \mathbf{Y} , 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

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

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

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

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

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

it can be seen that

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

Therefore, if we consider the transform

$$\Phi \equiv \mathbf{Q}_{\mathbf{Y}} \circ \mathbf{D}_{\mathbf{X}}, \quad (6)$$

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

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

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

Remark 1.1. From (2) and (4), the i -th component of Φ depends only on (x_1, \dots, x_i) and it is given by

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

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

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 diagonal elements strictly positive.

Remark 1.2. 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 diagonal elements strictly positive, 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 Section 2, we define and study a criteria of relative skewness based on convexity properties for the function Φ . We provide some properties and examples. In Section 3, we study the case of random vectors with the same copula, and we provide several examples for this case. Section 4 is devoted to define and study a new multivariate IFR [DFR] notion based on this new notion of relative skewness. 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). 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 have 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

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

There exist many interesting characterizations of convex functions in Literature. Along this paper we will assume the regularity conditions that make possible the following well known characterizations, when we use them.

Characterization 2.1. *If the function $\mathbf{f} : \mathbf{S} \rightarrow \mathbb{R}$, $\mathbf{S} \subseteq \mathbb{R}^k$, is differentiable in the interior of their 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 2.2. *If the function $\mathbf{f} : \mathbf{S} \rightarrow \mathbb{R}$, $\mathbf{S} \subseteq \mathbb{R}^k$, is twice differentiable in the interior of their 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 nonnegative.

Definition 2.3. *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 their support for all $i = 1, \dots, n$.*

Denoting by $J_{\Phi}(\mathbf{x})$ the Jacobian matrix of Φ at \mathbf{x} and using Characterization 2.1 for each Φ_i , $i = 1, \dots, n$, it is apparent that Definition 2.3, 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 1.1, the Jacobian matrix of Φ is a lower triangular matrix with

diagonal elements strictly positive having the following form

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

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 $\mathbf{u} = (u_1, \dots, u_n)$.

A meaningful interpretation of the mct-order is given by the following result.

Proposition 2.4. *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 in account that Φ maps the multivariate quantile transform of \mathbf{X} to the corresponding of \mathbf{Y} , the proof follows directly just observing the expressions (7) and (8) and recalling the definition of the univariate convex order. \square

Therefore, the mct-order implies the univariate convex order of the conditional distributions. Next, we will present two examples. First we consider an example to show that (10) and (11) in Proposition 2.4 are not sufficient conditions for the mct-order.

Example 2.5. Let $\mathbf{X} = (X_1, X_2)$ be a nonnegative bivariate distribution and let $m_1 \geq 1$ and $m_2 \geq 1$ two fixed constant. 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 1.2, 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 achieves strictly negative 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.

Now we present an example concerning to the multivariate normal distribution. We will show that all multivariate normal distributions are equal in the mct-order sense.

Example 2.6. Let $\mathbf{X} \rightsquigarrow N_n(\mu_1, \Sigma_1)$ and $\mathbf{Y} \rightsquigarrow N_n(\mu_2, \Sigma_2)$ be two multivariate normal distributions. Since Fernández-Ponce and Suárez-Llorens (2003), Example 4.1, the function Φ , defined in (6), that maps \mathbf{X} to \mathbf{Y} satisfies $J_\Phi = AB$ where A and B are two lower triangular matrices with $AA^t = \Sigma_2$ and $B^tB = \Sigma_1^{-1}$. Furthermore, according to Theorem 14.5.11 in Harville (1997), we have that

$$A^t = \mathbf{D}_A^{1/2}\mathbf{U}, \quad B = \mathbf{D}_B^{-1/2}(\mathbf{V}^{-1})^t \text{ and } \Sigma_1 = \mathbf{V}^t\mathbf{D}_B\mathbf{V},$$

with \mathbf{U} be the unique unit upper triangular matrix and $\mathbf{D}_A = \{d_i\}$ be the unique diagonal matrix such that

$$\Sigma_2 = \mathbf{U}^t\mathbf{D}_A\mathbf{U} \text{ and } \mathbf{D}_A^{1/2} = \{\sqrt{d_i}\}.$$

Similarly for the B matrix. The \mathbf{U} and \mathbf{V} matrices can be calculated using the Cholesky decomposition (see Harville, 1997). Consequently, if J_{Φ} is always a matrix, 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 $\mathbf{X} =_{mct} \mathbf{Y}$.

In previous example the function Φ is an affine transformation. Recalling the univariate convex order, given two univariate random variables X and Y then $X =_c Y$ if and only if $Y =_{st} aX + b$ for all $a > 0$ and real b . (see Proposition C.9 by Marshall and Olkin, 2007). 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 2.8. In order to prove it, we recall first the inverse function theorem.

Lemma 2.7. *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 2.8. *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 1.2 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 in account that \mathbf{A}^{-1} is also a lower triangular matrix with diagonal elements strictly positive, using again Remark 1.2, $\mathbf{Y} \leq_{mct} \mathbf{X}$ 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. Therefore,

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

and, from Lemma 2.7, we have

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

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 is true for $j = i$. Then we can write

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

and

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

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

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

From (13), we also have that

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

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

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

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. □

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 2.9. *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 $1 \leq i \leq n$ and $\mathbf{X}_I = (X_1, \dots, X_i)$ 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 2.10. *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 $(0, \infty)^2$. Since Remark 1.2, a straightforward computation shows that*

$$\begin{aligned} \Phi^{(1)}(x_1, x_2) &= \mathbf{Q}_{\mathbf{Y}}(\mathbf{D}_{\mathbf{X}}(x_1, x_2)) = (x_1^2, x_2^2), \\ \Phi^{(2)}(y_1, y_2) &= \mathbf{Q}_{\mathbf{Z}}(\mathbf{D}_{\mathbf{Y}}(y_1, y_2)) = (2y_1, -3y_1 + y_2), \\ \Phi(x_1, x_2) &= \mathbf{Q}_{\mathbf{Z}}(\mathbf{D}_{\mathbf{X}}(x_1, x_2)) = (2x_1^2, -3x_1^2 + x_2^2). \end{aligned}$$

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 Proposition B.7. in Marshall et al. (2011), and it has been omitted.

Proposition 2.11. *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 $\Phi^{(1)} \equiv \mathbf{Q}_{\mathbf{Y}} \circ \mathbf{D}_{\mathbf{X}}$ and $\Phi^{(2)} \equiv \mathbf{Q}_{\mathbf{Z}} \circ \mathbf{D}_{\mathbf{Y}}$ 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 (for a complete review about copulas, see Nelsen, 1999). 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 marginals distributions.

Theorem 3.1. *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

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

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

From Proposition 2.11 and Theorem 3.1 we obtain the following Corollary.

Corollary 3.2. *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 3.1 $\mathbf{Y} \leq_{mct} \mathbf{W}$ 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 2.11. \square

We would like to point out that Theorem 3.1 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 \leq_{mct} via the univariate comparison of the marginal distributions. Next we describe some situations where previous results 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 2.6), 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 3.1, 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 2.6 $\mathbf{X}' =_{mct} \mathbf{X}$ holds. Then, using Proposition 2.11 $\mathbf{X}' \leq_{mct} \mathbf{Y}$ also holds.

3.2 Multivariate distributions with IFR [DFR] marginals

Next corollary provides a situation where we can apply previous ideas.

Corollary 3.3. *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 $a_i > 0$, $b \in \mathbb{R}$ and $E \sim \exp(1)$, then $\mathbf{X} \leq_{mct} [\geq_{mct}] \mathbf{Y}$.*

Proof. The proof follows easily from Theorem 3.1 and the univariate characterization of the IFR [DFR] aging property given in (1). \square

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 3.4. Other examples can be taken from the literature. For instance, let \mathbf{X} be a Weibull distributed bivariate random vector, where the survival copula is given by a Gumbel copula, that is,

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

where $0 < \theta < 1$, and the margins are Weibull distributed like in the previous example. This bivariate distribution can be found in Lu and Bhattacharyya (1990). Now, if we consider \mathbf{Y} a Gumbel bivariate exponential distribution, i.e., a random vector with Gumbel survival copula and exponential margins (Gumbel, 1960), 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 Figure 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.

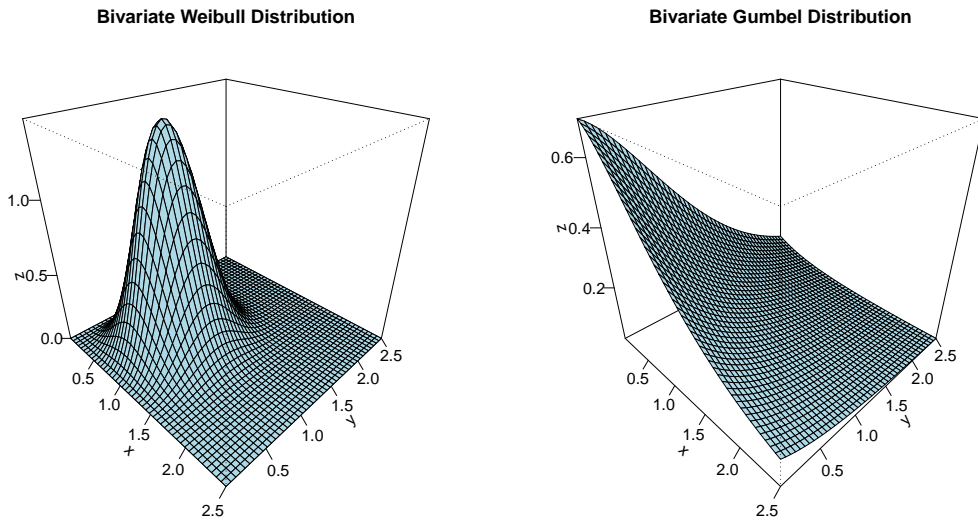


Figure 1: Joint density function of Weibull and Gumbel bivariate distributions.

Example 3.5. Another example is the following. Let us consider a random vector \mathbf{X} with a Clayton copula which is given by

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

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 and Pareto distributed margins (see Balakrishnan and Lai, 2009). It is known that the Pareto distribution is DFR. Hence, again from Corollary 3.3, we have $\mathbf{X} \leq_{mct} \mathbf{Y}$.

In Figure 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.

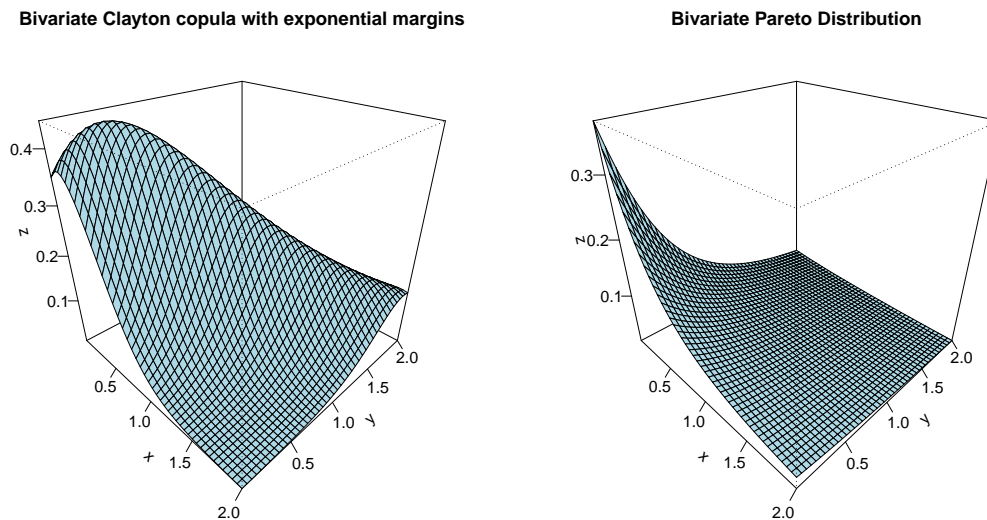


Figure 2: Joint density function of the Clayton copula with exponential margins and Pareto bivariate distributions.

3.3 Relative skewness for ordered data

The model of a random vector with ordered components arises in 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 nohomogeneous Poisson processes can be introduced as record values of a proper sequence of random variables, which is another typical example of ordered data. Given the similarity of several results for order statistics and record values Kamps (1995a) introduces 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) and (1995b).

Definition 3.6. 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** to 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* to 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 $F^{-1}(0) \leq x_1 \leq \dots \leq x_n \leq F^{-1}(1)$.

For GOS's, we have that two random vectors of GOS's with the same set of parameters and possibly based on different distributions have the same copula (see Belzunce et al., 2008).

Let us see now several 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 $X_{i:n} =_{st} X_{(i,n,0,1)}$.

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 GOS's 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 case of order statistics under multivariate imperfect repair (see Shaked and Shanthikumar, 1986).

Next, we show a property dealing with the comparison of relative skewness for generalized orders statistics.

Theorem 3.7. *Let $\mathbf{X} = (X_{(1,n,\tilde{m},k)}, \dots, X_{(n,n,\tilde{m},k)})$ and $\mathbf{Y} = (Y_{(1,n,\tilde{m},k)}, \dots, Y_{(n,n,\tilde{m},k)})$ be the random vectors of generalized order statistics based on distributions F and G from random variables X and Y , respectively. If*

$$X \leq_c Y,$$

then

$$\mathbf{X} \leq_{mct} \mathbf{Y}.$$

Proof. Since the vectors of generalized order statistics have the same copula, the result follows easily if $X_{(r,n,\tilde{m},k)} \leq_c (Y_{(r,n,\tilde{m},k)})$ for all $r = 1, \dots, n$. 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 (see also Belzunce et al., 2005). □

4 On a new definition of multivariate IFR [DFR] distributions

Among other interesting applications, Stochastic Orders are related to the study of aging (see, for example, Belzunce and Shaked, 2008a, and, 2008b). In particular, in all these well known

aging criteria we are comparing the relative performance of a unit under consideration with that of an exponential distribution, -it is worth to recall that a unit with an exponential life distribution suffers no aging-. In particular, as we have pointed out in (1), the well known IFR and DFR aging notions are characterized via the comparison in the univariate convex order sense. Roughly speaking, a random variable which represents the lifetime of a component or system is IFR [DFR] if, and only if, it is smaller [larger] than an exponential distribution in the convex order sense.

Next we propose a multivariate aging notion as a direct extension of the univariate characterization given in (1). Note that this new definition will be based on the mct-order and therefore also based on the **multivariate quantile transform**. It is worth to mention that the **multivariate quantile transform** has been already used in literature to define other aging notions, (see Arias-Nicolás et al., 2009 and 2010) and also it is interesting to note that Roy (2002), clearly inspired in the univariate case, proposed another type of multivariate aging notion based on certain comparisons with a multivariate distribution type Gumbel.

For our purpose, we first define the family of multivariate conditional exponential distributions.

Definition 4.1. *Given an n -dimensional random vector \mathbf{Y} , we will say that \mathbf{Y} follows a multivariate conditional exponential distribution with parameter matrix $\mathbf{A} = (a_{ij}) \in \mathbb{M}_{n \times n}$ and constant vector $\mathbf{b} \in \mathbb{R}^n$, denoted by $\mathbf{Y} \sim MCE(\mathbf{A}, \mathbf{b})$, if $\mathbf{Y} =_{st} \mathbf{A}\mathbf{E} + \mathbf{b}$, where \mathbf{A} is lower triangular with diagonal elements strictly positive and $\mathbf{E} = (E_1, \dots, E_n)$ denotes a multivariate random vector with i.i.d. exponential marginal distributions with common parameter $\lambda = 1$.*

Under the previous notation, it is apparent that $\mathbf{E} \sim MCE(\mathbf{I}_{n \times n}, \mathbf{0}_n)$, where $\mathbf{I}_{n \times n}$ and $\mathbf{0}_n$ represent the identity matrix and the null vector, respectively.

Remark 4.2. *Given $\mathbf{Y} \sim MCE(\mathbf{A}, \mathbf{b})$, it is remarkable that all marginal distributions are linear combinations of independent shifted exponential distributions. The name “conditional exponential” for Y comes from the property that all sequential conditional distributions are*

shifted exponential distributions, i.e.,

$$\begin{aligned}
Y_1 &\sim a_{11}E_1 + b_1, \\
(Y_2|Y_1 = y_1) &\sim a_{21}\alpha_1 + a_{22}E_2 + b_2, \\
(Y_3|Y_1 = y_1, Y_2 = y_2) &\sim a_{31}\alpha_1 + a_{32}\alpha_2 + a_{33}E_3 + b_3, \\
&\vdots \\
\left(Y_n \mid \bigcap_{j=1}^{n-1} Y_j = y_j\right) &\sim \sum_{j=1}^{n-1} a_{nj}\alpha_j + a_{nn}E_n + b_n,
\end{aligned} \tag{16}$$

where conditional distributions are well defined. A straightforward computation shows that

$$(\alpha_1, \dots, \alpha_{n-1})^t = \mathbf{A}_{n-1}^{-1}(y_1 - b_1, \dots, y_{n-1} - b_{n-1})^t,$$

where \mathbf{A}_{n-1} is a submatrix formed by deleting the n -th row and n -th column of \mathbf{A} .

Remark 4.3. Let $\mathbf{Y}_1 \sim MCE(\mathbf{A}_1, \mathbf{b}_1)$ and $\mathbf{Y}_2 \sim MCE(\mathbf{A}_2, \mathbf{b}_2)$ be two multivariate conditional exponential distributions. A straightforward computation shows that

$$\mathbf{Y}_2 =_{st} \mathbf{A}_2 \mathbf{A}_1^{-1}(\mathbf{Y}_1 - \mathbf{b}_1) + \mathbf{b}_2.$$

Just taking in account that \mathbf{A}_1^{-1} is a lower triangular matrix with diagonal elements strictly positive and then $\mathbf{A}_2 \mathbf{A}_1^{-1}$ is also a lower triangular matrix with diagonal elements strictly positive, using Proposition 2.8, all multivariate conditional exponential distributions are equal in the mct -order, i.e. $\mathbf{Y}_1 =_{mct} \mathbf{Y}_2$.

From the definition of the multivariate conditional exponential distribution and clearly inspired in (1), we define a new multivariate aging notion based on the \leq_{mct} order.

Definition 4.4. Let \mathbf{X} be an absolutely continuous random vector. Then \mathbf{X} is said to be mct -IFR [mct -DFR] if there exists a particular multivariate conditional exponential distribution, $\mathbf{Y} \sim MCE(\mathbf{A}, \mathbf{b})$, such that $\mathbf{X} \leq_{mct} [\geq_{mct}] \mathbf{Y}$. We will call the random vector \mathbf{Y} an upper bound exponential distribution associated with \mathbf{X} .

Remark 4.5. From Remark 4.3 one could expect that if \mathbf{X} is mct -IFR [mct -DFR], then \mathbf{X} could be ordered in the mct -order sense with any multivariate conditional exponential

distribution. However, due to the lack of transitivity of the multivariate transform convex order, recall Example 2.10, this is not necessarily true. On the other hand, it is also apparent, just changing the constant vector \mathbf{b} , that the upper bound distribution is not unique.

Propositions 4.6, 4.7 and 4.8 present some desirable properties associated with Definition 4.4.

Proposition 4.6. *Let $\mathbf{X} = (X_1, \dots, X_n)$ be an absolutely continuous random vector with independent components. Then \mathbf{X} is mct-IFR [mct-DFR] if, and only if, X_i is IFR [DFR] for all $i = 1, \dots, n$.*

Proof. We just need to prove that $\mathbf{E} \sim MCE(\mathbf{I}_{n \times n}, \mathbf{0}_n)$ is an upper bound exponential distribution associated with \mathbf{X} . Due to the fact that \mathbf{X} and \mathbf{E} share the same independence copula, just using Theorem 3.1 $\mathbf{X} \leq_{mct} [\geq_{mct}] \mathbf{E}$ holds if, and only if, $X_i \leq_c [\geq_c] E_i$ holds, for all $i = 1 \dots, n$. Since (1), the last condition holds if, and only if, X_i 's are IFR [DFR]. \square

Proposition 4.7. *Let $\mathbf{X} = (X_1, \dots, X_n)$ be an absolutely continuous mct-IFR distribution and let $\mathbf{Y} \sim MCE(\mathbf{A}, \mathbf{b})$ be an upper bound exponential distribution associated with \mathbf{X} . Let $\Phi \equiv \mathbf{Q}_{\mathbf{Y}} \circ \mathbf{D}_{\mathbf{X}}$ be the transformation described in (6). If $\Phi(\mathbf{x})$ is increasing for all \mathbf{x} in the support of \mathbf{X} , then the random vector $\mathbf{B}\mathbf{X}$ is also mct-IFR for all lower triangular matrix \mathbf{B} with strictly positive diagonal elements.*

Proof. Since Proposition 2.8, it is apparent that $\mathbf{B}\mathbf{X} =_{mct} \mathbf{X}$. Using the fact that \mathbf{X} is mct-IFR, $\mathbf{B}\mathbf{X} =_{mct} \mathbf{X} \leq_{mct} \mathbf{Y}$ holds. By hypothesis assumption and using Lemma 2.11 the proof is concluded. \square

Proposition 4.8. *Let $\mathbf{X} = (X_1, \dots, X_n)$ be an absolutely continuous mct-DFR distribution, then the random vector $\mathbf{B}\mathbf{X}$ is also mct-DFR for all lower triangular matrix $\mathbf{B} = \{a_{ij}\}$ such that $a_{ij} \geq 0$ with diagonal elements strictly positive.*

Proof. Let $\mathbf{Y} \sim MCE(\mathbf{A}, \mathbf{b})$ be the corresponding upper bound exponential distribution associated with \mathbf{X} . Since Proposition 2.8 it is apparent that $\mathbf{B}\mathbf{X} =_{mct} \mathbf{X}$. Using the fact that \mathbf{X} is mct-DFR, $\mathbf{B}\mathbf{X} =_{mct} \mathbf{X} \geq_{mct} \mathbf{Y}$ holds. Using Remark 1.2, the function $\Phi(\mathbf{x}) = \mathbf{B}\mathbf{x}$

has the form of (6) and trivially is increasing. The proof concludes just using Lemma 2.11. \square

The following result shows that if a random vector \mathbf{X} satisfies the mct-IFR [mct-DFR] aging property, then all sequential conditional distributions associated with \mathbf{X} are univariate IFR [DFR].

Theorem 4.9. *Let \mathbf{X} be a mct-IFR [mct-DFR] distribution. Then,*

$$X_1 \text{ is IFR [DFR]}$$

and the conditional distribution

$$\left(X_i \left| \bigcap_{j=1}^{i-1} \{X_j = Q_{\mathbf{X},j}(u_1, \dots, u_j)\} \right. \right) \text{ is IFR [DFR]}$$

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

Proof. Let $\mathbf{Y} \sim MCE(\mathbf{A}, \mathbf{b})$ be the corresponding upper bound exponential distribution associated with \mathbf{X} . From Proposition 2.4 and using (10)

$$X_1 \leq_c [\geq_c] Y_1,$$

and using (11)

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

for $i = 2, \dots, n$ and for all u_j such that $0 < u_j < 1$, $j = 1, \dots, i - 1$. The proof follows directly just taking in account that all sequential conditional distributions associated with \mathbf{Y} are univariate shifted exponential distributions, see Remark 4.2, and using (1). \square

Just recalling that the definition of the mct-order depends on the **multivariate quantile transform** which, in turn, is based on sequential conditional distributions, it is not surprising that Definition 4.4 is not invariant under permutation of the marginal distributions. Let us consider the following example: let X and Y be two independent DFR univariate random variables. It is apparent from Proposition 4.6 that the random vector $\mathbf{X} = (X, Y)$ is mct-DFR. If we consider now the lower triangular matrix $\mathbf{A} = (a_{ij}) \in \mathbb{M}_{2 \times 2}$, such that $a_{11} =$

$a_{21} = a_{22} = 1$, using Proposition 4.8 we obtain that $\mathbf{A}\mathbf{X} = (X, X + Y)$ is also mct-DFR. In general, however, we can not expect the mct-DFR property for the random vector $(X + Y, X)$ as can be seen using the following reasoning: if we assume that $(X + Y, X)$ is mct-DFR, using Theorem 4.9 the sum $X + Y$ is also DFR. Because we have not considered any restriction for X and Y , our first assumption is in contradiction with the well known fact that the sum of two independent DFR random variables is not necessarily DFR.

The above drawback naturally arises in the multivariate case and several authors have solved it in different ways. For example, it is common to restrict the study to the interesting case of exchangeable distributions, see Spizzichino (2001). Of course, Definition 4.4 can be also considered for all permutations of the marginal variables as a stronger condition. In conclusion, Definition 4.4 is a priori appropriated for random vectors with a natural order on their components, such as the family of generalized order statistics, exchangeable distributions and random vectors where the sequence of failures could determine the aging properties. We encounter situations where some components are “more relevant” than others. For instance, from a medical point of view, it is apparent that the sequence of multiple organ failure {heart, brain, liver, kidney} could have different aging properties than that by {kidney, liver, heart, brain}.

Now, we will provide some examples of classical distributions that satisfies the new multivariate aging property.

Theorem 4.10. *Let $\mathbf{X} \sim N_n(\mu, \Sigma)$ be a multivariate normal distribution. Then, \mathbf{X} is mct-IFR.*

Proof. We just need to prove that $\mathbf{E} \sim MCE(\mathbf{I}_{n \times n}, \mathbf{0}_n)$ is an upper bound exponential distribution associated with \mathbf{X} . Since Example 2.6, we obtain that $\mathbf{X} =_{mct} \mathbf{Z}$, where $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_{n \times n})$. Due to the fact that \mathbf{Z} and \mathbf{E} share the same independence copula, just using Theorem 3.1 $\mathbf{Z} \leq_{mct} [\geq_{mct}] \mathbf{E}$ holds if, and only if, $Z_i \leq_c [\geq_c] E_i$ holds, for all $i = 1, \dots, n$. Since (1), the last condition holds from the well-known fact that univariate normal distributions are IFR. The proof concludes just using Corollary 3.2. \square

Theorem 4.11. *Let X be a random variable with distribution function F , and let \mathbf{X} be*

a random vector of generalized order statistics based on F , $\mathbf{X} = (X_{(1,n,\tilde{m},k)}, \dots, X_{(n,n,\tilde{m},k)})$. Then \mathbf{X} is *mct-IFR* [*mct-DFR*] if, and only if, X is *IFR* [*DFR*].

Proof. Let \mathbf{Y} be a random vector of generalized order statistics based on $E \sim \exp(1)$, with the same parameters than \mathbf{X} , i.e. $\mathbf{Y} = (Y_{(1,n,\tilde{m},k)}, \dots, Y_{(n,n,\tilde{m},k)})$. First we will prove that \mathbf{Y} follows a conditional multivariate exponential distribution. From Theorem 3.10 in Kamps (1995), the normalized spacings of \mathbf{Y} are i.i.d. exponential distributions with parameter $\lambda = 1$, i.e. $\mathbf{D} = (D(1, n, \tilde{m}, k), \dots, D(n, n, \tilde{m}, k)) =_{st} \mathbf{E} \sim MCE(\mathbf{I}_{n \times n}, \mathbf{0}_n)$ where

$$\begin{aligned} D(1, n, \tilde{m}, k) &= \gamma_1 Y_{(1,n,\tilde{m},k)}, \\ D(r, n, \tilde{m}, k) &= \gamma_r (Y_{(r,n,\tilde{m},k)} - Y_{(r-1,n,\tilde{m},k)}), \end{aligned}$$

for $2 \leq r \leq n$. A straightforward computation shows that $\mathbf{D} = \mathbf{A}\mathbf{Y}$, where \mathbf{A} is a lower triangular matrix given by

$$\mathbf{A} = \begin{pmatrix} \gamma_1 & 0 & 0 & \dots & 0 \\ -\gamma_2 & \gamma_2 & 0 & \dots & 0 \\ 0 & -\gamma_3 & \gamma_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \gamma_n \end{pmatrix}.$$

Then just considering the inverse matrix, we obtain that $\mathbf{Y} \sim MCE(\mathbf{A}^{-1}, \mathbf{0}_n)$, where \mathbf{A}^{-1} is given by the lower triangular matrix

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{\gamma_1} & 0 & \dots & 0 \\ \frac{1}{\gamma_1} & \frac{1}{\gamma_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\gamma_1} & \frac{1}{\gamma_2} & \dots & \frac{1}{\gamma_n} \end{pmatrix}.$$

Due to the fact that \mathbf{X} and \mathbf{Y} have the same set of parameters, since Theorem 3.7 $\mathbf{X} \leq_{mct} \mathbf{Y}$ holds if, and only if, $F \leq_c [\geq_c] E$ holds. Since (1), the last condition holds if, and only if, F is *IFR* [*DFR*]. \square

References

- [1] Abtahi, A. and Towhidi, M. (2013). The new unified representation of multivariate. *Statistics* **47**, 126-140.
- [2] Aitchison J. and Brown, J.A.C. (1957). *The Lognormal Distribution*. Cambridge University Press, Cambridge.
- [3] Arellano-Valle, R.B. and Azzalini, A. (2006). On the unification of families of skew-normal distributions. *Scandinavian Journal of Statistics* **33**, 561-574.
- [4] Arias-Nicolás, J. P., Belzunce, F., Núñez-Barrera, O. and Suárez-Llorens, A. (2009). A multivariate IFR notion based on the multivariate dispersive ordering. *Applied Stochastic Models in Business and Industry* **25**, No. 3, 339-358.
- [5] Arias-Nicolás, J.P., Fernández-Ponce, J.M., Luque-Calvo, P. and Suárez-Llorens, A.(2005). Multivariate dispersion order and the notion of copula applied to the multivariate t -distribution. *Probability in the Engineering and Informational Sciences* **19**, 363-375.
- [6] Arias-Nicolás, J. P., Mulero, J., Núñez-Barrera, O. and Suárez-Llorens, A. (2010). New aging properties of the Clayton-Oakes model based on multivariate dispersion. *SORT*. **34**, No. 1, 79-94.
- [7] Arjas, E. and T. Lehtonen (1978). Approximating many server queues by means of single server queues. *Mathematics of Operations Research* **3**, 205-223.
- [8] Arnold, B.C., Castillo, E., Sarabia, J.M. (2006). Families of multivariate distributions involving the Rosenblatt construction. *Journal of the American Statistical Association* **101**, 1652-1662.
- [9] Azzalini, A. (2005) The skew-normal distribution and related multivariate families (with discussion). *Scandinavian Journal of Statistics* **32**, 159-188.
- [10] Azzalini, A. and Capitanò, A. (1999). Statistical applications of multivariate skew-normal distributions. *Journal of the Royal Statistical Society, Ser. B* **61**, 579-602.

- [11] Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83**, 715-726.
- [12] Balakrishnan, N. and Lai, C.D. (2009). *Continuous Bivariate Distributions*, Second Edition. Springer, New York.
- [13] Barlow, R.E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, Inc., New York.
- [14] Belzunce, F. (2013). Multivariate comparisons of ordered data in *Stochastic Ordes in Reliability and Risks*, eds. H. Li and X. Li, Springer, New York.
- [15] Belzunce, F., Ruíz, J.M. and Suárez-Llorens, A. (2008). On multivariate dispersion orderings based on the standard construction. *Statistics & Probability Letters* **78**, 271-281.
- [16] Belzunce, F. and Shaked, M. (2008a). Stochastic orders and aging notions in *Encyclopedia of Statistics in Quality and Reliability*, edited by F. Ruggeri, F. Faltin and R. Kenett, Wiley, London; 1931-1935.
- [17] Belzunce, F. and Shaked M. (2008b). Multivariate stochastic orders and aging in *Encyclopedia of Statistics in Quality and Reliability*, edited by F. Ruggeri, F. Faltin and R. Kenett, Wiley, London; 1224-1227.
- [18] Chandler, K.N. (1952). The distribution and frequency of record values. *Journal of the Royal Statistical Society* **14**, 220-228.
- [19] Fernández-Ponce, J. M. and Suárez-Llorens, A. (2003). A multivariate dispersion ordering based on quantiles more widely separated. *Journal of Multivariate Analysis* **85**, 40-53.
- [20] Ferreira, J.A.T.S. and Steel, M.F.J. (2006). A constructive representation of univariate skewed distributions, *Journal of the American Statistical Association* **101**, 823-829.
- [21] Gumbel, E.J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association* **55**, 698-707.

- [22] Harville, D.A. (1997), Matrix algebra from a statistician's perspective, Springer-Verlag, New York, Inc.
- [23] Kamps, U. (1995a). A concept of generalized order statistics. *Journal of Statistical Planning and Inference* **48**, 1-23.
- [24] Kamps, U. (1995b). *A Concept of Generalized Order Statistics*, B.G. Taubner, Stuttgart.
- [25] Lai, C.D. and Xie, M. (2006). *Stochastic Ageing and Dependence for Reliability*. Springer Series in Statistics, Springer, New York.
- [26] Ley, C. and Paindaveine, D. (2010). Multivariate skewing mechanisms: A unified perspective based on the transformation approach. *Statistics % Probability Letters* **80**, 1685-1694.
- [27] Lindley, D.V. and Singpurwalla, N.D. (1986). Multivariate distribution for the life lengths of a system sharing a common environment. *Journal of Applied Probability* **23**, 418-431.
- [28] Lu, J. and Bhattacharyya, G. K. (1990). Some new constructions of bivariate Weibull models. *Annals of the Institute of Statistical Mathematics* **42**, 543-559.
- [29] Marshall, A. and Olkin, I. (2007). *Life Distributions*. Springer Series in Statistics, Springer, New York.
- [30] Marshall, A., Olkin, I. and Arnold, B. C. (2011). *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics, Springer, New York.
- [31] Nelsen, R.B. (1999). *An Introduction to Copulas*. Lectures Notes in Statistics **139**. Springer-Verlag, New York.
- [32] O'Brien, G.L. (1975). The comparison method for stochastic processes. *The Annals of Probability* **3**, 80-88.
- [33] Roy, D. (2002). *Classification of multivariate life distributions based on partial ordering*, Probability in Engineering and Information Sciences **16**, 129-137.

- [34] Ruschendorf, L. (1981). Stochastically ordered distributions and monotonicity of the OC-function of sequential probability ratio tests. *Mathematische Operationsforschung und Statistik Series Statistics* **12**, 327-338.
- [35] Shaked, M. and Shanthikumar, J.G. (1986). Multivariate imperfect repair. *Operations Research* **34**, 437-448.
- [36] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*. Springer Series in Statistics. Springer, New York.
- [37] 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.
- [38] Spizzichino, F. (2001), Subjective Probability Models for Lifetimes. *Chapman and Hall/CRC*. New York.
- [39] van Zwet, W.R. (1964). *Convex Transformations of Random Variables*. Mathematical Centre Tracts, 7 Mathematisch Centrum, Amsterdam.