Statistical Methods & Applications (2022) 31:925–954
https://doi.org/10.1007/s10260-021-00613-2

ORIGINAL PAPER

Distortion representations of multivariate distributions

Jorge Navarro[1] · **Camilla Calì**[2] · **Maria Longobardi**[2] · **Fabrizio Durante**[3]

Accepted: 10 November 2021 / Published online: 20 January 2022

Abstract

The univariate distorted distributions were introduced in risk theory to represent changes (distortions) in the expected distributions of some risks. Later, they were also applied to represent distributions of order statistics, coherent systems, proportional hazard rate and proportional reversed hazard rate models, etc. In this paper we extend this concept to the multivariate setup. We show that, in some cases, they are a valid alternative to the copula representation, especially when the marginal distributions may not be easily handled. Several examples illustrate the applications of such representations in statistical modeling. They include the study of paired (dependent) ordered data, joint residual lifetimes, order statistics and coherent systems

Keywords Multivariate distributions · Copulas · Residual lifetimes · Order statistics · Coherent systems

Mathematics Subject Classification 62H05 · 62N05

✉ Maria Longobardi
malongob@unina.it

Jorge Navarro
jorgenav@um.es

Camilla Calì
camilla.cali@unina.it

Fabrizio Durante
fabrizio.durante@unisalento.it

1 Departamento Estadística e Investigación Operativa, Universidad de Murcia Campus de Espinardo, 30100 Murcia, Spain

2 Dipartimento di Biologia. Università di Napoli Federico II Via Cintia, I-80126 Napoli, Italy

3 Dipartimento di Scienze dell'Economia, Università del Salento Lecce, Lecce, Italy

1 Introduction

Distorted distributions were introduced in the theory of choice under risk (see Liu et al. 2021; López-Díaz et al. 2012; Sordo et al. 2016; Wang 1996; Yaari 1987) to model the changes in the distribution of the risk variable under study. The distorted distribution is defined as $F_d(t) = d(F(t))$, where F is the original distribution function and $d : [0, 1] \to [0, 1]$ is a distortion function, that is increasing, continuous and satisfies $d(0) = 0$ and $d(1) = 1$. However they can be applied in several contexts. For example, in reliability theory and survival analysis, they can be used to represent the distributions of coherent systems with identically distributed (ID) components (see, e.g., Navarro et al. 2013). These include both the cases of independent identically distributed (IID) and dependent identically distributed (DID) component lifetimes. In particular, they are also useful to represent the distributions of order statistics (i.e., the ordered data obtained from a sample) since they have the same distributions as k-out-of-n systems. Finally, they were also exploited to define classes of prior distributions in Bayesian statistics (see Arias-Nicolás et al. 2016) and to represent conditional distributions (see Navarro and Durante 2017; Navarro et al. 2017; Navarro and Sordo 2018).

The distorted distributions were extended in Navarro et al. (2016) to represent univariate distributions as distortions of $n \geq 2$ distribution functions. These representations were applied to study the distribution of a single coherent system formed of n components with different distributions. They can also be used to represent ordered data from different populations (or in presence of outliers) and to perform stochastic comparisons (see Navarro 2018; Navarro et al. 2016; Sordo and Suárez-Llorens 2011).

Several multivariate distortions have been proposed as well with the purpose of changing (shift) the distribution function of a given random vector $(X_1, \ldots, X_n)$. For example, the distortion of the first kind proposed in Valdez and Xiao (Valdez and Xiao 2011) maintains the copula and distorts the marginals (see Sect. 2.1). Alternatively, the distortion of the third kind proposed there maintains the marginals and replaces the copula by a distorted copula (see also Durante et al. 2010; Durante and Sempi 2016; Morillas 2005). Other authors propose alternative representations for a given multivariate distribution F to the classical ones based on copulas (see Durante and Sempi 2016; Nelsen 2006). For example, Klüppelberg and Resnick (Klüppelberg and Resnick 2008) proposed to use the Pareto-copula C_P to represent F (see Remark 2).

In this paper we introduce the new concept of *multivariate distorted distribution* (MDD) extending the univariate concept given above. The MDDs provide alternative representations for a given multivariate distribution that can be represented as *distortions* of univariate distributions. These representations are similar to the classical copula representations. Actually, the copula representations are included in this general model. The main difference is that the MDD representation may not be built from the univariate marginals of the considered model, but from any set of univariate distributions. This fact provides a wide flexibility and allows us to obtain different useful representations. For example, these representations can be used (instead of the copula representation) when the expressions for the marginals are not available or are too complex. Furthermore, two random vectors with different marginals could be compared by considering the MDD associated with the same set of univariate distributions. The MDD representations could also be used to built new multivariate models and to inspire novel two-stage inference procedures. Several examples of such representations are provided in Sect. 3. We must note that the purpose of these representations is not to change (distort) the original distribution F, but to provide alternative representations for it.

We provide several examples were these representations are useful. In the first example, we provide a representation for the residual lifetimes of the working components in a system at a given time $t > 0$ extending the results obtained in Longobardi and Pellerey (2019) and Navarro et al. (2017). In the second one, we study ordered paired data obtaining a representation for the joint distribution of the smallest and the largest data. This representation can be used to predict the largest order statistic from the smallest order statistic. This procedure can be applied, for instance, to study diseases of paired organs (eyes, kidneys, lungs, etc.). The representation can be extended to the general case of ordered data (order statistics) from a sample of dependent or independent identically distributed random variables. Finally, we show how they can also be applied in Reliability Theory to represent the joint distribution of two different coherent systems based on the same components. In particular, this representation can be used to compute the system reliability and the expected system residual lifetime at the time of the first component failure.

The rest of the paper is organized as follows. In the following Sect. 2 we define the multivariate distorted distributions obtaining their main properties. The relevant examples are placed in Sect. 3. Some illustrations of simulated ordered paired data sets are given in Sect. 4. The conclusions and open tasks for future research projects are in Sect. 5. The proofs and auxiliary technical results are collected in the Appendix.

2 Multivariate distorted distributions

Throughout the paper we use the terms ‘increasing’ and ‘decreasing’ in a wide sense, that is, they mean ‘non-decreasing’ and ‘non-increasing’, respectively. For example, a function $D : [0, 1]^n \to [0, 1]$ is increasing if $D(u_1, \ldots, u_n) \leq D(v_1, \ldots, v_n)$ whenever $0 \leq u_i \leq v_i \leq 1$ for all $i = 1, \ldots, n$.

2.1 Definition

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector over a probability space $(\Omega, \mathcal{S}, \Pr)$. Then the joint distribution function F of $\mathbf{X}$ is defined, for every $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$, by

$$F(x_1, \ldots, x_n) = \Pr(X_1 \leq x_1, \ldots, X_n \leq x_n).$$

The (marginal) distribution of X_i is given by

$$F_i(x_i) = \Pr(X_i \le x_i) = F(+\infty, \ldots, +\infty, x_i, +\infty, \ldots, +\infty)$$

for $i = 1, \ldots, n$. It is well known that the probability in the n-dimensional rectangle (box) determined by the points $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ can be computed from F as

$$\Pr(x_1 < X_1 \le y_1, \ldots, x_n < X_n \le y_n) = \triangle_{(x_1,\ldots,x_n)}^{(y_1,\ldots,y_n)} F,$$

where $x_i < y_i$ for every $i \in \{1, \ldots, n\}$,

$$\triangle_{(x_1,\ldots,x_n)}^{(y_1,\ldots,y_n)} F := \sum_{z_i \in \{x_i, y_i\}} (-1)^{\mathbf{1}(z_1,\ldots,z_n)} F(z_1, \ldots, z_n),$$

while $\mathbf{1}(z_1, \ldots, z_n) = \sum_{i=1}^n 1(z_i = x_i)$ and $1(A) = 1$ (respectively, 0) if A is true (respectively, false).

From Sklar's theorem (see, e.g., Sklar 1959 or p. 42 in Durante and Sempi 2016), we know that

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)) \tag{1}$$

for every $(x_1, \ldots, x_n) \in \mathbb{R}^n$, where $F_1, \ldots, F_n$ are the marginal distributions and C is a *copula* function. Moreover, if all these marginal distributions are continuous, then C is unique. For the basic properties of copulas we refer the reader to Durante and Sempi (2016); Joe (2014); Mai and Scherer (2017); Nelsen (2006) and references therein. The set of all copulas of dimension n will be denoted by $\mathcal{C}_n$. Any copula function can be extended to $\mathbb{R}^n$ to be a continuous multivariate distribution function with uniform marginals over the interval (0, 1).

Now, we introduce a slightly more general concept that can be used to represent a multivariate distribution function F in terms of arbitrary marginals.

Definition 1 A continuous function $D : [0,1]^n \to [0,1]$ is called *n–dimensional distortion function* if it satisfies the following properties:

(i) $D(u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_n) = 0$ for all $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n \in [0,1]$.
(ii) $D(1, \ldots, 1) = 1$.
(iii) D is n–increasing, i.e. for all $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ in $[0,1]^n$ with $x_i \le y_i$ for every $i \in \{1, \ldots, n\}$, it holds $\triangle_{\mathbf{x}}^{\mathbf{y}} D \ge 0$.

In particular, $D(0, \ldots, 0) = 0$. Moreover, in view of (Nelsen 2006, Lemma 2.1.5), it follows that D is increasing in each variable. As for copulas, D can be extended to $\mathbb{R}^n$ to be a continuous multivariate distribution function with support contained in $[0,1]^n$. The set of all distortion functions of dimension n will be represented by $\mathcal{D}_n$. Obviously, $\mathcal{C}_n \subseteq \mathcal{D}_n$.

Notice that the distortion D is also called *pseudo-copula* in Fermanian and Wegkamp (2012), where the concept has been mainly motivated in the context of conditional time-varying models.

Definition 2 A multivariate distribution function F is said to be a *multivariate distorted distribution* (MDD) of the univariate distribution functions $G_1, \ldots, G_n$ if there exists a distortion $D \in \mathcal{D}_n$ such that, for every $(x_1, \ldots, x_n) \in \mathbb{R}^n$, it holds

$$F(x_1, \ldots, x_n) = D(G_1(x_1), \ldots, G_n(x_n)) \tag{2}$$

We write $F \equiv MDD(G_1, \ldots, G_n)$, when F is a MDD of $G_1, \ldots, G_n$.

For instance, any multivariate distribution function F is a MDD of its own univariate marginals via the copula representation (1). However, for some other choices of $G_1, \ldots, G_n$, F may be expressed as a MDD of $G_1, \ldots, G_n$.

Remark 1 An example of MDD representation has been described in Valdez and Xiao (2011) under the name of *copula distortion of the first kind*. Given a multivariate distribution function F expressed in terms of its copula representation (1), we can change the original multivariate distribution F to the distorted one

$$F_{d_1,\ldots,d_n}(x_1, \ldots, x_n) := C(d_1(F_1(x_1)), \ldots, d_n(F_n(x_n))),$$

for given univariate distortion functions $d_1, \ldots, d_n$. In order to compare F and $F_{d_1,\ldots,d_n}$ we could represent this latter function as a MDD of $(F_1, \ldots, F_n)$ with

$$F_{d_1,\ldots,d_n}(x_1, \ldots, x_n) = D(F_1(x_1), \ldots, F_n(x_n)),$$

where

$$D(u_1, \ldots, u_n) := C(d_1(u_1), \ldots, d_n(u_n))$$

for every $(u_1, \ldots, u_n) \in [0, 1]^n$. Obviously, if $d_1 = \ldots = d_n$ is the identity function, then D is a copula.

Remark 2 While the copula representation has provided to be very useful, in the literature, different specifications have been used to represent a multivariate distribution function F (see, e.g., Embrechts 2009). For instance, if G is a continuous univariate distribution function and C is a copula, we can define the function C_G via

$$C_G(v_1, \ldots, v_n) = C(G(v_1), \ldots, G(v_n))$$

for every $(v_1, \ldots, v_n) \in \mathbb{R}^n$. Such a C_G need not be a copula since its common marginal distribution is G (it is only a copula when G is the standard uniform distribution). Moreover, it can be used to obtain the following representation for a distribution function F expressed as in (1),

$$F(x_1, \ldots, x_n) = C_G(G^{-1}(F_1(x_1)), \ldots, G^{-1}(F_n(x_n))). \tag{3}$$

For instance, if G is the standard Pareto distribution, then C_G is the Pareto copula proposed in Klüppelberg and Resnick (2008). Similar definitions can be proposed for other relevant distributions, such as Gaussian (see, e.g., Joe 2014). However, note that C_G is not a distortion function, since it is not supported on $[0, 1]^n$ and, hence, (3) does not provide a MDD representation for F.

The main properties of MDD representations are given in the next subsections. Several useful examples are included in Sect. 3.

2.2 Main properties

According to Sklar's theorem (Sklar 1959), any multivariate distribution function can be expressed in terms of its univariate marginal distributions via the representation (1). Invoking the same type of arguments, in Fermanian and Wegkamp (2012, Theorem 1) it is shown that, under some mild conditions, any multivariate distribution function can be expressed in terms of arbitrary univariate marginal distributions $G_1, \ldots, G_n$ via the representation (2).

In the following result, we state a similar Sklar-type theorem for continuous distribution functions.

Proposition 1 *Let* $(X_1, \ldots, X_n)$ *be a random vector with joint continuous distribution function* F*. Let* $G_1, \ldots, G_n$ *be arbitrary continuous distribution functions and let us assume that* G_i *is strictly increasing in the support of* X_i *for* $i = 1, \ldots, n$*. Then there exists a unique distortion* $D \in \mathcal{D}_n$ *such that (2) holds.*

Proof See the Appendix. □

Remark 3 From the proof of Proposition 1, it follows that D is the multivariate distribution function of $(G_1(X_1), \ldots, G_n(X_n))$, so that each $V_i = G_i(X_i)$ is a componentwise increasing transformation of X_i for $i = 1, \ldots, n$. Thus, for any measure of concordance κ (as Kendall's tau or Spearman's rho, see, e.g. Taylor 2016), $\kappa(V_1, \ldots, V_n) = \kappa(X_1, \ldots, X_n)$. In essence, D contains all the information about the (rank-invariant) dependence structure of F.

Remark 4 From Proposition 1, for instance, if $X_1, \ldots, X_n$ are nonnegative random variables, then we can choose to represent the joint distribution function F as a MDD of a common standard exponential distribution with $G(t) = 1 - e^{-t}$ for every $t \geq 0$.

For a continuous distribution function F, it is possible to link the copula representation and any MDD representation via the following result.

Proposition 2 *Let* $(X_1, \ldots, X_n)$ *be a random vector with joint continuous distribution function* F*. Let* $G_1, \ldots, G_n$ *be arbitrary continuous distribution functions. Suppose that* $F \equiv MDD(G_1, \ldots, G_n)$ *with distortion* D*. Then, for every* $(u_1, \ldots, u_n) \in [0, 1]^n$,

$$D(u_1, \ldots, u_n) = C(F_1(G_1^{-1}(u_1)), \ldots, F_n(G_n^{-1}(u_n))),$$

where G_i^{-1} is the quasi-inverse of G_i for $i = 1, \ldots, n$.

Proof See the Appendix. □

The converse of Proposition 1 can be stated as follows.

Proposition 3 *If $D \in \mathcal{D}_n$, then the function F defined by (2) is a multivariate distribution function for all univariate distribution functions $G_1, \ldots, G_n$.*

Proof See the Appendix. □

The previous result gives a useful recipe to built new multivariate probability models. In fact, we can obtain multivariate distributions by changing the univariate distribution functions $G_1, \ldots, G_n$ for a fixed D, or by changing $D \in \mathcal{D}_n$ for fixed $G_1, \ldots, G_n$ (as in copula representations).

Remark 5 Notice that the distortion (or aggregation) functions considered in Navarro et al. (2016) to get univariate distribution functions from n univariate distribution functions do not necessarily belong to $\mathcal{D}_n$ (i.e., in many cases, they are not multivariate distortion functions). For example, the aggregation function $Q(u_1, \ldots, u_n) = (u_1 + \ldots + u_n)/n$ is not a multivariate distortion function since $Q(0, u_2, \ldots, u_n) = (u_2 + \ldots + u_n)/n > 0$ for all $u_2, \ldots, u_n \in (0, 1]$. On the other hand, $\mathcal{D}_2$ is included in the class of 2–increasing aggregation functions considered in Durante et al. (2007).

In the next proposition we show that, if (2) holds, then a similar representation holds for the joint survival (or reliability) function

$$\bar{F}(x_1, \ldots, x_n) = \Pr(X_1 > x_1, \ldots, X_n > x_n).$$

Proposition 4 *Let $(X_1, \ldots, X_n)$ be a random vector with distribution function F. If (2) holds for $G_1, \ldots, G_n$ and $D \in \mathcal{D}_n$, then the joint survival function of $(X_1, \ldots, X_n)$ can be written as*

$$\bar{F}(x_1, \ldots, x_n) = \hat{D}(\bar{G}_1(x_1), \ldots, \bar{G}_n(x_n)) \tag{4}$$

for all $x_1, \ldots, x_n$, where $\bar{G}_i = 1 - G_i$ is the survival function associated to G_i for $i = 1, \ldots, n$ and $\hat{D} \in \mathcal{D}_n$.

The proof is immediate since the probability $\Pr(X_1 > x_1, \ldots, X_n > x_n)$ can be obtained as the probability in the n-dimensional rectangle determined by $(x_1, \ldots, x_n)$ and $(+\infty, \ldots, +\infty)$. The distortion function $\hat{D}$ determined by this formula can be called *dual (or survival) distortion function* as in the univariate case. Note that $\hat{D}$ is determined by D (and vice versa). In the univariate case we have $\hat{D}(u) = 1 - D(1 - u)$. However, when $n > 1$,

$$\hat{D}(u_1, \ldots, u_n) \neq 1 - D(1 - u_1, \ldots, 1 - u_n).$$

The formula to get $\hat{D}$ from D is similar to the expression for the survival copula $\hat{C}$ in term of the distributional copula C. Thus, if $n = 2$, we have

$$\begin{aligned}\bar{F}(x_1,x_2) &= \Pr(X_1 > x_1, X_2 > x_2)\\ &= 1 - F(x_1,+\infty) - F(+\infty,x_2) + F(x_1,x_2)\\ &= 1 - D(G_1(x_1),1) - D(1,G_2(x_2)) + D(G_1(x_1),G(x_2))\\ &= 1 - D(1-\bar{G}_1(x_1),1) - D(1,1-\bar{G}_2(x_2)) + D(1-\bar{G}_1(x_1),1-\bar{G}(x_2))\\ &= \hat{D}(\bar{G}_1(x_1),\bar{G}_2(x_2)),\end{aligned}$$

where $\hat{D}(u_1,u_2) = 1 - D(1-u_1,1) - D(1,1-u_2) + D(1-u_1,1-u_2)$ for all $u_1,u_2 \in [0,1]$. Note that if $D \in \mathcal{D}_2$ is the distribution function of the random vector (U_1,U_2), then $\hat{D}$ is the distribution function of $(1-U_1,1-U_2)$ whose support is included in $[0,1]^2$ as well. So $\hat{D} \in \mathcal{D}_2$. Representations (2) and (4) are equivalent but, depending on the application, it could be better to use (4) instead of (2) (or vice versa). Some examples are given in the next section.

2.3 Marginal distributions

A relevant property of the MDD representation $F \equiv MDD(G_1,\ldots,G_n)$ is that all the multivariate marginal distributions of F are also MDD of suitable univariate distributions selected from $G_1,\ldots,G_n$.

In fact, let $(X_1,\ldots,X_n)$ be a random vector with distribution function F. Let $F_{1,\ldots,m}$ be the distribution function of $(X_1,\ldots,X_m)$ for $1 \le m \le n$. The expressions for the other marginals can be obtained in a similar way. The following result holds.

Proposition 5 *If $F \equiv MDD(G_1,\ldots,G_n)$ and $1 \le m \le n$, then the joint distribution function $F_{1,\ldots,m}$ can be written as*

$$F_{1,\ldots,m}(x_1,\ldots,x_m) = D_{1,\ldots,m}(G_1(x_1),\ldots,G_m(x_m)) \tag{5}$$

for all $(x_1,\ldots,x_m) \in \mathbb{R}^m$, where $D_{1,\ldots,m}(u_1,\ldots,u_m) := D(u_1,\ldots,u_m,1,\ldots,1)$ for all $(u_1,\ldots,u_m) \in [0,1]^m$ and $D_{1,\ldots,m} \in \mathcal{D}_m$.

Proof See the Appendix. □

In particular, the ith marginal distribution function of X_i can be written as

$$F_i(x_i) = D(1,\ldots,1,G_i(x_i),1,\ldots,1) = D_i(G_i(x_i)) \tag{6}$$

for all $x_i \in \mathbb{R}$, where $D_i(u) := D(1,\ldots,1,u,1,\ldots,1)$ and the value u is placed at the ith position. Clearly, we have $G_i = F_i$ for a fixed $i \in \{1,\ldots,n\}$ when $D_i(u) = u$ for all $u \in [0,1]$, that is, when the ith marginal of D is a uniform distribution over the interval (0, 1).

2.4 Probability density function and conditional distributions

Let us assume in this subsection that F is absolutely continuous with joint probability density function (PDF) f, where $f = \partial_{1,\ldots,n}F$ a.e., that is, almost everywhere. Here, $\partial_i F$ represents the partial derivative of F with respect to its ith variable, $\partial_{i,j}F := \partial_i\partial_j F$, and so on. Then the joint PDF of a multivariate distorted distribution can be obtained as follows.

Proposition 6 *If $F \equiv MDD(G_1,\ldots,G_n)$ for absolutely continuous distribution functions $G_1,\ldots,G_n$ with PDFs $g_1,\ldots,g_n$, respectively, and a distortion function D that admits continuous mixed derivatives of order n, then the joint PDF f of $(X_1,\ldots,X_n)$ is*

$$f(x_1,\ldots,x_n) = g_1(x_1)\ldots g_n(x_n)\ \partial_{1,\ldots,n}D(G_1(x_1),\ldots,G_n(x_n)). \tag{7}$$

The proof is immediate from (2) and $f = \partial_{1,\ldots,n}F$ (a.e.). Note that, if $D \in \mathcal{D}_n$ and it is absolutely continuous, then $d := \partial_{1,\ldots,n}D$ is a PDF of D.

Remark 6 The PDF representation (7) expresses the PDF f as a product of the PDF d and the marginal PDFs g_i. Thus, as in copula representations (Joe 2005), we can argue that the log-likelihood of such models (when all the PDFs are expressed in a parametric form) can be maximized in a two-stage procedure: the first stage involves maximum likelihood from the univariate distributions, and the second stage involves maximum likelihood of the dependence parameters with the univariate parameters held fixed from the first stage. This is extremely important in practice, as it happens for copula models when the maximum likelihood is computationally difficult to handle. Moreover, now we can choose arbitrary models for $g_1,\ldots,g_n$, from which we may have some partial knowledge to be taken into account.

We can also prove that all the conditional distributions (when they exist) of $F \equiv MDD(G_1,\ldots,G_n)$ can be expressed as a suitable MDD representation. To simplify the notation we just consider the conditional distribution $(X_2|X_1 = x_1)$. The result in such a case can be stated as follows.

Proposition 7 *Let (X_1,X_2) be a random vector with joint distribution function F. If $F \equiv MDD(G_1,G_2)$ for two absolutely continuous distribution functions G_1 and G_2 and a distortion function $D \in \mathcal{D}_2$ that admits continuous mixed derivatives of order 2, then the distribution function $F_{2|1}$ of $(X_2|X_1 = x_1)$ can be written as*

$$F_{2|1}(x_2|x_1) = D_{2|1}(G_2(x_2)|G_1(x_1)) \tag{8}$$

whenever $\lim_{v\to 0^+} \partial_1 D(G_1(x_1),v) = 0$, where $D_{2|1}$ is a distortion function given by

$$D_{2|1}(v|G_1(x_1)) := \frac{\partial_1 D(G_1(x_1),v)}{\partial_1 D(G_1(x_1),1)}$$

for $0<v<1$ and x_1 such that $\partial_1 D(G_1(x_1),1) > 0$.

Proof See the Appendix. □

A similar expression can be obtained from $\hat{D}$ for the conditional survival function by using (4). Note that the PDF of $(X_2|X_1 = x_1)$ can be written as

$$f_{2|1}(x_2|x_1) = g_2(x_2)d_{2|1}(G_2(x_2)|G_1(x_1)), \tag{9}$$

where

$$d_{2|1}(v|G_1(x_1)) := D'_{2|1}(v|G_1(x_1)) = \frac{\partial_{1,2}D(G_1(x_1), v)}{\partial_1 D(G_1(x_1), 1)}$$

for $0<v<1$ (zero elsewhere). Hence, from (9), the regression curve to predict X_2 from X_1, $m_{2|1}(x_1) := E(X_2|X_1 = x_1)$, can be obtained as

$$m_{2|1}(x_1) = \int_{-\infty}^{+\infty} x_2 g_2(x_2)d_{2|1}(G_2(x_2)|G_1(x_1))dx_2.$$

If X_2 is nonnegative (almost surely), then an alternative expression can be obtained from the conditional survival function. Another option to predict X_2 from X_1 is to use the conditional median regression curve $\tilde{m}_{2|1}(x_1) := F_{2|1}^{-1}(0.5|x_1)$ (see Koenker 2005 or Nelsen 2006, p. 217). Note that this function can be computed from (8) as

$$F_{2|1}^{-1}(v|x_1) = G_2^{-1}(D_{2|1}^{-1}(v|G_1(x_1)))$$

for $0<v<1$, provided that the inverse functions of G_2 and $D_{2|1}(v|G_1(x_1))$ can be computed. Moreover, we can obtain α-confidence bands in a similar way (see Koenker (2005) taking $0\leq\beta_1<\beta_2\leq 1$ such that $\beta_2-\beta_1=\alpha$. For example, the centered 50% and 90% quantile-confidence bands for $(X_2|X_1 = x_1)$ are determined, respectively, by $(F_{2|1}^{-1}(0.25|x_1), F_{2|1}^{-1}(0.75|x_1))$ and $(F_{2|1}^{-1}(0.05|x_1), F_{2|1}^{-1}(0.95|x_1))$. Some examples are given in Sect. 4.

2.5 Stochastic comparisons

Let us assume now that two random vectors $\mathbf{X} = (X_1,\ldots,X_n)$ and $\mathbf{Y} = (Y_1,\ldots,Y_n)$ have multivariate distorted distributions with respective distortion functions D_X and D_Y and with the same baseline distribution functions $G_1,\ldots,G_n$. Hence we have the following immediate results for the lower orthant $\leq_{lo}$ and upper orthant $\leq_{uo}$ orders. For the definitions and main properties of these stochastic orders see, e.g., Shaked and Shanthikumar (2007), pages 308–314.

Proposition 8 *Let $\mathbf{X}$ and $\mathbf{Y}$ have MDD of $G_1,\ldots,G_n$ with respective distortion functions D_X and D_Y.*

(i) If $D_X \geq D_Y$, then $\mathbf{X}\leq_{lo}\mathbf{Y}$.
(ii) If $\hat{D}_X \leq \hat{D}_Y$, then $\mathbf{X}\leq_{uo}\mathbf{Y}$.

Analogously, if they have the same distortion, we get the following results.

Proposition 9 *If $\mathbf{X}$ and $\mathbf{Y}$ have MDD with the same distortion function D from $G_1,\ldots,G_n$ and $H_1,\ldots,H_n$, respectively, and $G_i\geq H_i$ holds for $i=1,\ldots,n$, then $\mathbf{X}\leq_{lo}\mathbf{Y}$ and $\mathbf{X}\leq_{uo}\mathbf{Y}$.*

The proof is immediate from (2) and (4) by noting that, if they have the same distortion function D, then they have the same dual distortion function $\hat{D}$ as well. Note that we can combine both propositions to compare MDD with different distortions and different univariate baseline distributions.

3 Illustrative examples

The purpose of this section is to show some examples where the MDD representations can be useful.

3.1 Joint residual lifetimes

In this section, we assume that $X_1, \ldots, X_n$ represent the lifetimes of n components in an engineering or biological system. So we assume that they are nonnegative almost surely.

For $i = 1, \ldots, n$, we denote by $(X_i - t | X_i > t)$ the univariate residual lifetimes at time $t > 0$ whose survival functions are given by

$$\bar{F}_{i,t}(x) := \Pr(X_i - t > x | X_i > t) = \frac{\bar{F}_i(t + x)}{\bar{F}_i(t)}$$

for every $x \geq 0$, whenever $\bar{F}_i(t) > 0$. The univariate residual lifetimes play a central role in the study of the reliability of a system. For example, the mean residual lifetime (MRL) function $m_i(t) = E(X_i - t | X_i > t)$ is used to define a stochastic order (the MRL order) and two aging classes (the increasing/decreasing MRL classes, denoted as IMRL and DMRL, respectively).

Analogously, for the random vector $\mathbf{X} = (X_1, \ldots, X_n)$ and for $t \geq 0$, we can consider the random vector of the residual lifetimes

$$\mathbf{X}_t = (X_1 - t, \ldots, X_n - t | X_1 > t, \ldots, X_n > t)$$

whose survival function is given, for all $x_1 \geq t, \ldots, x_n \geq t$, by

$$\bar{F}_t(x_1, \ldots, x_n) := \Pr(X_1 > x_1, \ldots, X_n > x_n | X_1 > t, \ldots, X_n > t),$$

where it is assumed that $\Pr(X_1 > t, \ldots, X_n > t) > 0$. Note that it is natural to consider a common time t for all the components in a system. Here we just consider that, at a time t, all the components are working (e.g. we can be in a plane and to know that all the engines are working after t hours). Let us remark that, when we write $Z = X|Y$ we mean the random variable whose distribution coincides with the conditional distribution function of X given $Y = y$. Some results for the residual lifetime of the system under this assumption were obtained in Navarro (2018) and Navarro and Durante (2017). We will consider other options later.

In the following proposition we prove that, for all $t \geq 0$, F_t admits a MDD representation in terms of $\bar{F}_{1,t}, \ldots, \bar{F}_{n,t}$. The analogous expression for the joint distribution function can be obtained in a similar way.

Proposition 10 *If $\bar{F}(t,\ldots,t)>0$ for some $t\geq 0$, then*

$$\bar{F}_t(x_1,\ldots,x_n)=\hat{D}_t(\bar{F}_{1,t}(x_1),\ldots,\bar{F}_{n,t}(x_n)) \tag{10}$$

for all $x_1,\ldots,x_n\geq t$ and the following dual distortion function

$$\hat{D}_t(u_1,\ldots,u_n):=\frac{\hat{C}(\bar{F}_1(t)u_1,\ldots,\bar{F}_n(t)u_n)}{\hat{C}(\bar{F}_1(t),\ldots,\bar{F}_n(t))},\ u_1,\ldots,u_n\in[0,1], \tag{11}$$

which depends on $\bar{F}_1(t),\ldots,\bar{F}_n(t)$.

Proof See the Appendix. □

Remark 7 Note that the survival functions $\bar{F}_{1,t},\ldots,\bar{F}_{n,t}$ of the marginal residual lifetimes at time t are not the marginal survival functions of the random vector $\mathbf{X}_t$. The ith marginal survival function $\bar{H}_{i,t}$ of $\mathbf{X}_t$ is

$$\begin{aligned}\bar{H}_{i,t}(x)&=\Pr(X_i-t>x|X_1>t,\ldots,X_n>t)\\&=\frac{\Pr(X_1>t,\ldots,X_{i-1}>t,X_i>t+x,X_{i+1}>t,\ldots,X_n>t)}{\Pr(X_1>t,\ldots,X_n>t)}\\&=\frac{\bar{F}(t,\ldots,t,t+x,t,\ldots,t)}{\bar{F}(t,\ldots,t)}\\&=\frac{\hat{C}(\bar{F}_1(t),\ldots,\bar{F}_{i-1}(t),\bar{F}_i(t+x),\bar{F}_{i+1}(t+x)\ldots,\bar{F}_n(t))}{\hat{C}(\bar{F}_1(t),\ldots,\bar{F}_n(t))}.\end{aligned}$$

Hence representation (10) is not a copula representation. To obtain the survival copula $\hat{C}_t$ of $\mathbf{X}_t$ from (10) we need the inverse functions of $\bar{H}_{1,t},\ldots,\bar{H}_{n,t}$ (see, for instance, Durante et al. (2008)). Note that, in many models, it is not possible to get these inverse functions in closed forms. However, representation (10) always holds and it is based on the univariate residual survival functions $\bar{F}_{1,t},\ldots,\bar{F}_{n,t}$. Roughly speaking, in representation (10), the marginal distributions $\bar{F}_{1,t}$ are conditioned to their own past history, while $H_{i,t}$ are conditioned to an event that takes into account the past history of all components of the system. In a dynamic copula model setting, the former approach could be preferred; see, e.g., (Fermanian and Wegkamp 2012).

We can obtain (in a similar way) MDD representations for other residual lifetimes. For example, if we know that at time t, the first $n-1$ components are alive, but the nth component has failed, then the joint survival function $\bar{F}_t^{(n)}$ of the random vector

$$\mathbf{X}_t^{(n,\leq)}:=(X_1-t,\ldots,X_n-t|X_1>t,\ldots,X_{n-1}>t,X_n\leq t) \tag{12}$$

defined for $t>0$ such that $\Pr(X_1>t,\ldots,X_{n-1}>t,X_n\leq t)>0$, can be written as a MDD (see Appendix).

Similar representations can be obtained from (10) for

$$(X_1 - t, \ldots, X_{n-1} - t | X_1 > t, \ldots, X_n > t)$$

(a marginal of $\mathbf{X}_t$) and

$$(X_1 - t, \ldots, X_{n-1} - t | X_1 > t, \ldots, X_{n-1} > t)$$

(the $n - 1$ dimensional case). More interestingly, we can compare these two random vectors with $\mathbf{X}_t^{(n)}$ just by comparing their distortion functions since their representations are based on the same survival functions $\bar{F}_{1,t}, \ldots, \bar{F}_{n-1,t}$. This is not the case if we use copula representations since these random vectors have different marginal distributions when $X_1, \ldots, X_n$ are dependent. Let us see an example.

Example 1 Let (X_1, X_2, X_3) be a random vector of identically distributed marginals with survival function $\bar{F}$ and a Farlie-Gumbel-Morgenstern (FGM) (survival) copula

$$\hat{C}(u_1, u_2, u_3) = u_1 u_2 u_3 [1 + \theta(1 - u_1)(1 - u_2)(1 - u_3)]$$

for $-1 \le \theta \le 1$. Consider the residual lifetimes

$$\begin{aligned}
\mathbf{X}_t &:= (X_1 - t, X_2 - t | X_1 > t, X_2 > t), \\
\mathbf{X}_t^{(3,\le)} &:= (X_1 - t, X_2 - t | X_1 > t, X_2 > t, X_3 \le t), \\
\mathbf{X}_t^{(3,>)} &:= (X_1 - t, X_2 - t | X_1 > t, X_2 > t, X_3 > t).
\end{aligned}$$

If $t > 0$ and $k := \bar{F}(t) > 0$, then the dual distortion functions are given, respectively, by

$$\begin{aligned}
\hat{D}_t(u_1, u_2) &= \frac{\hat{C}(ku_1, ku_2, 1)}{\hat{C}(k, k, 1)} = u_1 u_2 \\
\hat{D}_t^{(3,\le)}(u_1, u_2) &= \frac{\hat{C}(ku_1, ku_2, 1) - \hat{C}(ku_1, ku_2, k)}{\hat{C}(k, k, 1) - \hat{C}(k, k, k)} = u_1 u_2 \frac{1 - \theta k(1 - ku_1)(1 - ku_2)}{1 - \theta k(1 - k)^2}, \\
\hat{D}_t^{(3,>)}(u_1, u_2) &= \frac{\hat{C}(ku_1, ku_2, k)}{\hat{C}(k, k, k)} = u_1 u_2 \frac{1 + \theta(1 - ku_1)(1 - ku_2)(1 - k)}{1 + \theta(1 - k)^3}.
\end{aligned}$$

A straightforward calculation shows that $\hat{D}_t^{(3,>)} \le \hat{D}_t \le \hat{D}_t^{(3,\le)}$ when $\theta \le 0$, and that $\hat{D}_t^{(3,>)} \ge \hat{D}_t \ge \hat{D}_t^{(3,\le)}$ when $\theta \ge 0$. Hence $\mathbf{X}_t^{(3,>)} \le_{uo} \mathbf{X}_t \le_{uo} \mathbf{X}_t^{(3,\le)}$ for every $t > 0$, every $\bar{F}$ and every $\theta \le 0$. The reverse orderings hold when $\theta \ge 0$.

Similar representations for other conditional residual lifetimes can be given as well. For example, an analogous representation can be obtained for the residual lifetime vector $(X_1 - t, \ldots, X_{n-1} - t | X_1 > t, \ldots, X_{n-1} > t, X_n = t_1)$ when $0 < t_1 < t$ by using the techniques used in Navarro and Durante (2017) and Navarro and Sordo (2018). The comparisons of these random vectors can be used to study the effect of the information available at time t in the residual lifetimes of the working components at this time. We can study inactivity times as well by using the procedures introduced in Navarro and Calì (2019) and Navarro et al. (2017).

3.2 Ordered paired data

In this section, we assume that X and Y have a common absolutely continuous distribution function F and an absolutely continuous copula C. Hence the joint distribution function is $F_{X,Y}(x,y) = C(F(x), F(y))$. In some cases, we may also assume that C is permutation symmetric. In this case, (X, Y) is exchangeable (EXC), that is, (X, Y) and (Y, X) have the same joint distribution. Moreover, we assume that we have some information about $L = \min(X, Y)$, and we would like to estimate $U = \max(X, Y)$. In particular, we would like to predict the regression curve $m(t) = E(U|L = t)$ or its conditional survival function $S(x|t) = \Pr(U > x|L = t)$ that can be used to compute the median regression curve and its confidence bands. Note that the target random variable is U (not Y).

For example, in Biostatistics, X and Y may represent disease lifetimes for paired organs (breast, lung, eyes, etc.). We assume that they are observed in a training sample $(X_1, Y_1), \ldots, (X_m, Y_m)$ from the random vector (X, Y). However, for other individuals, we may just know $L = \min(X, Y)$ and we want to estimate $U = \max(X, Y)$. Note that both F and C can be estimated from the training sample by using parametric models or empirical or kernel type estimators (see. e.g., Omelka et al. 2009; Sumarjaya 2017 and references therein).

We want to obtain a MDD representation for the random vector (L, U) in terms of F and C. First, its joint distribution function $G(x, y) = \Pr(L \le x, U \le y)$ can be computed as

$$G(x,y) = \Pr(U \le y) = \Pr(X \le y, Y \le y) = C(F(y), F(y)),$$

when $y \le x$, and as

$$G(x,y) = \Pr(L \le x, U \le y) = \Pr((\{X \le x\} \cup \{Y \le x\}) \cap \{X \le y\} \cap \{Y \le y\}),$$

when $x < y$. Hence, by using the inclusion-exclusion formula, we get

$$\begin{aligned} G(x,y) &= \Pr(X \le x, Y \le y) + \Pr(X \le y, Y \le x) - \Pr(X \le x, Y \le x) \\ &= C(F(x), F(y)) + C(F(y), F(x)) - C(F(x), F(x)) \end{aligned}$$

for $x < y$. Therefore, $G \equiv MDD(F, F)$, i.e.

$$G(x,y) = D(F(x), F(y)) \tag{13}$$

for the following distortion function

$$D(u,v) = \begin{cases} C(v,v) & \text{for} \quad v \le u; \\ C(u,v) + C(v,u) - C(u,u) & \text{for} \quad u < v. \end{cases} \tag{14}$$

Then the marginal distributions of (L, U) can be written as

$$G_1(x) := \Pr(L \le x) = D(F(x), 1) = D_1(F(x))$$

and

$$G_2(y) := \Pr(U \le y) = D(1, F(y)) = D_2(F(y)),$$

where $D_1(u) = D(u, 1) = 2u - C(u, u)$ and $D_2(u) = D(1, u) = C(u, u)$ for all $u \in [0, 1]$ (see (6)). Here, D_1 and D_2 are univariate distortion functions for every copula C. For example, if X and Y are independent, then $D_1(u) = D(u, 1) = 2u - u^2 \neq u$ and $D_2(u) = D(1, u) = u^2 \neq u$ for all $u \in (0, 1)$. We notice that, in representation (13), D is not a copula and the marginals G_1 and G_2 of G do not appear.

From (8) and (13), the distribution function of $(U|L = x)$ can be obtained as

$$G_{2|1}(y|x) = D_{2|1}(F(y)|F(x)) \tag{15}$$

for $y \geq x$, where

$$D_{2|1}(v|F(x)) := \frac{\partial_1 D(F(x), v)}{\partial_1 D(F(x), 1)},$$

$\partial_1 D(u, v) = 0$ for $v < u$ and

$$\partial_1 D(u, v) = \partial_1 C(u, v) + \partial_2 C(v, u) - \partial_1 C(u, u) - \partial_2 C(u, u),$$

for $v > u$. Hence $\lim_{v \to 0^+} \partial_1 D(u, v) = 0$ for all $0 < u < 1$. In particular, in the EXC case, we have $\partial_1 D(u, v) = 2\partial_1 C(u, v) - 2\partial_1 C(u, u)$ and in the IID case $\partial_1 D(u, v) = 2(v - u)$ for $u \leq v \leq 1$. In this last case, we get $D_{2|1}(v|F(x)) = (v - F(x))/\bar{F}(x)$ for $F(x) \leq v \leq 1$.

In the general case, the PDF of $G_{2|1}$ is $g_{2|1}(y|x) = f(y)\ d_{2|1}(F(y)|F(x))$ for $y \geq x$ (while it is equal to 0, elsewhere), where $f = F'$,

$$d_{2|1}(v|F(x)) := D'_{2|1}(v|F(x)) = \frac{\partial_{1,2} D(F(x), v)}{\partial_1 D(F(x), 1)}.$$

Here, $d(u, v) := \partial_{1,2} D(u, v) = c(u, v) + c(v, u)$, for $v \geq u$, and $d(u, v) := 0$, elsewhere, is the PDF of D and $c(u, v) = \partial_{1,2} C(u, v)$ for $0 \leq u, v \leq 1$ (zero elsewhere) is the PDF of C. Hence, the regression curve $m_{2|1}(x) := E(U|L = x)$, can be obtained as

$$m_{2|1}(x) = \int_x^{+\infty} y f(y) d_{2|1}(F(y)|F(x)) dy = \int_x^{+\infty} y f(y) \frac{\partial_{1,2} D(F(x), F(y))}{\partial_1 D(F(x), 1)} dy.$$

Remark 8 In the IID case, we get $d(u, v) = \partial_{1,2} D(u, v) = 2$ for $0 \leq u \leq v \leq 1$ (i.e. a uniform distribution over this triangle). Moreover, we get

$$m_{2|1}(x) = \int_x^{+\infty} z f(z) \frac{2}{2(1 - F(x))} dz = E(X|X > x),$$

that is, $E(U - x|L = x) = E(X - x|X > x)$ which is the MRL of X (or Y). Therefore, the residual lifetime of U from $x = L$ is equal to the residual lifetime of a component when we know that it is alive at time x and $m_{2|1}$ uniquely determines F. In particular, if F is an exponential distribution, then $m_{2|1}(x) = x + E(X)$.

Remark 9 An alternative expression for the regression curve $m_{2|1}$ can be obtained from the survival function of $(U|L=x)$, given by $\bar{G}_{2|1}(y|x) = \hat{D}_{2|1}(\bar{F}(y)|\bar{F}(x))$ for $y \geq x$, where $\hat{D}_{2|1}(v|\bar{F}(x)) := 1 - D_{2|1}(1-v|F(x))$. As $(U-x|L=x) \geq 0$, then

$$m_{2|1}(x) = x + E(U-x|L=x) = x + \int_x^{+\infty} \bar{G}_{2|1}(y|x)dy.$$

For example, in the IID case, we have

$$\hat{D}_{2|1}(v|\bar{F}(x)) = 1 - D_{2|1}(1-v|F(x)) = 1 - \frac{(1-v)-F(x)}{1-F(x)} = \frac{v}{\bar{F}(x)}$$

for $0 \leq v \leq \bar{F}(x)$. Hence, $\bar{G}_{2|1}(y|x) = \hat{D}_{2|1}(\bar{F}(y)|\bar{F}(x)) = \bar{F}(y)/\bar{F}(x)$ for $y \geq x$, that is, $(U|L=x) =_{st} (X|X>x)$ and so $m_{2|1}(x) = E(X|X>x)$ (as above).

In the following, we show an example in the DID case.

Example 2 Consider the random pair (X, Y) where X and Y have a common exponential distribution with mean $\mu = 1$ and a FGM survival copula

$$\hat{C}(u,v) = uv[1+\theta(1-u)(1-v)] \qquad \text{with } -1 \leq \theta \leq 1.$$

Then

$$\hat{D}_{2|1}(v|\bar{F}(x)) = 1 - D_{2|1}(1-v|F(x)) = \frac{v+\theta v(1-v)(1-2\bar{F}(x))}{\bar{F}(x)+\theta\bar{F}(x)(1-\bar{F}(x))(1-2\bar{F}(x))}$$

for $0 \leq v \leq \bar{F}(x)$, and

$$\begin{aligned} m_{2|1}(x) &= x + \int_x^{+\infty} \bar{G}_{2|1}(y|x)dy \\ &= x + \int_x^{+\infty} \frac{\bar{F}(y)+\theta\bar{F}(y)(1-\bar{F}(y))(1-2\bar{F}(x))}{\bar{F}(x)+\theta\bar{F}(x)(1-\bar{F}(x))(1-2\bar{F}(x))}dy \\ &= x + \frac{(1+\theta-2\theta\bar{F}(x))\int_x^{+\infty}\bar{F}(y)dy - \theta(1-2\bar{F}(x))\int_x^{+\infty}\bar{F}^2(y)dy}{\bar{F}(x)+\theta\bar{F}(x)(1-\bar{F}(x))(1-2\bar{F}(x))}, \end{aligned}$$

where $\int_x^{+\infty}\bar{F}(y)dy = \bar{F}(x)E(X-x|X>x)$. Since F is a standard exponential, then

$$m_{2|1}(x) = x + \frac{1+\theta-2.5\theta e^{-x}+\theta e^{-2x}}{1+\theta-3\theta e^{-x}+2\theta e^{-2x}}$$

for $x \geq 0$. In Fig. 1 we plot the conditional survival functions $\bar{G}_{2|1}(y|1)$ of U given $L=1$ and the regression curves to predict U from $L=x$ for different values of the copula parameter θ. In general, the influence of the dependence parameter is small. Moreover, the provided region from the regression curve may serve to have bounds for the predictions of U from $L=x$ when θ is unknown. A similar procedure can be applied to other one-parameter families of copulas. If F is an exponential distribution with mean $\mu > 0$, then $E(U|L=x) = \mu E(U^*|L^*=x/\mu)$ for $x \geq 0$, where $L^* = L/\mu$ and $U^* = U/\mu$ are obtained from standard exponential distributions.

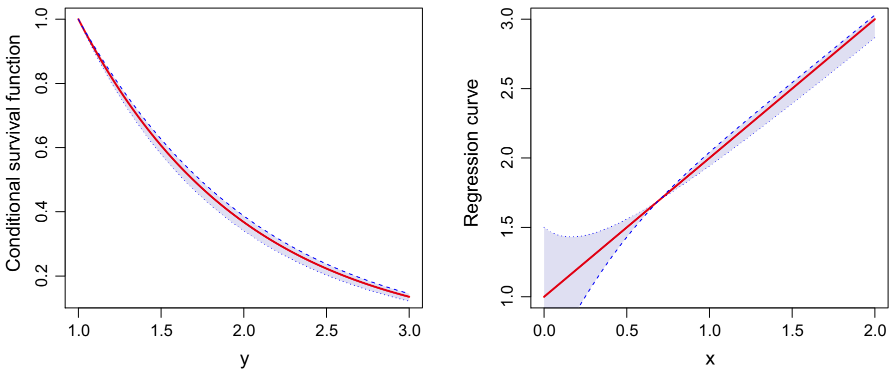

Fig. 1 Conditional survival functions $\bar{G}_{2|1}(y|1)$ (left) and regression curves (right) for (L, U) in Example 2 when $\hat{C}$ is a FGM copula with $\theta \in \{-1, 0, 1\}$ (dotted blue, red, dashed blue). The blue regions illustrate the possible values when $\theta \in [-1, 1]$

Another option to predict U from L is to use the conditional median curve $\tilde{m}_{2|1}(x) := G_{2|1}^{-1}(0.5|x)$, which can be obtained from (15) and the inverse function of $D_{2|1}(u|F(x))$. Furthermore, we can obtain quantile-confidence bands (see Sect. 2.4). In the IID case, we get

$$\tilde{m}_{2|1}(x) = F^{-1}(F(x) + 0.5\bar{F}(x)). \tag{16}$$

In particular, if F is a standard exponential distribution, then

$$\tilde{m}_{2|1}(x) = x - E(X)\ln(0.5) \approx x + 0.6931472E(X) < m_{2|1}(x) = x + E(X) \tag{17}$$

and, hence, $\tilde{m}_{2|1}(x) < m_{2|1}(x) = x + E(X)$. An example with an EXC copula is given in Sect. 4.

Finally, the joint PDF of (L, U) can be obtained from (7) as

$$\begin{aligned} g(x, y) &= f(x)f(y)d(F(x), F(y)) \\ &= f(x)f(y)[c(F(x), F(y)) + c(F(y), F(x))] \end{aligned} \tag{18}$$

for $x \leq y$ while it is equal to 0 elsewhere. For a graphical representation in terms of contour plots see, for instance, Fig. 4.

3.3 Order statistics

Let us consider now the ordered data $X_{1:n}, \ldots, X_{n:n}$ obtained from a sample $(X_1, \ldots, X_n)$ of n possibly dependent identically distributed (DID) random variables with absolutely continuous copula C and marginal distribution F. The usual order statistics obtained from IID random variables are obtained when C is the independence copula. Recent results for conditional distributions in this case can be seen in Ahmadi and Nagaraja (2020). Clearly, the support of $(X_{1:n}, \ldots, X_{n:n})$ is included in the set $S = \{(x_1,\ldots,x_n) : x_1 \leq \ldots \leq x_n\}$. Then we can state the following result.

Proposition 11 *The random vector* $(X_{1:n},\ldots,X_{n:n})$ *has a MDD from a distortion function D and F, that is, its joint distribution function G can be written as*

$$G(x_1,\ldots,x_n) = D(F(x_1),\ldots,F(x_n)) \qquad \text{for every } (x_1,\ldots,x_n) \in S.$$

The explicit expression of D is quite complicated. If $n = 2$ we obtain the distortion function of (L, U) given in the preceding subsection (see (14)). The proof of the preceding proposition for $n = 3$ is given in the Appendix.

If F and C are absolutely continuous with respective PDFs f and c, then the PDF g of G is

$$g(x_1,\ldots,x_n) = f(x_1)\ldots f(x_n) d(F(x_1),\ldots,F(x_n))$$

for all $(x_1,\ldots,x_n) \in S$ (zero elsewhere), where

$$d(u_1,\ldots,u_n) = \sum_{\sigma \in \mathcal{P}_n} c(u_{\sigma(1)},\ldots,u_{\sigma(n)})$$

for $0 \leq u_1 \leq \ldots \leq u_n \leq 1$ (zero elsewhere) and $\mathcal{P}_n$ is the set of permutations of dimension n. The expression for d can be obtained by changing variables from $X_1,\ldots,X_n$ to $X_{1:n},\ldots,X_{n:n}$. If C is EXC, then the expression of d can be simplified to

$$d(u_1,\ldots,u_n) = n!\ c(u_1,\ldots,u_n), \quad \text{for } 0 \leq u_1 \leq \ldots \leq u_n \leq 1.$$

Obviously, the well-known expression for the IID case is obtained with the independence copula, i.e. when $c = 1$.

3.4 Coherent systems

A *system* is a Boolean function $\psi : \{0,1\}^n \to \{0,1\}$ where $\psi(x_1,\ldots,x_n)$ represents the state of the system that is completely determined by the components' states $x_1,\ldots,x_n$. Here $\psi(x_1,\ldots,x_n) = 1$ means that the system works and $\psi(x_1,\ldots,x_n) = 0$ that it has failed. A system is *semi-coherent* if ψ is increasing, $\psi(0,\ldots,0) = 0$ and $\psi(1,\ldots,1) = 1$. It is *coherent* if ψ is increasing and none of its components is irrelevant, i.e., for every $i = 1,\ldots,n$,

$$\psi(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n) < \psi(x_1,\ldots,x_{i-1},1,x_{i+1},\ldots,x_n)$$

for at least one $(x_1,\ldots,x_n) \in \{0,1\}^{n-1}$. In particular, $\psi(0,\ldots,0) = 0$ and $\psi(1,\ldots,1) = 1$. For the basic properties of systems we refer the reader to the classic monograph Barlow and Proschan (1975). In particular, we can see that $\psi(x_1,\ldots,x_n) = \min_{i=1,\ldots,s} \max_{j \in \mathcal{C}_i} x_j$, where $\mathcal{C}_1,\ldots,\mathcal{C}_s$ are the minimal cut sets of the system. A set $A \subseteq \{1,\ldots,n\}$ is a *cut set* of ψ if $\psi(x_1,\ldots,x_n) = 0$ when $x_j = 0$ for all $i \in A$. A cut set is *minimal* if it does not contain other cut sets. The preceding expression can be used to extend ψ to $\mathbb{R}^n$, so that the random system lifetime T can be obtained as

$$T = \psi(X_1,\ldots,X_n) = \min_{i=1,\ldots,s} \max_{j\in\mathcal{C}_i} X_j,$$

where $X_1,\ldots,X_n$ represent the random component lifetimes. Note that the random variable $X^{\mathcal{C}_i} := \max_{j\in\mathcal{C}_i} X_j$ represents the lifetime of the parallel system formed with the components belonging to the set $\mathcal{C}_i$.

It is well known (see, e.g., Navarro 2018 and Navarro et al. 2016) that the distribution function of a coherent (or semi-coherent) system T can be obtained as a distortion of the components' distribution functions. The joint distribution of two semi-coherent systems with some shared components was studied in Navarro and Balakrishnan (2010) and Navarro et al. (2010). Connections between dependence (copula) properties and ordering properties of systems were studied in Navarro et al. (2020).

In this section, we consider two semi-coherent systems based on the same components with a common marginal distribution function F. Then we prove that the joint distribution of these two systems admits a MDD representation in terms of F. The result can be stated as follows. The proof is similar to that of Lemma 3.1 in Navarro and Balakrishnan (2010), that is stated only for systems with IID components.

Proposition 12 *Let $(X_1,\ldots,X_n)$ be a vector of continuous lifetimes with a common marginal distribution function F and copula C. Let $T = \psi(X_1,\ldots,X_n)$ and $T^* = \psi^*(X_1,\ldots,X_n)$ be the lifetimes of two semi-coherent systems. Then the joint distribution function $\mathbf{F}$ of (T,T^*) can be written as*

$$\mathbf{F}(x,y) = D(F(x),F(y)) \tag{19}$$

for all $x\geq 0$, $y\geq 0$, where $D\in\mathcal{D}_2$.

Proof See the Appendix. □

In the proof we can see that the distortion function $D(u, v)$ has different expressions for $u\leq v$ and for $u > v$. Moreover, note that $\mathbf{F}$ can have a singular part since the systems can fail at the same time (even if the joint distribution of the component lifetimes $(X_1,\ldots,X_n)$ is absolutely continuous). Similar results can be stated for the joint distribution and the joint survival functions of m coherent (or semi-coherent) systems based on the same components.

In practice, the above result will be typically applied to study the joint distribution of a system T^* and a related system $T < T^*$ a.s.. The most usual situation is to consider $T = X_{1:n} = \min(X_1,\ldots,X_n)$. In this case we want to study (predict) the system lifetime T^* when we know the failure time T (i.e. the first components' failure time). To this aim, we can use here (19) and all the results given in Sect. 2.

For example, let us consider the system $T^* = \max(X_1,\min(X_2,X_3))$ and $T = X_{1:3}$. If the joint distribution of (X_1,X_2,X_3) is absolutely continuous, then $T^* < T$. Hence the joint survival function $\bar{\mathbf{F}}$ of (T,T^*) can be written as $\bar{\mathbf{F}}(x,y) =$

$\bar{D}(\bar{F}(x), \bar{F}(y))$ for all $x \le y$, where $\bar{D}(u, v) = \hat{C}(u, v, v) + \hat{C}(v, u, u) - \hat{C}(v, v, v)$ for $0 \le v \le u \le 1$.

Another interesting example is to consider the parallel system with three components $T^* = X_{3:3} = \max(X_1, X_2, X_3)$ and its first components' failure (series system) $T = X_{1:3}$. Note that the distribution F of (T, T^*) is a marginal of the random vector $(X_{1:3}, X_{2:3}, X_{3:3})$ (studied in the preceding subsection). Note that F can be written as (19) for

$$D(u, v) = C(u, v, v) + C(v, u, v) + C(v, v, u) - C(u, u, v) - C(u, v, u) \\ - C(v, u, u) + C(u, u, u)$$

for $0 \le u \le v \le 1$ and $D(u, v) = C(v, v, v)$ for $0 \le v \le u \le 1$. If C is EXC, then

$$D(u, v) = 3C(u, v, v) - 3C(u, u, v) + C(u, u, u)$$

for $0 \le u \le v \le 1$. In particular, if the components are IID, we get

$$D(u, v) = 3uv^2 - 3u^2v + u^3$$

for $0 \le u \le v \le 1$. Note that T and T^* are dependent even if X_1, X_2, X_3 are IID.

These expressions can be used to compute the joint PDF, the marginal and conditional distributions and the conditional median (jointly with the associated confidence regions).

For example, in the last IID case, the joint PDF is

$$g(x, y) = f(x)f(y)d(F(x), F(y))$$

where $d(x, y) = \partial_{1,2}D(u, v) = 6(v - u)$ for $0 \le u \le v \le 1$ (zero elsewhere). The marginal distributions are $G_1(x) = \Pr(X_{1:3} \le x) = D_1(F(x))$ and $G_2(y) = \Pr(X_{3:3} \le y) = D_2(F(y))$, where $D_1(u) = D(u, 1) = 3u - 3u^2 + u^3$ and $D_2(u) = D(1, u) = u^3$ for $u \in [0, 1]$. Analogously, from (8), the conditional distribution function of T^* given T can be written as

$$G_{2|1}(y|x) = \Pr(X_{3:3} \le y | X_{1:3} = x) = D_{2|1}(F(y)|F(x)),$$

where $D_{2|1}$ is given by

$$D_{2|1}(v|u) := \frac{\partial_1 D(u, v)}{\partial_1 D(u, 1)} = \frac{v^2 - 2uv + u^2}{1 - 2u + u^2}$$

for $0 < u \le v \le 1$ since $\lim_{v \to 0^+} \partial_1 D(u, v) = 0$ for all $u \in (0, 1)$.

4 An illustration about paired ordered data

We consider the problem stated in Sect. 3.2. We assume here that we have a sample $(X_1, Y_1), \ldots, (X_m, Y_m)$ from a random vector (X, Y) with joint absolutely continuous distribution function. We also assume that X and Y have a common marginal F. Given some information about $L = \min(X, Y)$, we would like to predict $U =$

$\max(X, Y)$ from L by using the regression curve $E(U|L = t)$ or the median regression curve obtained from the conditional survival function $\Pr(U > y|L = x)$.

To illustrate the problem, we simulate $m = 100$ random points (x_i, y_i) from the joint distribution function $H(x, y) = C(F(x), F(y))$, where C is a suitable copula, while F is either Gaussian with $\mu = 60$ and $\sigma = 5$ or exponential with $\mu = 60$.

First, we consider two independent normal distributions $N(60, 5)$ in Fig. 2, left, and the associated ordered data in Fig. 2, right. Notice that, as can be graphically issued, L and U are dependent even if X and Y are independent. In the right plot, we also include the conditional median curve (green) and the centered 90% and 50% quantile-confidence bands computed from (16).

The analogous plots obtained from two independent exponential distributions $Exp(60)$ are in Fig. 3. Note that the conditional median curve and the quantile-confidence bands in the right plot are determined by means of (17). Note that many of these data could be censored data in practice (for instance, when they are greater than 100) due to the large dispersion of the exponential model and the independence assumption.

Let us consider now that (X, Y) are DID with a copula C and a common marginal distribution F. We consider again the above Gaussian and exponential models for F. Moreover, we assume that

$$C(u, v) = \frac{uv}{u + v - uv} \tag{20}$$

for $(u, v) \in [0, 1]^2$, which represents a positive symmetric dependence between X and Y. Such a copula appears as limiting case in various Archimedean families (see, e.g., Nelsen 2006, p. 116) and it is indicated by the symbol $\Pi/(\Sigma - \Pi)$. This copula is associated with a Kendall's tau equal to 1/3.

We simulate a random sample from (X, Y) by using the inverse transform method (see Example 2.20 in Nelsen 2006). These data are used to get the data from (L, U) with $L = \min(X, Y)$ and $U = \max(X, Y)$ for the normal and exponential models plotted in Fig. 4.

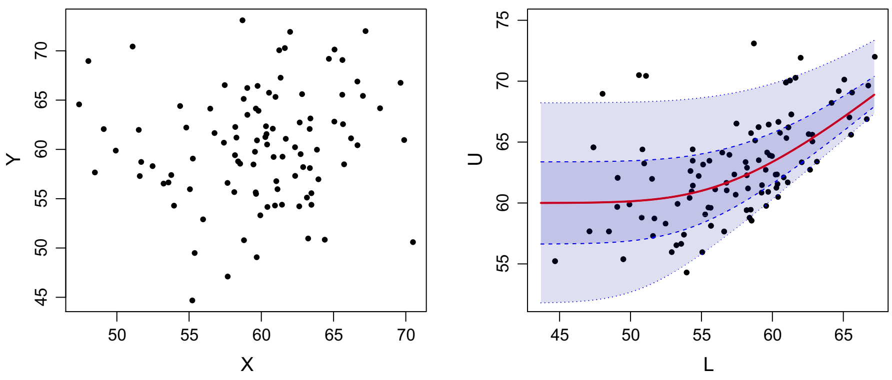

Fig. 2 Scatter plot of a random sample of size $m = 100$ from (X, Y) (left) and (L, U) (right), when X and Y have independent normal distributions with $\mu = 60$ and $\sigma = 5$. In the right plot, the conditional median curve (red) and the 90% (dotted blue) and 50% (dashed blue) quantile-confidence bands are showed

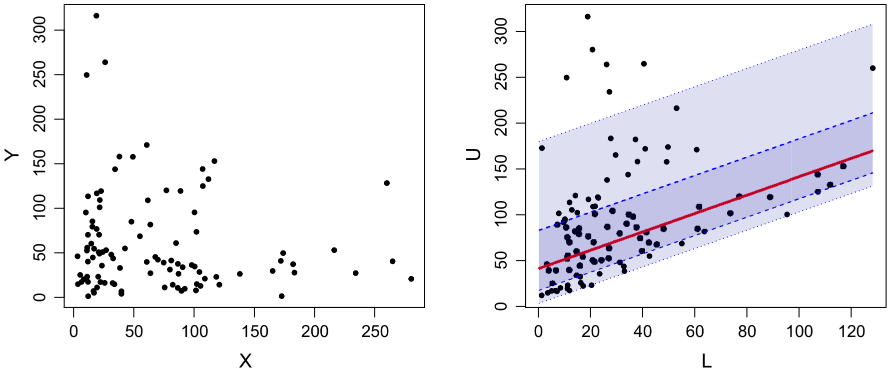

Fig. 3 Scatter plot of a random sample of size $m = 100$ from (X, Y) (left) and (L, U) (right), when X and Y have independent exponential distributions with $\mu = 60$. In the right plot, the conditional median curve (red) and the 90% (dotted blue) and 50% (dashed blue) quantile-confidence bands are showed

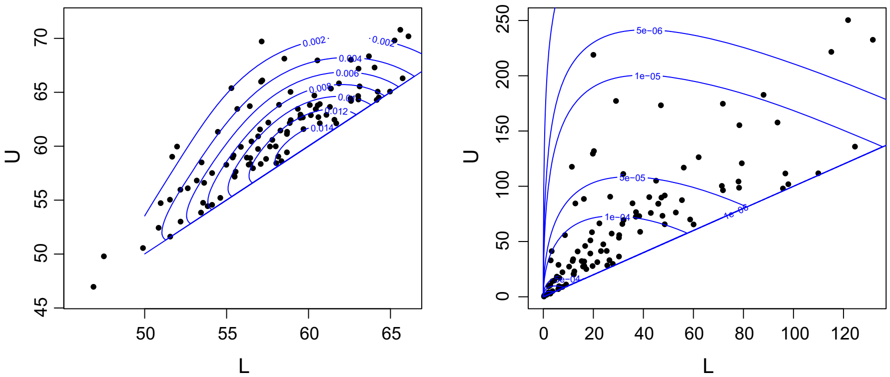

Fig. 4 Contour plots for the joint PDF g of (L, U) when X and Y have the Clayton copula C in (20) and normal marginal distributions with $\mu = 60$ and $\sigma = 5$ (left) and exponential distributions with $\mu = 60$ (right)

In order to obtain a joint model for (L, U) we can use the joint PDF g given in (18) to plot the level curves for (L, U). For the above copula of (20) we have

$$g(x, y) = 2f(x)f(y)c(F(x), F(y))$$

for $x \leq y$, where

$$c(u, v) = \partial_{1,2}C(u, v) = \frac{2uv}{(u + v - uv)^3}$$

for $(u, v) \in [0, 1]^2$. When the common marginal is normal and exponential distributed, respectively, we get the contour plots given in Fig. 4. Analogously, we can plot the marginal PDFs of L and U (see Fig. 5). Note that, for the above $\Pi/(\Sigma - \Pi)$

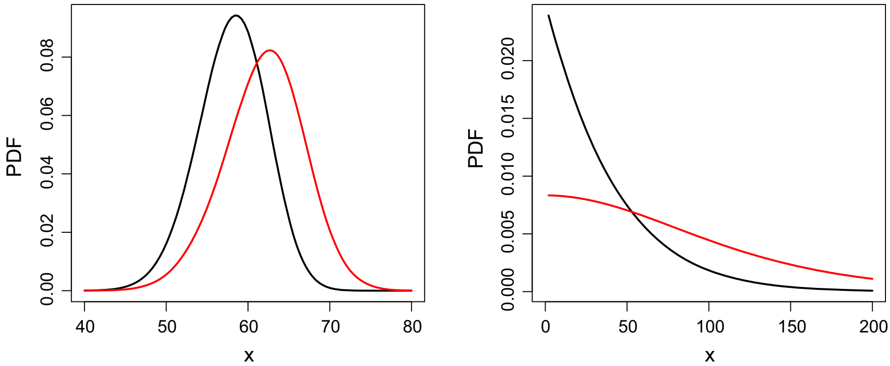

Fig. 5 Plots for the PDF of L (black) and U (red) when X and Y have the Clayton copula C in (20) and normal distributions with $\mu = 60$ and $\sigma = 5$ (left) and exponential distributions with $\mu = 60$ (right)

copula we obtain $D_1(u) = (3u - 2u^2)/(2 - u)$ and $D_2(u) = u/(2 - u)$ for $u \in [0, 1]$. So L and U do not have neither normal nor exponential distributions.

To get the conditional median curves and the confidence bands for (L, U) we need the conditional distribution $G_{2|1}(y|x)$ of $(U|L = x)$ that can be obtained from (15) with

$$\partial_1 D(u, v) = 2\partial_1 C(u, v) - 2\partial_1 C(u, u) = \frac{2v^2}{(u + v - uv)^2} - \frac{2}{(2 - u)^2}$$

for $v \geq u$. Then, to compute the inverse of $G_{2|1}$, we need to solve in y the equation $G_{2|1}(y|x) = q$ for $q \in (0, 1)$. This leads to

$$\frac{F^2(y)}{(F(x) + F(y) - F(x)F(y))^2} = \frac{1 - q + q(2 - F(x))^2}{(2 - F(x))^2}.$$

Therefore

$$y = F^{-1}\left(\frac{F(x)}{F(x) - 1 + \frac{2 - F(x)}{\sqrt{1 - q + q(2 - F(x))^2}}}\right).$$

We use this expression to plot in Fig. 6 the conditional median curves and the associated 90% and 50% confidence bands for the above normal (left) and exponential (right) models. In these figures we also include the empirical regression lines to predict U from L (purple lines). The curves obtained in these models are quite different (as expected, the data from the exponential model are more dispersed). Moreover, the regression line provide a poor fit in the exponential case. Note that in practice, we expect to have a few of censored realizations from the data in Fig. 6, right. For instance, if the data represent the ages (in years) for a disease in paired organs (e.g. breast cancer), the censure means that the patients died before they

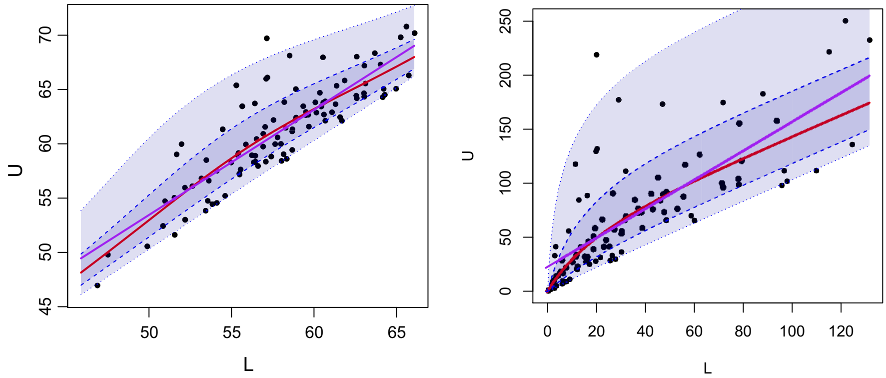

Fig. 6 Scatter plots of size $m = 100$ from (L, U) obtained when X and Y have a copula C of (20) and normal marginal distributions with $\mu = 60$ and $\sigma = 5$ (left) and exponential distributions with $\mu = 60$ (right). There are included the regression line (purple), the conditional median curve (red) and the 90% (dotted blue) and 50% (dashed blue) quantile-confidence bands

suffer this disease in this organ or that they do not have the disease when the experiment ended.

5 Conclusions

The main purpose of this paper is to provide an alternative representation based on distortions to the classical copula representation for the joint distribution F of a random vector. The new representations are more flexible since they are not based on the marginals of the original model. However, a disadvantage is that they do not separate the dependence structure and the marginals. The examples in Sect. 3 show that, in some cases, it could be better to use the new representation instead of the copula representation. The MDD representations for fixed continuous distribution functions $G_1, \ldots, G_n$ are unique (see Proposition 1). Moreover, the distortion function D is uniquely determined by the copula C (see Proposition 2).

As in the copula setting, MDDs can be used to study the lower-dimensional marginals of F and the associated the conditional distributions. They can also be used to obtain the regression and median regression curves and the associated confidence bands.

We provide several examples where these representations are useful. Additional examples can be obtained in a similar way. We also include a simulation study for paired ordered data from independent and dependent variables. This procedure can be used to predict, for instance, the second failure in paired organs from the first one.

This paper represents a first step into the analysis of MDD representations. The main task for future research projects is to develop the appropriate inference procedures (and their properties) to apply these representations to real data sets. In this respect, the related investigations about dynamic copula models could be helpful (see Fermanian and Wegkamp 2012).

Appendix

Proof of proposition1 For every $i = 1, \ldots, n$, G_i is continuous and, hence, its range $Ran(G_i)$ contains the interval (0, 1). Let D be the distribution function of $(G_1(X_1), \ldots, G_n(X_n))$. Then it can be checked that D satisfies properties (i), (ii) and (iii) of Definition 1. Thus, $D \in \mathcal{D}_n$.

Moreover, for every $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$, it follows that

$$\begin{aligned} F(x_1, \ldots, x_n) &= \Pr(X_1 \le x_1, \ldots, X_n \le x_n) \\ &= \Pr(G_1(X_1) \le G_1(x_1), \ldots, G_n(X_n) \le G_n(x_n)) \\ &= D(G_1(x_1), \ldots, G_n(x_n)), \end{aligned}$$

where in the second equality we use that G_i is strictly increasing in the support of X_i. Hence is $F \equiv MDD(G_1, \ldots, G_n)$. □

***Proof of proposition* 2** Since F is continuous, (1) holds for a unique copula C. Thus, for every $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$, it follows that

$$D(G_1(x_1), \ldots, G_n(x_n)) = C(F_1(x_1), \ldots, F_n(x_n)).$$

For every $i = 1, \ldots, n$, since G_i is continuous, there exists $u_i \in (0, 1)$ such that $x_i = G_i^{-1}(u_i)$, where G_i^{-1} is quasi–inverse of G (see, e.g., Durante and Sempi 2016). Thus, it follows that

$$D(u_1, \ldots, u_n) = C(F_1(G_i^{-1}(u_1)), \ldots, F_n(G_i^{-1}(u_n))),$$

which is the desired assertion. □

***Proof of proposition* 3** Clearly, if (2) holds for some distribution functions $G_1, \ldots, G_n$ and $D \in \mathcal{D}_n$, then

$$\lim_{x_i \to -\infty} F(x_1, \ldots, x_n) = \lim_{x_i \to -\infty} D(G_1(x_1), \ldots, G_n(x_n)) = 0$$

for $i = 1, \ldots, n$ and

$$\begin{aligned} \lim_{\min(x_1,\ldots,x_n) \to +\infty} F(x_1, \ldots, x_n) &= \lim_{\min(x_1,\ldots,x_n) \to +\infty} D(G_1(x_1), \ldots, G_n(x_n)) \\ &= D(1, \ldots, 1) = 1 \end{aligned}$$

since $D \in \mathcal{D}_n$. Moreover, F is right-continuous in each variable since D is continuous and $G_1, \ldots, G_n$ are right-continuous.

Let us consider now $(x_1, \ldots, x_n) \in \mathbb{R}^n$ and $(y_1, \ldots, y_n) \in \mathbb{R}^n$ such that $x_i \le y_i$ for $i = 1, \ldots, n$. Then we define $u_i = G_i(x_i)$ and $v_i = G_i(y_i)$ for $i = 1, \ldots, n$. As G_i is a distribution function, we have $0 \le u_i \le v_i \le 1$ for $i = 1, \ldots, n$. Therefore

$$\triangle_{(x_1,\ldots,x_n)}^{(y_1,\ldots,y_n)} F = \triangle_{(u_1,\ldots,u_n)}^{(v_1,\ldots,v_n)} D \ge 0$$

since D satisfies property (iii) in Definition 1. Therefore, F is a proper multivariate distribution function. □

Proof of Proposition 5 The joint distribution function of $(X_1,\ldots,X_m)$ can be written as

$$F_{1,\ldots,m}(x_1,\ldots,x_m)=F(x_1,\ldots,x_m,+\infty,\ldots,+\infty)$$

for all $(x_1,\ldots,x_m)\in\mathbb{R}^m$. Then (5) is obtained from (2) taking into account that $G_i(+\infty)=1$ for any distribution function G_i and $i=m+1,\ldots,n$. Finally, (5) implies $D_{1,\ldots,m}\in\mathcal{D}_m$. □

Proof of Proposition 7 The conditional PDF of $(X_2|X_1=x_1)$ can be written as

$$f_{2|1}(x_2|x_1)=\frac{f(x_1,x_2)}{f_1(x_1)}$$

for all x_1,x_2 such that $f_1(x_1)>0$. Then by using (7) and the fact that

$$f_1(x_1)=g_1(x_1)D_1'(G_1(x_1))>0,$$

where $D_1(u):=D(u,1)$ and $D_1'(u)=\partial_1 D(u,1)$, we obtain

$$f_{2|1}(x_2|x_1)=g_2(x_2)\frac{\partial_{1,2}D(G_1(x_1),G_2(x_2))}{\partial_1 D(G_1(x_1),1)}.$$

Thus, the conditional distribution function can be obtained as

$$F_{2|1}(x_2|x_1)=\int_{-\infty}^{x_2}f_{2|1}(z|x_1)dz=\int_{-\infty}^{x_2}g_2(z)\ \frac{\partial_{1,2}D(G_1(x_1),G_2(z))}{\partial_1 D(G_1(x_1),1)}dz.$$

Now, if we assume $\lim_{v\to 0^+}\partial_1 D(G_1(x_1),v)=0$, then

$$F_{2|1}(x_2|x_1)=\left[\frac{\partial_1 D(G_1(x_1),G_2(z))}{\partial_1 D(G_1(x_1),1)}\right]_{z=-\infty}^{x_2}=\frac{\partial_1 D(G_1(x_1),G_2(x_2))}{\partial_1 D(G_1(x_1),1)}.$$

Hence, (8) holds. □

Proof of Proposition10 First we note that $\Pr(X_i>t)\geq\Pr(X_1>t,\ldots,X_n>t)>0$ for $i=1,\ldots,n$. So we can consider the survival functions $\bar{F}_{1,t},\ldots,\bar{F}_{n,t}$ of the marginal residual lifetimes at time t. Then we note that $\bar{F}_t$ can be written as

$$\begin{aligned}\bar{F}_t(x_1,\ldots,x_n)&=\Pr(X_1-t>x_1,\ldots,X_n-t>x_n|X_1>t,\ldots,X_n>t)\\&=\frac{\Pr(X_1>t+x_1,\ldots,X_n>t+x_n)}{\Pr(X_1>t,\ldots,X_n>t)}\\&=\frac{\bar{F}(t+x_1,\ldots,t+x_n)}{\bar{F}(t,\ldots,t)}\end{aligned}$$

for $x_1,\ldots,x_n\geq 0$. Now we use the following copula representation for $\bar{F}$ (obtained from Sklar's theorem) $\bar{F}(x_1,\ldots,x_n)=\hat{C}(\bar{F}_1(x_1),\ldots,\bar{F}_n(x_n))$, where $\hat{C}$ is a continuous survival copula of $\bar{F}$. Hence

$$\begin{aligned}\bar{F}_t(x_1,\ldots,x_n) &= \frac{\bar{F}(t+x_1,\ldots,t+x_n)}{\bar{F}(t,\ldots,t)}\\ &= \frac{\hat{C}(\bar{F}_1(t+x_1),\ldots,\bar{F}_n(t+x_n))}{\hat{C}(\bar{F}_1(t),\ldots,\bar{F}_n(t))}\\ &= \frac{\hat{C}(\bar{F}_1(t)\bar{F}_{1,t}(x_1),\ldots,\bar{F}_n(t)\bar{F}_{n,t}(x_n))}{\hat{C}(\bar{F}_1(t),\ldots,\bar{F}_n(t))}\\ &= \hat{D}_t(\bar{F}_{1,t}(x_1),\ldots,\bar{F}_{n,t}(x_n))\end{aligned}$$

and (10) holds for the function D_t in (11). Hence $\hat{D}_t \in \mathcal{D}_n$. □

Survival function of $\mathbf{X}_t^{(n)}$ ***given by*** (12) We aim at calculating the survival function of

$$\mathbf{X}_t^{(n,\leq)} := (X_1 - t, \ldots, X_n - t | X_1 > t, \ldots, X_{n-1} > t, X_n \leq t)$$

To this end, consider that

$$\begin{aligned}&\bar{F}_t^{(n,\leq)}(x_1,\ldots,x_{n-1})\\ &\quad= \Pr(X_1 - t > x_1,\ldots,X_{n-1} - t > x_{n-1} | X_1 > t,\ldots,X_{n-1} > t, X_n \leq t)\\ &\quad= \frac{\Pr(X_1 > t+x_1,\ldots,X_{n-1} > t+x_{n-1}, X_n \leq x_n)}{\Pr(X_1 > t,\ldots,X_{n-1} > t, X_n \leq t)}\\ &\quad= \frac{\Pr(X_1 > t+x_1,\ldots,X_{n-1} > t+x_{n-1})}{\Pr(X_1 > t,\ldots,X_{n-1} > t, X_n \leq t)}\\ &\qquad- \frac{\Pr(X_1 > t+x_1,\ldots,X_{n-1} > t+x_{n-1}, X_n > t)}{\Pr(X_1 > t,\ldots,X_{n-1} > t, X_n \leq t)}\\ &\quad= \frac{\bar{F}(t+x_1,\ldots,t+x_{n-1},0) - \bar{F}(t+x_1,\ldots,t+x_{n-1},t)}{\Pr(X_1 > t,\ldots,X_{n-1} > t, X_n \leq t)}\\ &\quad= \frac{\bar{F}(t+x_1,\ldots,t+x_{n-1},0) - \bar{F}(t+x_1,\ldots,t+x_{n-1},t)}{\bar{F}(t,\ldots,t,0) - \bar{F}(t,\ldots,t)}\\ &\quad= \frac{\hat{C}(\bar{F}_1(t+x_1),\ldots,\bar{F}_{n-1}(t+x_{n-1}),1)}{\hat{C}(\bar{F}_1(t),\ldots,\bar{F}_{n-1}(t),1) - \hat{C}(\bar{F}_1(t),\ldots,\bar{F}_n(t))}\\ &\qquad- \frac{\hat{C}(\bar{F}_1(t+x_1),\ldots,\bar{F}_{n-1}(t+x_{n-1}),\bar{F}_n(t))}{\hat{C}(\bar{F}_1(t),\ldots,\bar{F}_{n-1}(t),1) - \hat{C}(\bar{F}_1(t),\ldots,\bar{F}_n(t))}\end{aligned}$$

for $x_1,\ldots,x_{n-1} \geq 0$. Hence it can be written as

$$\bar{F}_t^{(n,\leq)}(x_1,\ldots,x_{n-1}) = D_t^{(n,\leq)}(\bar{F}_{1,t}(x_1),\ldots,\bar{F}_{n-1,t}(x_{n-1})),$$

where $\bar{F}_{1,t},\ldots,\bar{F}_{n-1,t}$ are the survival functions of the univariate residual lifetimes,

$$D_t^{(n,\le)}(\mathbf{u}) = \frac{\hat{C}(\bar{F}_1(t)u_1,\ldots,\bar{F}_{n-1}(t)u_{n-1},1) - \hat{C}(\bar{F}_1(t)u_1,\ldots,\bar{F}_{n-1}(t)u_{n-1},\bar{F}_n(t))}{\hat{C}(\bar{F}_1(t),\ldots,\bar{F}_{n-1}(t),1) - \hat{C}(\bar{F}_1(t),\ldots,\bar{F}_n(t))}$$

for $\mathbf{u} = (u_1,\ldots,u_n) \in [0,1]^n$ and $D_t^{(n,\le)} \in \mathcal{D}_{n-1}$. □

***Proof of Proposition* 11** The distortion function for $n = 3$ can be obtained as follows for $0 \le u_1 \le u_2 \le u_3 \le 1$. If we assume $F(x) = x$ for $x \in [0,1]$, then

$$\begin{aligned} D(u_1,u_2,u_3) &= \Pr(X_{1:3} \le u_1, X_{2:3} \le u_2, X_{3:3} \le u_3) \\ &= \Pr((A_1 \cup A_2 \cup A_3) \cap (A_{1,2} \cup A_{1,3} \cup A_{2,3}) \cap A_{1,2,3}), \end{aligned}$$

where $A_i = \{X_i \le u_1\}$, $A_{i,j} = \{X_i \le u_2\} \cap \{X_j \le u_2\}$, and $A_{1,2,3} = \{X_1 \le u_3\} \cap \{X_2 \le u_3\} \cap \{X_3 \le u_3\}$ for $i,j \in \{1,2,3\}$. Hence

$$D(u_1,u_2,u_3) = \Pr(B_1 \cup \ldots \cup B_9), \tag{21}$$

where $B_1 = A_1 \cap A_{1,2} \cap A_{1,2,3}$, $B_2 = A_2 \cap A_{1,2} \cap A_{1,2,3}$, $B_3 = A_3 \cap A_{1,2} \cap A_{1,2,3}$, $B_4 = A_1 \cap A_{1,3} \cap A_{1,2,3}$, $B_5 = A_2 \cap A_{1,3} \cap A_{1,2,3}$, $B_6 = A_3 \cap A_{1,3} \cap A_{1,2,3}$, $B_7 = A_1 \cap A_{2,3} \cap A_{1,2,3}$, $B_8 = A_2 \cap A_{2,3} \cap A_{1,2,3}$, and $B_9 = A_3 \cap A_{2,3} \cap A_{1,2,3}$. Hence, the formula for D is obtained by applying the inclusion-exclusion formula to (21) taking into account that all these probabilities can be computed from C. For example

$$\Pr(B_1) = \Pr(A_1 \cap A_{1,2} \cap A_{1,2,3}) = \Pr(X_1 \le u_1, X_2 \le u_2, X_3 \le u_3) = C(u_1,u_2,u_3)$$

and

$$\begin{aligned} \Pr(B_1 \cap B_2) &= \Pr(A_1 \cap A_2 \cap A_{1,2} \cap A_{1,2,3}) \\ &= \Pr(X_1 \le u_1, X_2 \le u_1, X_3 \le u_3) \\ &= C(u_1,u_1,u_3) \end{aligned}$$

$0 \le u_1 \le u_2 \le u_3 \le 1$. The other probabilities can be obtained in a similar way. Clearly, this procedure can also be applied to the n dimensional case (but the expression for D gets really involved).

***Proof of Proposition* 12** Let $\mathcal{C}_1,\ldots,\mathcal{C}_s$ and $\mathcal{C}_1^*,\ldots,\mathcal{C}_{s^*}^*$ be the minimal cut sets of T and T^*, respectively. Hence the joint distribution $F(x,y) = \Pr(T \le x, T^* \le y)$ can be written as

$$\begin{aligned} F(x,y) &= \Pr(\min_{i=1,\ldots,s} \max_{k\in\mathcal{C}_i} X_k \le x, \min_{j=1,\ldots,s^*} \max_{k\in\mathcal{C}_j^*} X_k \le y) \\ &= \Pr((A_1 \cup \ldots \cup A_s) \cap (A_1^* \cup \ldots \cup A_{s^*}^*)) \\ &= \Pr\left(\cup_{i=1}^{s} \cup_{j=1}^{s^*} B_{i,j}\right), \end{aligned}$$

where $A_i := \{\max_{k\in\mathcal{C}_i} X_k \le x\}$, $A_j^* := \{\max_{k\in\mathcal{C}_j^*} X_k \le y\}$ and $B_{i,j} := A_i \cap A_j^*$. Now, we can apply the inclusion-exclusion formula to the union of the sets $B_{i,j}$. Moreover we note that, if $x \le y$, then

$$\Pr(B_{i,j}) = \Pr\left(\max_{k\in\mathcal{C}_i} X_k \le x, \max_{k\in\mathcal{C}_j^*} X_k \le y\right)$$

$$= \Pr\left(\max_{k\in\mathcal{C}_i} X_k \le x, \max_{k\in\mathcal{C}_j^*-\mathcal{C}_i} X_k \le y\right) = C_{\mathcal{C}_i,\mathcal{C}_j^*}(F(x), F(y)),$$

where $\mathcal{C}_j^* - \mathcal{C}_i = \mathcal{C}_j^* \cap \bar{\mathcal{C}_i}$ ($\bar{A}$ is the complementary set of the set A), $C_{\mathcal{C}_i,\mathcal{C}_j^*}(u,v) := C(u_1,\ldots,u_n)$, $u_k = F(x)$ if $k \in \mathcal{C}_i$, $u_k = F(y)$ if $k \in \mathcal{C}_j^* - \mathcal{C}_i$, and $u_k = 1$ if $k \notin \mathcal{C}_i \cup \mathcal{C}_j^*$. Similar expressions can be obtained for the other probabilities in the inclusion-exclusion formula as $\Pr(B_{i,j} \cap B_{\ell,r})$, ... and for $x > y$. Hence, we obtain (19). □

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10260-021-00613-2.

Acknowledgements The authors would like to thank the anonymous reviewers for the careful reading and several useful suggestions. JN is partially supported by Ministerio de Ciencia e Innovación of Spain under grant PID2019-103971GB-I00/AEI/10.13039/501100011033, CC and ML are partially supported by the GNAMPA research group of INDAM (Istituto Nazionale di Alta Matematica) and CC, ML and FD are also partially supported by MIUR-PRIN 2017, Project "Stochastic Models for Complex Systems" (No. 2017JFFHSH)

References

Ahmadi J, Nagaraja HN (2020) Conditional properties of a random sample given an order statistic. Stat Pap 61:1971–1996

Arias-Nicolás P, Ruggeri F, Suárez-Llorens A (2016) New classes of priors based on stochastic orders and distortion functions. Bayesian Anal 11:1107–1136

Barlow RE, Proschan F (1975) Statistical theory of reliability and life testing. International Series in Decision Processes, Holt, Rinehart and Winston Inc, New York

Durante F, Foschi R, Sarkoci P (2010) Distorted copulas: constructions and tail dependence. Commun Stat-Theory Methods 39:2288–2301

Durante F, Foschi R, Spizzichino F (2008) Threshold copulas and positive dependence. Stat Probab Lett 78:2902–2909

Durante F, Mesiar R, Papini PL, Sempi C (2007) 2-increasing binary aggregation operators. Inf Sci 177:111–129

Durante F, Sempi C (2016) Principles of copula theory. CRC/Chapman & Hall, London

Embrechts P (2009) Copulas: a personal view. J Risk Insur 76:639–650

Fermanian JD, Wegkamp MH (2012) Time-dependent copulas. J Multivar Anal 110:19–29

Joe H (2005) Asymptotic efficiency of the two-stage estimation method for copula-based models. J Multivar Anal 94:401–419

Joe H (2014) Dependence modeling with copulas. CRC/Chapman & Hall, Boca Raton, FL

Klüppelberg C, Resnick SI (2008) The Pareto copula, aggregation of risks and the emperor's socks. J Appl Probab 45:67–84

Koenker R (2005) Quantile Regression. Cambridge University Press

Liu P, Schied A, Wang R (2021) Distributional transforms, probability distortions, and their applications. Mathematics of Operations Research (**in press**)

Longobardi M, Pellerey F (2019) On the role of dependence in residual lifetimes. Stat Probab Lett 153:56–64

López-Díaz M, Sordo MA, Suárez-Llorens A (2012). On the Lp-metric between a probability distribution and its distortion. *Insurance: Mathematics and Economics* 51, 257–264

Mai J-F, Scherer M (2017) Simulating copulas, 2nd edn. World Scientific, Singapore

Morillas PM (2005) A method to obtain new copulas from a given one. Metrika 61:169–184

Navarro J (2018) Distribution-free comparisons of residual lifetimes of coherent systems based on copula properties. Stat Pap 59:781–800
Navarro J, Balakrishnan N (2010) Study of some measures of dependence between order statistics and systems. J Multivar Anal 101:52–67
Navarro J, Calì C (2019) Inactivity times of coherent systems with dependent components under periodical inspection. Appl Stoch Models Bus Ind 35:871–892
Navarro J, Durante F (2017) Copula-based representations for the reliability of the residual lifetimes of coherent systems with dependent components. J Multivar Anal 158:87–102
Navarro J, del Águila Y, Sordo MA, Suárez-Llorens A (2013) Stochastic ordering properties for systems with dependent identically distributed components. Appl Stochastic Models Bus Ind 29:264–278
Navarro J, del Águila Y, Sordo MA, Suárez-Llorens A (2016) Preservation of stochastic orders under the formation of generalized distorted distributions. Methodol Comput Appl Probab 18:529–545
Navarro J, Durante F, Fernández-Sánchez J (2021) Connecting copula properties with reliability properties of coherent systems. Appl Stochast Models Bus Ind 37:496–512
Navarro J, Longobardi M, Pellerey F (2017) Comparison results for inactivity times of k-out-of-n and general coherent systems with dependent components. Test 26:822–846
Navarro J, Samaniego FJ, Balakrishnan N (2010) The joint signature of coherent systems with shared components. J Appl Probab 47:235–253
Navarro J, Sordo MA (2018) Stochastic comparisons and bounds for conditional distributions by using copula properties. Depend Modeling 6:156–177
Nelsen RB (2006) An introduction to copulas, 2nd edn. Springer, New York
Omelka M, Gijbels I, Veraverbeke N (2009) Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing. Ann Stat 37:3023–3058
Shaked M, Shanthikumar JG (2007) Stochastic orders. Springer Series in Statistics, Springer, New York
Sklar A (1959) Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris 8:229–231
Sordo MA, Castaño-Martínez A, Pigueiras G (2016). A family of premium principles based on mixtures of TVaRs. *Insurance: Mathematics and Economics* 70, 397–405
Sordo MA, Suárez-Llorens A (2011). Stochastic comparisons of distorted variability measures. *Insurance: Mathematics and Economics* 49, 11–17
Sumarjaya IW (2017). A survey of kernel-type estimators for copula and their applications. *Journal of Physics: Conference Series 893 012027.*https://doi.org/10.1088/1742-6596/893/1/012027
Taylor MD (2016) Multivariate measures of concordance for copulas and their marginals. Depend Modeling 4:224–236
Valdez EA, Xiao Y (2011) On the distortion of a copula and its margins. Scand Actuar J 4:292–317
Wang S (1996) Premium calculation by transforming the layer premium density. ASTIN Bull 26:71–92
Yaari ME (1987) The dual theory of choice under risk. Econometrica 55:95–115