METRON - *International Journal of Statistics*
2003, vol. LXI, n. 3, pp. 323-353

KOUROS OWZAR – PRANAB KUMAR SEN

Copulas: concepts and novel applications

Summary - A bivariate copula can be statistically interpreted as a bivariate distribution function with uniform marginals. Sklar (1959) argues that for any bivariate distribution function, say $\mathbb{H}$ with marginals F and G, there exists a copula functional, say $\mathbb{C}$, such that

$$\mathbb{H}[x, y] = \mathbb{C}[F[x], G[y]],$$

for $(x, y)^T$ in the support of $\mathbb{H}$. What is to presented is a self-contained review, mainly from a statistical point of view, of the concept of copulas vis-a-vis multivariate distributions and dependence and to motivate their utility via a number of applications to the design of clinical trials, microarray studies with survival endpoints and the analysis of dependent Receiver Operator Curves (ROC).

Key Words - Copula; Dependence; Semi-parametrics; Clinical Trials; Microarrays; Survival; ROC.

1. Introduction and problem setting

1.1. Introductory remarks

The goal of this paper is to present a self-contained review of the concept behind copulas vis-a-vis multivariate distributions and dependence with a view mainly towards biostatistical applications. To keep the presentation both readable as well as manageable in volume, with some notable exceptions, most technical, whether of complicated or trivial nature, have been omitted. Instead, a representative, although by no means exhaustive, list of references has been provided. For notational simplicity and clarity, we have focused our attention to bivariate distributions, although the bulk of the material, more often than not, trivially extends to higher dimensions. We will begin by introducing the notion behind a copula including a number of key examples. We will continue by outlining estimation techniques. Thereafter, we will introduce additional key properties and useful ideas related to copulas. Finally, we will motivate the utility of the presented material by considering a number of applications.

Received September 2003 and revised November 2003.

1.2. Motivational examples

The concept of copulas in relation to multivariate distributions and dependence can be motivated in many different ways. We have chosen to motivate these relationships through the following two standard examples of bivariate distributions. It is a matter of simple algebra to exhibit that any bivariate normal distribution function $\mathbb{H}$, with marginals F and G, both univariate normal distributions, and correlation coefficient $\theta \in [-1, 1]$, is expressible as

$$\mathbb{H}[x, y] = \mathbb{C}_{\mathbb{N}}[F[x], G[y], \theta]\,, \tag{1}$$

where

$$\mathbb{C}_{\mathbb{N}}[u, v] = \int_{-\infty}^{\Phi^{-1}[u]} \int_{-\infty}^{\Phi^{-1}[v]} \frac{1}{2\pi\sqrt{1-\theta^2}} \exp\left[-\frac{x^2 - 2\theta xy + y^2}{2(1-\theta^2)}\right] \mathrm{d}x\mathrm{d}y\,, \tag{2}$$

and $\Phi[z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp[-\frac{1}{2}a^2]\mathrm{d}a$ denotes standardized univariate normal distribution function. Similarly, a bivariate exponential distribution (see for example Gumbel (1960)) $\mathbb{H}$, with marginals F and G, both univariate exponential distributions, and association parameter $\theta \in [0, 1]$ is expressible in the form

$$\mathbb{H}[x, y] = \mathbb{C}_{\mathbb{E}}[F[x], G[y], \theta]\,, \tag{3}$$

where

$$\mathbb{C}_{\mathbb{E}}[u, v, \theta] = 1 - u - v + uve^{-\theta \log[u] \log[v]}\,, \tag{4}$$

for $(u, v) \in [0, 1]^2$. Needless to say, we have not presented anything new about these standard examples of bivariate distributions. The key point, however, to note here is that in both of these examples, the joint distribution functions were expressible as functions of their respective marginals, F and G, and a finite-dimensional θ which controls the degree of dependence between the two components of the distribution. There are more general bivariate distributions, most notably the family of bivariate elliptical distribution, which yield themselves to this form. The functions $\mathbb{C}_{\mathbb{N}}$ in (2) and $\mathbb{C}_{\mathbb{E}}$ in (4) couple the joint distribution to their respective marginals and as such have been aptly named copulas. It turns out that there is a rich family of copulas that lend bivariate distributions to the aforementioned form. Most of the concepts to be presented will easily generalize to higher dimensions. However, to keep the notation simple and more manageable, we will mostly limit our attention to the bivariate case.

Remark 1. Throughout this presentation, $\mathbb{H}$ will denote the target bivariate distribution function with marginals F and G respectively. The supports of these marginals are to be denoted by $\mathbb{F}$ and $\mathbb{G}$ which belong to spaces $\mathcal{F}$ and $\mathcal{G}$ respectively. The corresponding bivariate and marginal survival functions are denoted by $\overline{\mathbb{H}}$, $\overline{F}$ and $\overline{G}$ respectively.

1.3. Axiomatic definitions

A copula can be defined from both axiomatic and statistical points of view. Although, the emphasis in this exposition will focus on the statistical perspective, for the sake of completeness and those interested in more technical definitions, we will present axiomatic definitions for bivariate distributions and copulas. As such, this section may be skipped by those primarily interested in statistical applications of copulas. A bivariate distribution function can be axiomatically defined as follows.

Definition 1 [Bivariate distribution]**.** A mapping $\mathbb{H} := \mathbb{F} \times \mathbb{G} \to [0, 1]$, where $\mathbb{F}, \mathbb{G} \subset \mathbb{R}$, is a bivariate distribution function if it satisfies the following three conditions:

1. $\mathbb{H}[x, -\infty] = \mathbb{H}[-\infty, y] = 0$, for all $x \in \mathbb{F}$ and $y \in \mathbb{G}$.
2. $\mathbb{H}[+\infty, +\infty] = 1$.
3. $\mathbb{H}[x_2, y_2]-\mathbb{H}[x_1, y_2]-\mathbb{H}[x_2, y_1]+\mathbb{H}[x_1, y_1] \geq 0$ for every $(x_1, y_1)^T, (x_2, y_2)^T \in \mathbb{F} \times \mathbb{G}$ satisfying $x_1 \leq x_2$ and $y_1 \leq y_2$.

Axiomatically, a copula can be defined as follows.

Definition 2 [Copula]**.** A (two-dimensional) copula $\mathbb{C}$ is a mapping from $[0, 1]\times[0, 1]$ (the unit square) onto $[0, 1]$ (the unit interval) which satisfies the following three conditions:

1. $\mathbb{C}[u, 0] = \mathbb{C}[0, u] = 0$ for every $u \in [0, 1]$;
2. $\mathbb{C}[u, 1] = \mathbb{C}[1, u] = u$ for every $u \in [0, 1]$;
3. $\mathbb{C}[u_2, v_2] - \mathbb{C}[u_1, v_2] - \mathbb{C}[u_2, v_1] + \mathbb{C}[u_1, v_1] \geq 0$ for every $u_1, u_2, v_1, v_2 \in [0, 1]$ satisfying $u_1 \leq u_2$ and $v_1 \leq v_2$.

A trivial example of a copula is the function defined as

$$\mathbb{C}_{\mathrm{I}}[u, v] = uv\,, \tag{5}$$

for $(u, v)^T \in [0, 1]^2$, which is often called the independence copula for reasons that will become clear later on. The picture that comes to mind, by virtue of this definition and perhaps by using figure 1 as a visual aid, is that of a surface bounded within the unit cube which is tied down along the two axes in the first quadrant. This surface is non-decreasing as imposed by the above inequality, which often referred to as the rectangle inequality (a generalization of the triangle inequality for real numbers to $\mathbb{R}^2$).

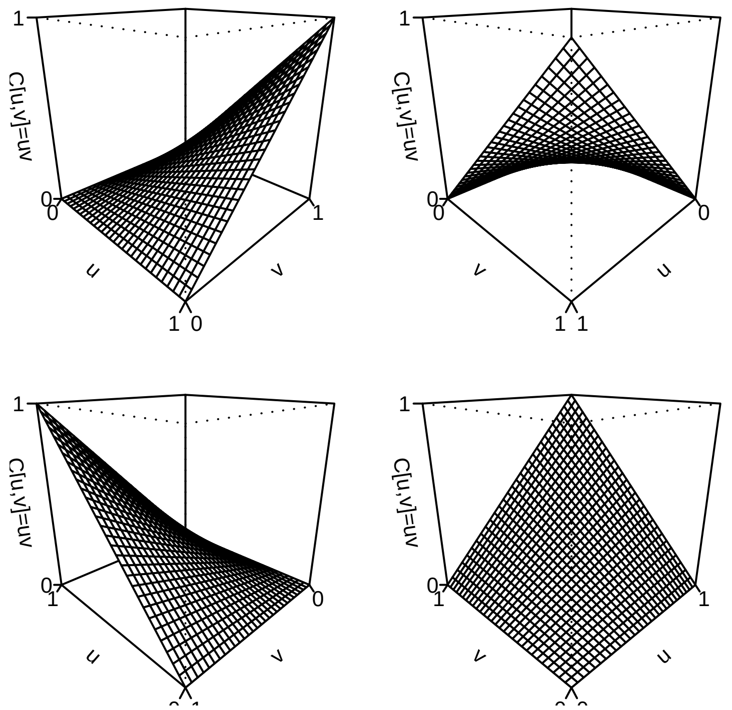

Figure 1. Surface plot of the independence copula $\mathbb{C}_I[u, v] = uv$ for $(u, v)^T \in [0, 1]^2$.

In light of definition (1), a copula, as defined by (2), is a bivariate distribution function with uniform marginals. This will yield a statistical definition for a copula which will be provided in a formal manner shortly.

1.4. Copulas and multivariate distributions

The first order of business is to formally establish the relationship between copulas and bivariate distributions. To that end, we will present a result that has generally been attributed to Sklar (1959), who argues that any given bivariate distribution function is expressible as copula of its marginals. A formal presentation is provided next.

Theorem 1. *Given is a distribution function*, *say* $\mathbb{H}$, *with marginals marginals* F *and* G. *There exists a function* $\mathbb{C}$ *satisfying the properties of definition* (2), *namely a copula*, *such that*

$$\mathbb{H}[x, y] = \mathbb{C}[F[x], G[y]], \tag{6}$$

for every $(x, y)^T \in \mathbb{F} \times \mathbb{G}$. *Furthermore*, *if the marginals* F *and* G *are both continuous*, *then representation* (6) *is unique. Conversely*, *given any copula* *function*, *say* $\tilde{\mathbb{C}}$, *and any pair of continuous univariate distribution functions*, *say* $\tilde{F}$ *and* $\tilde{G}$, *the function defined as*

$$\tilde{\mathbb{H}}[x, y] := \tilde{\mathbb{C}}[\tilde{F}[x], \tilde{G}[y]], \tag{7}$$

for each $(x, y)^T \in \tilde{\mathbb{F}} \times \tilde{\mathbb{G}}$, *is a bivariate distribution function as defined in* (1).

The representation in (6) suggests the appropriateness of the term copula to describe $\mathbb{C}$ as it couples the marginals F and G to the joint distribution function $\mathbb{H}$. Note that by virtue of the converse to this result, if uniform marginals, which are of course continuous, are plugged into any copula, the resulting function is a bivariate distribution function with uniform marginals. This reconciles, as mentioned in the last section, with comparing the axiomatic definitions for a bivariate distribution (1) and that of a bivariate copula (2) and yields the following statistical definition.

Definition 3 [Copula]**.** A copula is a bivariate distribution function with uniform marginals.

Furthermore, the representation (6) suggests that if the copula $\mathbb{C}$ were known, then substituting continuous marginal estimators for F and G would yield a plug-in estimate of their associated joint distribution function $\mathbb{H}$. Moreover, in light of Sklar's result with arrive at the following functional definition of a copula

Definition 4 [Copula]**.** Given a bivariate distribution function $\mathbb{H}$ with marginals F and G, the function defined as

$$\mathbb{C}_{\mathbb{H}}[u, v] = \mathbb{H}[F^{-}[u], G^{-}[u]], \tag{8}$$

for $(u, v)^T \in [0, 1]^2$, where F^{-} and G^{-} are the inverse functions of F and G respectively, is the copula corresponding to $\mathbb{H}$.

We conclude this section by presenting the following important comment.

Remark 2. Given a bivariate distribution function $\mathbb{H}$ with continuous marginals F and G, let $\mathbb{C}$ the unique copula expressing the distribution function as $\mathbb{H}[x, y] = \mathbb{C}[F[x], G[y]]$. We will say that $\mathbb{C}$ is generated by $\mathbb{H}$. What should be noted is that the pair $(F[X], G[Y])^T$ is distributed according to the copula $\mathbb{C}$. Similarly, given a pair $(U, V)^T$ drawn from the uniform copula $\mathbb{C}$, the pair defined as $(F^{-}[U], G^{-}[V])^T$ is distributed according to $\mathbb{H}$.

1.5. Copulas, dependence, independence and exchangeability

Sklar's theorem provides a formal framework for exploring the relationship between copulas and multivariate distributions. To explore the relationship between copulas and dependence, we will introduce the notions of the Frechet-Hoeffding bounds and tail dependence next. We begin by formally defining

$$\mathbb{C}_{\mathrm{L}}[u, v] = \max[u + v - 1, 0]\,, \tag{9}$$
$$\mathbb{C}_{\mathrm{U}}[u, v] = \min[u, v]\,, \tag{10}$$

which we will refer to as the Frechet-Hoeffding lower and upper bounds. Some fundamental relationships between copulas and these bounds and the independence copula $\mathbb{C}_{\mathrm{I}}$ defined in (5) are itemized next.

Theorem 2. *Suppose that $(X, Y)^T$ is a random pair from a distribution $\mathbb{H}$ with continuous marginals and let $\mathbb{C}$ denote the corresponding copula.*

i. *$\mathbb{C}_{\mathrm{L}}$, $\mathbb{C}_{\mathrm{U}}$ and $\mathbb{C}_{\mathrm{I}}$ are all bona-fide copulas.*
ii. *For each $(u, v)^T \in [0, 1]^2$,*

$$\mathbb{C}_{\mathrm{L}}[u, v] = \max[u + v - 1, 0] \le \mathbb{C}[u, v] \le \min[u, v] = \mathbb{C}_{\mathrm{U}}[u, v]\,. \tag{11}$$

iii. *X and Y are independent if and only if $\mathbb{C} = \mathbb{C}_{\mathrm{I}}$ (the independence copula (5)).*
iv. *Let h be any arbitrary mapping of $\mathcal{F}$ into $\mathcal{G}$. Then h is almost surely strictly increasing (or decreasing) on $\mathcal{G}$ if and only if $\mathbb{C} = \mathbb{C}_{\mathrm{U}}$ (or $\mathbb{C} = \mathbb{C}_{\mathrm{L}}$), where $\mathbb{C}_{\mathrm{U}}$ and $\mathbb{C}_{\mathrm{L}}$ are the upper and lower Frechet-Hoeffding bounds defined by (10) and (9) respectively.*

Remark 3. Sklar's theorem and the function definition of copulas trivially extend to higher dimensions. The results in the previous theorem also extend to higher dimensions by defining the Frechet-Hoeffding bounds and the independence copula, as for example is the three-dimensional case, by $\mathbb{C}_{\mathrm{L}}[u, v, w] = \max[u + v + w - 1, 0]$, $\mathbb{C}_{\mathrm{U}}[u, v, w] = \min[u, v]$ and $\mathbb{C}_{\mathrm{I}}[u, v, w] = uvw$. We do point out that the lower Frechet-Hoeffding bound, in higher dimensions, is not a copula.

In turns out that copulas are invariant under increasing transformations. Furthermore, under decreasing transformations the resulting copula, can be expressed these properties are to be formalized in the following result.

Theorem 3. *Given is a pair of random variables $(X, Y)^T$ distributed according to $\mathbb{H}$ with continuous marginals F and G. Let $\mathbb{C}_{XY}$ denote the copula generated by $(X, Y)^T$.*

i. *Suppose that $h_1 : \mathbb{F} \to \mathbb{R}$ and $h_2 : \mathbb{G} \to \mathbb{R}$ are strictly increasing and define $Z_1 = h_1[X]$ and $Z_2 = h_2[Y]$. Then*

$$\mathbb{C}_{XY} = \mathbb{C}_{Z_1 Z_2}\,, \tag{12}$$

where $\mathbb{C}_{Z_1 Z_2}$ denotes the copula generated by $(Z_1, Z_2)^T$.

ii. *Suppose that $h_1 : \mathbb{F} \to \mathbb{R}$ and $h_2 : \mathbb{G} \to \mathbb{R}$ are both strictly decreasing and define $Z_1 = h_1[X]$ and $Z_2 = h_2[Y]$. Then*

$$\mathbb{C}_{Z_1 Y}[u, v] = v - \mathbb{C}_{XY}[1-u, v]\,, \tag{13}$$

$$\mathbb{C}_{X Z_2}[u, v] = u - \mathbb{C}_{XY}[u, 1-v]\,, \tag{14}$$

$$\mathbb{C}_{Z_1 Z_2}[u, v] = u + v - 1 + \mathbb{C}_{XY}[u, v]\,. \tag{15}$$

Next, we will introduce the concept of tail-dependence and illustrate its intimate relationship to copulas. Suppose that $(X, Y)^T$ is a random pair from some distribution $\mathbb{H}$. Tail-dependence, as formalized in the following definition, can be used to quantify the likelihood that X attains an extreme value given that Y has attained an extreme value in its support.

Definition 5 [Tail-Dependence]**.** $\mathbb{H}$ with continuous marginals F and G respectively. If the limit

$$\lambda_{\mathbb{L}}[\mathbb{H}] := \lim_{u \downarrow 0} \lambda_{\mathbb{L}}[\mathbb{H}; u] = \lim_{u \downarrow 0} \mathbb{P}[X \leq F^{-}[u] | Y \leq G^{-}[u]]\,, \tag{16}$$

exists, then $\mathbb{H}$ is said to have lower tail-dependence (LTD) if $\lambda_{\mathbb{L}}[\mathbb{H}] \in (0, 1]$ or to have lower tail-independence (LTI) if $\lambda_{\mathbb{L}}[\mathbb{H}] = 0$. Similarly, if the limit

$$\lambda_{\mathbb{U}}[\mathbb{H}] := \lim_{u \uparrow 1} \lambda_{\mathbb{U}}[\mathbb{H}; u] = \lim_{u \uparrow 1} \mathbb{P}[X > F^{-}[u] | Y > G^{-}[u]]\,, \tag{17}$$

exists, then $\mathbb{H}$ is said to have upper tail-dependence (UTD) if $\lambda_{\mathbb{U}}[\mathbb{H}] \in [0, 1)$ or to have upper tail-independence (UTI) if $\lambda_{\mathbb{U}}[\mathbb{H}] = 0$.

It is a matter of some algebraic manipulations to show that the functions $\lambda_{\mathbb{L}}[\mathbb{H}; u]$ in (16) and $\lambda_{\mathbb{U}}[\mathbb{H}; u]$ in (17) maybe expressed, in terms of the copula $\mathbb{C}$ generated by $\mathbb{H}$ as

$$\lambda_{\mathbb{L}}[\mathbb{H}; u] = \frac{1}{u} \mathbb{C}[u, u] \tag{18}$$

and

$$\lambda_{\mathbb{U}}[\mathbb{H}; u] = \frac{1}{1-u} \left\{ \mathbb{C}[u, u] - 2u + 1 \right\}\,, \tag{19}$$

which suggests tail-dependence, or lack thereof, for bivariate distributions may be studied using copulas.

It turns out that copulas are related to some of the standard non-parametric measures of dependence. Again, assume that $Z = (X, Y)^T$ is a random variable with joint distribution $\mathbb{H}$ and marginals F and G, which admits a copula model of the form $\mathbb{H}[x, y] = \mathbb{C}[F[x], G[y]]$, for $(x, y)^T \in \mathbb{F} \times \mathbb{G}$. Using the transformation $(u, v)^T = (F[x], G[y])^T$, in the integrals below, we obtain

$$\begin{aligned} r[\mathbb{H}] &:= \frac{1}{\sqrt{V_1 V_2}} \int_{\mathbb{R}} \int_{\mathbb{R}} \{\mathbb{H}[x, y] - F[x]G[y]\} \mathrm{d}x \mathrm{d}y \\ &= \frac{1}{\sqrt{V_1 V_2}} \int_0^1 \int_0^1 \{\mathbb{C}[u, v] - uv\} \mathrm{d}F_1^{-1}[u] \mathrm{d}F_2^{-1}[v] , \end{aligned} \tag{20}$$

$$\begin{aligned} \rho[\mathbb{H}] &:= 12 \int_{\mathbb{R}} \int_{\mathbb{R}} \{\mathbb{H}[x, y] - F[x]G[y]\} \mathrm{d}F[x] \mathrm{d}G[y] \\ &= 12 \int_0^1 \int_0^1 \{\mathbb{C}[u, v] - uv\} \mathrm{d}u \mathrm{d}v \\ &= 12 \int_0^1 \int_0^1 \mathbb{C}[u, v] \mathrm{d}u \mathrm{d}v - 3 \end{aligned} \tag{21}$$

$$\begin{aligned} \tau[\mathbb{H}] &= 4 \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{H}[x, y] \mathrm{d}\mathbb{H}[x, y] - 1 \\ &= 4 \int_0^1 \int_0^1 \mathbb{C}[u, v] \mathrm{d}\mathbb{C}[u, v] - 1 , \end{aligned} \tag{22}$$

$$\begin{aligned} \Delta[\mathbb{H}] &= \int_{\mathbb{R}} \int_{\mathbb{R}} \{\mathbb{H}[x, y] - F[x]G[y]\}^2 \mathrm{d}H[x, y] \\ &= \int_0^1 \int_0^1 \{\mathbb{C}[u, v] - uv\}^2 \mathrm{d}\mathbb{C}[u, v] , \end{aligned} \tag{23}$$

$$\begin{aligned} \gamma[\mathbb{H}] &= 2 \int_{\mathbb{R}} \int_{\mathbb{R}} \{|F[x] + G[y] - 1| - |F[x] - G[y]|\} \mathrm{d}\mathbb{H}[x, y] \\ &= 2 \int_0^1 \int_0^1 \{|u + v - 1| - |u - v|\} \mathrm{d}\mathbb{C}[u, v] , \end{aligned} \tag{24}$$

where $\mathbb{C}[u, v] = \mathbb{H}[F^{-1}[u], G^{-1}[v]]$ and where V_1 and V_2 in (20) denote the variances of the marginals to F and G respectively. We note that r, ρ, τ, Δ and γ above are the famous Pearson Correlation, Spearman's ρ, Kendall's τ, Hoeffding's Δ and Gini's γ respectively.

We have already brought to attention the obvious fact that two random variables are independent if and only if their generating copula is $\mathbb{C}_{\mathrm{I}}$. We end this section by formalizing the notion of a copula in the context of exchangeable random variables.

Theorem 4. *Given is a random pair (X, Y) from a distribution $\mathbb{H}$ with continuous marginals F and G with common support $\mathbb{F}$. X and Y are exchangeable if and only if $F[x] = G[x]$ for every $x \in \mathbb{F}$ and $\mathbb{C}[u, v] = \mathbb{C}[v, u]$ for all $(u, v)^T \in [0, 1]^2$.*

1.6. Parametric and archimedean copulas

In the introduction, we exemplified bivariate distributions which were expressible as a function of their marginals and a finite dimensional parameter which controlled the degree of dependence. The one-parameter copula family, to be discussed in some detail shortly, generalizes the aforementioned type of bivariate distributions. In particular, we will assume that the joint distribution function $\mathbb{H}$ is expressible as

$$H[x, y] = \mathbb{C}[F[x], G[y], \theta], \tag{25}$$

where $\theta \in \Theta$ is an unknown parameter, which as in the previously discussed bivariate normal and exponential examples, controls the degree of dependence between the two components. Next, we will present a discussion on four standard parametric copulas. For three of the examples, we have indicated what is believed to be the original or at least an early reference for that copula.

Example 1 [Normal copula]**.** The normal copula, also referred to as the Gaussian copula, is defined as

$$\mathbb{C}_{\mathbb{N}}[u, v, \theta] = \Phi_2[\Phi_1^{-1}[u], \Phi_1^{-1}[v], \theta], \tag{26}$$

where Φ_2, the standardized bivariate distribution function with correlation coefficient $\theta \in [-1, 1]$, explicitly defined in (2), and Φ_1^{-1} denotes the inverse function for a univariate standard normal distribution function. For this copula, $\lim_{\theta\uparrow\downarrow 0} \mathbb{C}[u, v, \theta] \uparrow\downarrow \mathbb{C}[u, v, 0] = \mathbb{C}_{\mathrm{I}}$, $\lim_{\theta\uparrow 1} \mathbb{C}[u, v, \theta] \uparrow \mathbb{C}[u, v, 1] = \mathbb{C}_{\mathrm{U}}$, and $\lim_{\theta\downarrow -1} \mathbb{C}[u, v, \theta] \downarrow \mathbb{C}[u, v, -1] = \mathbb{C}_{\mathrm{L}}$.

Example 2 [Gumbel's copula]**.** The Gumbel copula (see Gumbel (1961)) with parameter θ is defined as

$$\mathbb{C}_{\mathbb{G}}[u, v, \theta] = \exp[-((-\log[u])^{\theta} + (-\log[v])^{\theta})^{\frac{1}{\theta}}] \tag{27}$$

where the parameter $\theta \in [1, \infty)$. For this copula if $\theta = 1$ attains $\mathbb{C}_{\mathrm{I}}$ and $\lim_{\theta\uparrow\infty} \mathbb{C}[u, v, \theta] \uparrow \mathbb{C}_{\mathrm{U}}$. As such, it only admits positive dependence. It also has upper tail dependence.

Example 3 [Frank copula]**.** Frank's copula (see Frank (1979)) with parameter $\theta \in (-\infty, \infty) - \{0\}$ is given by

$$\mathbb{C}_{\mathbb{F}}[u, v, \theta] = -\frac{1}{\theta} \log\left[1 - \frac{(1 - \exp[-\theta u])(1 - \exp[-\theta v])}{1 - \exp[-\theta]}\right]. \tag{28}$$

For this copula, $\lim_{\theta\uparrow\downarrow 0} \mathbb{C}[u, v, \theta] \uparrow\downarrow \mathbb{C}_{\mathrm{I}}$, $\lim_{\theta\uparrow\infty} \mathbb{C}[u, v, \theta] \uparrow \mathbb{C}_{\mathrm{U}}$, and $\lim_{\theta\downarrow-\infty} \mathbb{C}[u, v, \theta] \downarrow \mathbb{C}_{\mathrm{L}}$. Note that similar to the normal copula, it admits positive as well as negative dependence. What should be noted is that if one were to plug in normal marginals in (28), the resulting function, by virtue to the converse of Sklar's result, is a bona-fide bivariate distribution with normal marginals. It is, however, not a bivariate normal distribution function.

Example 4 [Clayton's copula]**.** Clayton's copula (see Clayton (1978)) with parameter $\theta \in (1, \infty)$ is given by

$$\mathbb{C}[u, v, \theta] = \{u^{1-\theta} + v^{1-\theta} - 1\}^{\frac{1}{1-\theta}} . \tag{29}$$

For this copula, $\lim_{\theta \downarrow 1} \mathbb{C}[u, v, \theta] \downarrow \mathbb{C}_\mathrm{I}$ and $\lim_{\theta \uparrow \infty} \mathbb{C}[u, v, \theta] \uparrow \mathbb{C}_\mathrm{U}$. Unlike the normal and Frank's copula, it only admits positive dependence.

An important subclass of parametric copulas is the Archimedean Copula (AC) class. This rich class of copulas does not only enjoy a very simple representation, as shown below, but also can be used, as will be discussed later, as a goodness of fit tool. Let us start with providing a formal definition.

Definition 6 [AC copulas and distributions]**.** A copula function $\mathbb{C}$ is said to be Archimedean if there exists a convex function ϕ with $\phi[1] = 0$, such that the copula $\mathbb{C}$ is expressible as

$$\mathbb{C}[u, v] = \phi^{-1}[\phi[u] + \phi[v]] , \tag{30}$$

for all $(u, v)^T \in [0, 1]^2$. A bivariate distribution function with continuous marginals is said to be AC if its generating copula is AC. In this case, the function ϕ is the called the generator.

It is easy to see, for example, that for the independence copula $\mathbb{C}_\mathrm{I}[u, v] = uv$, the generator is given by $\phi[u] = -\log[u]$ and that for Gumbel copula (2) the generator is given by $\phi[u; \theta] = (-\log[u])^\theta$.

We will end this section by making the following remark.

Remark 4. If a copula $\mathbb{C}$ coincides with one of the copulas $\mathbb{C}_\mathrm{I}, \mathbb{C}_\mathrm{L}$ and $\mathbb{C}_\mathrm{U}$, the results of Theorem 2 would be useful in studying dependence. In practical settings, one would need to assess how close a copula $\mathbb{C}$ is to one of these three copulas. This would require the comparison between two surfaces (i.e., the copula $\mathbb{C}$ versus $\mathbb{C}_\mathrm{I}, \mathbb{C}_\mathrm{L}$ or $\mathbb{C}_\mathrm{U}$). In empirical settings, such comparisons, although possible by considering empirical copulas (to be discussed later), may not be desirable. In the context of parametric copulas, such as those discussed in this section, this difficulty is circumvented as closeness between the copula and three reference copulas is quantified by a single parameter.

1.7. General literature and historical notes

As pointed out in Nelsen (1998), a comprehensive monograph on copulas, early work on concepts related on copulas can be found in works by Hoeffding, who considered bivariate standardized distributions whose support is $[-\frac{1}{2}, \frac{1}{2}]$ instead that of copulas which is $[0, 1]^2$. Hoeffding's work did not receive a lot of attention, as it was published in not readily available European journals, until his collected works were translated and reprinted in a volume edited by Fisher and Sen (1994). Other insightful accounts for historical perspectives on copulas can be found in Schweizer (1991) and Sklar (1996). Another monograph by Joe (1997) provides not only a comprehensive treatment of copulas, from a statistical point of view, but is also a general reference on multivariate models and dependence. Insightful and extensive coverage on the subject of dependence in the context of copulas, in particular in relation to the standard non-parametric measures of monotone dependence (20)–(24), are furnished in Schweizer and Wolff (1981), Schweizer and Sklar (1983) and Scarsini (1984). The latter article also discusses the implication of discrete distributions on dependence, an issue also taken up in Joe (1997) and and Nelsen (1998), which is often, as in the case of this paper, ignored in the literature. Cifarelli *et al.* (1996) provide a comprehensive discussion on the topic of monotone dependence and suggest a general index for monotone dependence of which Gini's γ (24) and Spearman's ρ (21) are special cases. Chapter 5 of Nelsen (1998) also provides a detailed discussion on these topics of including a thorough treatment of tail-dependence. Both Nelsen (1998) (pages 94-97) and Joe (1997) (Section 5.1) provide comprehensive lists of parametric copulas and their properties including the generators. A discussion on copulas can also be found in an entry by Fisher (1997) in the Encyclopedia of Statistical Sciences. Fisher and Switzer (2001) provide discussions on graphical methods for representing dependence in the context of copulas. For early work on families of general bivariate distributions see Mardia (1970).

2. Estimation

2.1. Model

The underlying model of interest is expressed as

$$\mathbb{H}[x, y] = \mathbb{C}[F[x], G[y], \theta], \tag{31}$$

where F and G are a pair of absolutely continuous marginals with densities f and g respectively. We do point out that often additional smoothness and regularity conditions must be imposed. What is observed is a random sample of pairs

$$Z^n = \{Z_1, \dots, Z_n\} = \{(X_1, Y_1)^T, \dots, (X_n, Y_n)^T\}, \tag{32}$$

assumed to be drawn from the target distribution $\mathbb{H}$ as specified in (31). Corresponding to this observed sample Z^n, we will also define

$$W^n = \{W_1, \dots, W_n\} = \{(U_1, V_1)^T, \dots, (U_n, V_n)^T\}, \tag{33}$$

where $(U_i, V_i)^T = (F[X_i], G[Y_i])^T$. What should be noted is that the elements of W^n, unlike those of Z^n, are unobservable as the marginals are unknown and that by virtue of Remark 2 the elements of W^n are distributed according to the generating copula $\mathbb{C}$.

2.2. Semi-parametric model

The formulation of model (31), yields a semi-parametric model in the sense that it is non-parametric in the marginals $F \in \mathcal{F}$ and $G \in \mathcal{G}$ and parametric in the dependence parameter $\theta \in \Theta$. The log-likelihood function for the parameter $\eta = (\theta, F, G)^T \in \Theta \times \mathcal{F} \times \mathcal{G}$ given set of observed pairs, Z^n is given by

$$\begin{aligned} \ell^n[\theta, F, G] &= \ell^n[\theta|F, G] + \ell^n[F, G] \\ &= \sum_{i=1}^{n} \log c[F[X_i], G[Y_i], \theta] + \sum_{i=1}^{n} \log\{f[X_i]g[Y_i]\}, \end{aligned} \tag{34}$$

where

$$c[u, v, \theta] = \frac{\partial^2}{\partial u \partial v}\mathbb{C}[u, v, \theta], \tag{35}$$

is the density function of the distribution function $\mathbb{C}[u, v, \theta]$. Note that if the marginals F and G were known, then $\ell^n[\theta, F, G] \propto \ell^n[\theta|F, G]$ as $\ell^n[F, G]$ in (34) does not depend on θ. Consequently, in this case, the log-likelihood function would effectively be reduced to

$$\ell^n[\theta|F, G] = \sum_{i=1}^{n} \ell[\theta|F[X_i], G[Y_i]] = \sum_{i=1}^{n} \log c[F[X_i], G[Y_i], \theta]. \tag{36}$$

The MLE-estimator of θ, then could be canonically represented as

$$\hat{\theta}_n[F, G] = \underset{\theta \in \Theta}{\operatorname{argmax}} \;\; \ell^n[\theta|F, G], \tag{37}$$

or be equivalently obtained as the solution of the estimation equation

$$\Psi^n[\theta|F, G] = \frac{1}{n}\dot{\ell}^n[\theta|F, G] = \frac{1}{n}\sum_{i=1}^{n} \dot{\ell}[\theta|F[X_i], G[Y_i]], \tag{38}$$

where $\dot{\ell}[\theta, u, v] = \frac{\partial}{\partial \theta}\ell[\theta, u, v]$. As the summands in (38) are mutually independent and identically distributed, the asymptotic properties of the $\hat{\theta}_n[F, G]$, under standard regularity conditions, can be obtained from standard results for M-estimators. In particular, the M-estimator $\hat{\theta}_n[F, G]$, subject to certain regularity conditions, is consistent and is asymptotically normal in the sense that

$$\theta_n[F, G] \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} \theta\,, \tag{39}$$

and

$$\sqrt{n}(\theta_n[F, G] - \theta) \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} \mathbb{N}_1[0, \sigma^2_{\mathrm{M}}[\theta]]\,, \tag{40}$$

where

$$\sigma^2_{\mathrm{M}}[\theta] = \frac{\mathbb{V}[\dot{\ell}[\theta, F[X], G[Y]]]}{\mathbb{E}[\dot{\ell}[\theta, F[X], G[Y]]^2]^2} = \frac{\mathbb{V}[\dot{\ell}[\theta, U, V]]}{\mathbb{E}[\dot{\ell}[\theta, U, V]^2]^2}\,, \tag{41}$$

for a generic pair $(U, V)^T$ from the uniform copula $\mathbb{C}[u, v, \theta]$. Note that the above asymptotic variance depends explicitly on the dependence parameter θ but not on the marginals F and G. In practice, however, the marginals F and G, are unknown. What can, however, be done is to estimate the marginals F and G by a pair of empirical estimators, say $\hat{F}_n$ and $\hat{G}_n$ and then approximate $\ell^n[\theta|F, G]$ by $\ell^n[\theta|\hat{F}_n, \hat{G}_n]$. Analogous to (37), one then obtains a pseudo-ML estimator of θ as

$$\theta_n[\hat{F}_n, \hat{G}_n] = \underset{\theta\in\Theta}{\operatorname{argmax}}\, \ell^n[\theta|\hat{F}_n, \hat{G}_n]\,, \tag{42}$$

$$\Psi_n[\theta|\hat{F}_n, \hat{G}_n] = \frac{1}{n}\dot{\ell}^n[\theta|\hat{F}_n, \hat{G}_n] = \frac{1}{n}\sum_{i=1}^{n}\dot{\ell}[\theta|\hat{F}_n[X_i], \hat{G}_n[Y_i]]\,. \tag{43}$$

Note that the $\ell^n[\theta|\hat{F}_n, \hat{G}_n]$ is not necessarily a sum of independent terms as, for example, $\hat{F}_n[X_i]$ is not independent of $\hat{F}_n[X_j]$. Consequently, the standard results for M-estimators do not necessarily apply. The question that needs to be addressed is what kind of marginal estimators, $\hat{F}_n$ and $\hat{G}_n$, of F and G, under some additional assumptions, would preserve the asymptotic properties enjoyed by (37). To that end, we will consider a special class of marginal estimators as defined next.

Definition 7 [UCAN]**.** For $\mathbb{H}$ given in (31), let $\hat{F}_n$ and $\hat{G}_n$ denote a pair of marginal estimators for its marginals $F \in \mathcal{F}$ and $G \in \mathcal{G}$. The estimator $\hat{M}_n = (\hat{F}_n, \hat{G}_n)^T$ of $M = (F, G)^T$ is said to be UCAN (Uniformly Convergent and Asymptotically Normal) if

$$\sup_{z\in\mathcal{F}\times\mathcal{G}} |\hat{M}_n[z] - M[z]| \underset{n\to\infty}{\overset{\text{a.s.}}{\to}} (0, 0)^T \tag{44}$$

and

$$\sqrt{n}(\hat{M}_n[z] - M[z]) \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} \mathbb{N}_2[0, \Sigma_{\mathbb{H}}[z]] \tag{45}$$

for $z = (x, y)^T \in \mathbb{F} \times \mathbb{G}$. where

$$\begin{aligned}\Sigma_{\mathbb{H}}[z] &= \begin{pmatrix} M_{11}[x] & H[x,y] - M_{12}[z] \\ H[x,y] - M_{12}[z] & M_{22}[y] \end{pmatrix} \\ &= \begin{pmatrix} M_{11}[x] & \mathbb{C}[F[x], G[y], \theta] - M_{12}[z] \\ \mathbb{C}[F[x], G[y], \theta] - M_{12}[z] & M_{22}[y] \end{pmatrix},\end{aligned} \tag{46}$$

where $M_{11} = F(1 - F)$, $M_{22} = G(1 - G)$ and $M_{12} = FG$,

It is noted that the standard empirical distribution function and smoothed kernel estimators, under certain conditions as discussed in for example Yamato (1972), Winter (1979), Reiss (1981), Falk (1983) and Fernholz (1991) are UCAN. The abovementioned call of estimators will ensure that the resulting M-estimator of θ is consistent and asymptotically normal as stated in the following result.

Theorem 5. *Given that $\hat{F}^n$ is UCAN estimator of F, the plugin estimator $\theta_n[\hat{F}^n]$ is consistent and asymptotically normal. More specifically,*

$$\theta_n[\hat{F}_n, \hat{G}_n] - \theta \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0\,, \tag{47}$$

and

$$\sqrt{n}(\theta_n[\hat{F}_n, \hat{G}_n] - \theta) \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} \mathbb{N}_1[0, \sigma^2[\theta]]\,, \tag{48}$$

where

$$\begin{aligned}\sigma^2[\theta] &= \frac{\mathbb{V}[\dot{\ell}[\theta, U, V]]}{\mathbb{E}[\dot{\ell}[\theta, U, V]^2]^2} + \frac{\mathbb{V}[D_{10}[U|\theta] + D_{01}[V|\theta]]}{\mathbb{E}[\dot{\ell}[\theta, U, V]^2]^2}\,, \\ &= \sigma^2_{\mathrm{M}}[\theta] + \sigma^2_{\mathrm{P}}[\theta]\end{aligned} \tag{49}$$

where $(U, V)^T$ is a generic pair from the uniform copula $\mathbb{C}[u, v, \theta]$, and where

$$D_{10}[u^*|\theta] = \int_0^1 \int_0^1 \mathbb{I}[u^* \le u]\dot{\ell}_{10}[u, v, \theta]c[u, v, \theta]\mathrm{d}u\mathrm{d}v\,, \tag{50}$$

$$D_{01}[v^*|\theta] = \int_0^1 \int_0^1 \mathbb{I}[v^* \le v]\dot{\ell}_{01}[u, v, \theta]c[u, v, \theta]\mathrm{d}u\mathrm{d}v\,, \tag{51}$$

for all $(u^*, v^*) \in [0, 1]^2$, where $\dot{\ell}_{10}[u, v, \theta] = \frac{\partial}{\partial u}\dot{\ell}[u, v, \theta]$ and $\dot{\ell}_{01}[u, v, \theta] = \frac{\partial}{\partial u}\dot{\ell}[u, v, \theta]$ for all $(u, v) \in [0, 1]^2$.

The asymptotic variance is decomposed into the sum of $\sigma^2_{\mathrm{M}}[\theta]$, the asymptotic variance of the MLE estimator, for which the marginals were known, and $\sigma^2_{\mathrm{P}}[\theta]$ which can be thought of as the added variance induced due to plugging-in marginal estimators for the marginals F and G. This estimation method, in the special case where F and G are estimated by their respective standard marginal empirical distribution functions, is discussed in Genet *et al.* (1995) and Shih and Louis (1995). Two papers that consider estimation of copulas in the context of AC copulas are Genest and Rivest (1993) and Wang and Wells (2000).

2.3. Parametric model

A fully parametric counterpart of the copula model is obtained, if one were to make parametric assumptions on the marginals F and G. For the sake of completeness, we will provide a short discussion on estimating this parametric model. In particular, let us assume that F and G are indexed by λ_1 and λ_2 belonging to parameter spaces Λ_1 and Λ_2 respectively. It is noted that the likelihood function is expressible as

$$\ell_n[\theta, \lambda_1, \lambda_2] = \sum_{i=1}^{n} \log c[F_{\lambda_1}[X_i], G_{\lambda_2}[Y_i], \theta]] + \sum_{i=1}^{n} \log\{ f_{\lambda_1}[X_i] g_{\lambda_1}[Y_i]\} . \quad (52)$$

Given that the parameter space is finite-dimensional, one may attempt to estimate the parameter $\eta = (\theta, \lambda_1, \lambda_2)^T$ via MLE by

$$\hat{\eta}_n = \underset{\eta\in\Theta\times\Lambda_1\times\Lambda_2}{\operatorname{argmax}} \ \ell_n[\eta] \ \underset{\eta\in\Theta\times\Lambda_1\times\Lambda_2}{\operatorname{argmax}} \ \ell_n[\theta, \lambda_1, \lambda_2] . \quad (53)$$

Alternatively, one may employ an approach similar to that discussed for the semi-parametric case. Specifically, given that the marginals are known, which in this case is equivalent to having the parameters λ_1 and λ_2 known, the log-likelihood reduces to

$$\ell_n[\theta|\lambda_1, \lambda_2] \sim \sum_{i=1}^{n} \log c[F_{\lambda_1}[X_i], G_{\lambda_2}[Y_i], \theta]] , \quad (54)$$

from which one-step plug-in estimator of θ is defined explicitly as

$$\theta_n[\lambda_1, \lambda_2] = \underset{\theta\in\Theta}{\operatorname{argmax}} \, \ell_n[\theta|\lambda] , \quad (55)$$

or implicitly as the solution of the estimating equation

$$\Psi^n[\theta|\lambda_1, \lambda_2] = \frac{1}{n} \dot{\ell}^n[\theta|\lambda_1, \lambda_2] = \frac{1}{n} \sum_{i=1}^{n} \dot{\ell}[\theta|\lambda_1, \lambda_2] . \quad (56)$$

Analogous, to the approach taken in the semi-parametric case, the marginal parameters are the estimated using, for example maximum likelihood, and then are substituted into (54) to yield an estimator of θ and the

$$\theta_n[\hat{\lambda}_1, \hat{\lambda}_2] = \operatorname*{argmax}_{\theta\in\Theta} \ell_n[\theta|\hat{\lambda}_n]\,. \tag{57}$$

A comprehensive treatment of this fully parametric model is given in Joe (1997). Shih and Louis (1995) also provide a discussion on this model.

2.4. Additional remarks

One can consider variations of the types of models discussed above by looking at estimation methods given that for example the two marginals are modeled under a location-shift context. Also, one may consider estimation methods assuming that the marginals are equivalent or if one is stochastically larger than the other.

3. Additional topics and general applications

3.1. Empirical copula

Given a random sample

$$Z^n = \{(X_1, Y_1)^T, \dots, (X_n, Y_n)^T\}\,, \tag{58}$$

from a distribution $\mathbb{H}$ with continuous marginals, we define the empirical copula as function on the lattice

$$I_n = \left\{\left(\frac{i}{n}, \frac{j}{n}\right) : i, j \in \{1, \dots, n\}\right\}\,, \tag{59}$$

defined as

$$\mathbb{C}_n\left[\frac{i}{n}, \frac{j}{n}\right] = \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}[X_k \le X_{n:i}, Y_k \le Y_{n:j}]\,, \tag{60}$$

where $X_{n:1}, \dots, X_{n:n}$ denote the order statistics of $X_1, \dots, X_n$. It is easy to show that

$$\mathbb{C}_n[u, v] = \mathbb{H}_n[\hat{F}_n^-[u], \hat{G}_n^-[u]]\,, \tag{61}$$

where $\hat{F}_n^-[u]$ and $\hat{G}_n^-[u]$ are the standard empirical distribution functions of F and G and where

$$\mathbb{H}_n[x, y] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}[X_i \le x, Y_i \le y]\,, \tag{62}$$

is the bivariate empirical distribution function of $\mathbb{H}$. It can be shown that $\mathbb{C}_n$ converges strongly to $\mathbb{C}$. Also, under some conditions, as for example listed in Lemma 3.9.28 in van der Vaart and Wellner (1996), the copula is, in suitable sense differentiable, ans as such $\sqrt{n}(\mathbb{C}_n - \mathbb{C})$ converges in law to a function of tight Brownian Bridge. This copula can be used to generate non-parametric tests for independence as discussed. See for example Deheuvels (1981a, b). It can be used to construct empirical counterparts to the measures of dependence (20)–(23). For example, we recall that Spearman's ρ in terms of copulas can be expressed as

$$\rho[\mathbb{H}] = 12 \int_0^1 \int_0^1 \{\mathbb{C}[u, v] - uv\} \mathrm{d}u \mathrm{d}v \,,$$

for which an empirical version, based on the empirical copula, is given by

$$\rho_n[H] = \frac{12}{n^2 - 1} \sum_{i=1}^{n} \sum_{j=1}^{n} \left\{ \mathbb{C}_n \left[\frac{i}{n}, \frac{j}{n} \right] - \frac{ij}{n^2} \right\} . \tag{63}$$

We conclude this section by the elaborating on the comments made in Remark 2.

Remark 5. Each pair $(X_i, Y_i)^T$ of Z^n is distributed according to $\mathbb{H}$, while each pair $(F[X_i], G[Y_i])^T$ is distributed according to the uniform copula $\mathbb{C}$. Given that $\hat{F}_n$ and $\hat{G}_n$ are the standard empirical distribution functions of F and G, the pair $(\hat{F}_n[X_i], \hat{G}_n[Y_i])^T$ is distributed according to the empirical copula $\mathbb{C}_n$.

3.2. Continuity and differentiability

Two important properties of copulas are that they are uniformly continuous and have bounded first-order partial derivatives.

Theorem 6. *Given is a copula* $\mathbb{C}$*.*

i. $\mathbb{C}$ *satisfies the Lipshitz condition*

$$|\mathbb{C}[u_1, v_1] - \mathbb{C}[u_2, v_2]| \leq |u_1 - v_1| - |u_2 - v_2| \,, \tag{64}$$

for $u_1, u_2, v_1, v_2 \in [0, 1]$.

ii. $\mathbb{C}$ *is uniformly continuous on* $[0, 1]^2$.

iii. *The partial derivatives*

$$\mathbb{C}_{10}[u, v] = \frac{\partial}{\partial u}\mathbb{C}[u, v] \in [0, 1] \;\; \textit{and} \;\; \mathbb{C}_{01}[u, v] = \frac{\partial}{\partial u}\mathbb{C}[u, v] \in [0, 1] \,, \tag{65}$$

exist for every pair $(u, v)^T \in [0, 1]^2$ *and our bounded.*

Both continuity and differentiability properties follow quite easily from the Lipshitz condition coupled with the axiomatic properties of a copula.

3.3. Conditional distribution and simulation

Given that (U, V) is a pair from the copula $\mathbb{C}$, then for every $u \in [0, 1]$, the conditional distribution of function of V given $U = u$ is given by

$$\mathbb{P}[V \le v|U = u] = \mathbb{C}_{10}[u, v] , \tag{66}$$

for every $v \in [0, 1]$. Analogously, one obtains ins the conditional distribution of U given V using $\mathbb{C}_{01}$ in (65). Consequently, given a random pair $(X, Y)^T$ from a distribution $\mathbb{H}$, assuming that the marginals are continuous, the conditional distribution of X given $Y = y \in \mathcal{G}$ can be expressed in terms of the conditional copula as

$$\mathbb{P}[X \le x|Y = y] = \mathbb{C}_{01}[F[x], G[y]] , \tag{67}$$

for $x \in \mathcal{F}$. For any copula $\mathbb{C}$, given that U^* and V^* are two independent uniform [0,1] variates, the pair

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} \mathbb{C}_{01}^{-}[U^*, V^*] \\ V^* \end{pmatrix} , \tag{68}$$

where $\mathbb{C}_{01}^{-1}$ denotes the quantile function of $\mathbb{C}_{01}$, is distributed according to the $\mathbb{C}$. One could analogously generate a random pair from $\mathbb{C}$, using the pair $(U, V)^T = (U^*, \mathbb{C}_{10}^{-}[U^*, V^*])^T$. Furthermore, given that $(X, Y)^T$ is a random pair from a distribution $\mathbb{H}$ with continuous marginals F and G, the pair

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} F^{-}[U] \\ G^{-}[V] \end{pmatrix} , \tag{69}$$

where F^- and G^- denote the inverse functions of F and G respectively. is distributed according to $\mathbb{H}$. One could analogously generate a random pair from $\mathbb{C}$, using the pair $(U, V)^T = (U^*, \mathbb{C}_{10}^{-}[U^*, V^*])^T$. We should point out, that with a few exceptions, the conditional copula $\mathbb{C}_{01}$ does not admit a closed-form inverse function. As such, the employment of this method will more often that not in practice necessitate the use of numerical inversion methods. Copulas and bivariate distributions, as for example discussed in Marshal and Olkin (1988, 1991), can be constructed by mixtures. This construction will yield and alternate method, discussed in the latter reference, for simulating from copulas which does not require numerical inversion. The monograph by Johnson (1987) is a general purpose reference on simulation for multivariate distributions. Also see Lee (1993). The conditional copula can be used to generate updated estimators of the marginals. One notes that

$$F[x] = \int_y \mathbb{P}[X \le x|Y = y] \mathrm{d}G[y] , \tag{70}$$

which can be equivalently presented, in terms of the conditional copula, as

$$F[x] = \int_y \mathbb{C}_{01}[x, y, \theta] \mathrm{d}G[y] \,. \tag{71}$$

What this representation suggests is that if one has a current estimate of the marginal G and the parameter θ, then one can obtain an updated estimate of the marginal F

$$F^{(1)}[x] = \int_y \mathbb{C}_{01}[x, y, \theta^{(0)}] \mathrm{d}G^{(0)} \,. \tag{72}$$

Similarly, one may obtain an updated estimate of the marginal G and subsequently obtain an updated estimate of θ based on $F^{(1)}$ and $G^{(1)}$. On a related topic Zheng and Klein (1994, 1995, 1996) propose an estimation routine for estimating the marginals in the context of a competing risk model.

3.4. Survival function

There is a sizeable amount of literature concerning the estimation, from non- and semi-parametric points of view, of the dependence and joint distribution of censored survival variables. Some papers on the estimation of the dependence, in a semi-parametric setting, are Clayton (1978), Oakes (1982, 1986, 1994), O' Quigley and Prentice (1991), Jung *et al.* (1995) and Shih and Louis (1995, 1996). Dabrowska (1988, 1989), Prentice and Cai (1992), Oakes (1994) and Gill *et al.* (1995) constitute a list of representative papers. A comprehensive treatment of issues concerning estimation for multivariate survival functions, including material on copulas, is presented in a monograph by Hougaard (2000). Parametric copulas can be used not only to model the dependence between censored survival variables but also to estimate the bivariate survival function. The latter can be accomplished by substituting appropriate smooth estimators for the marginal survivals. The relationship between the copula generated by a distribution function, $\mathbb{H}$, and that generated by its corresponding survival function $\bar{\mathbb{H}}$ can be seen from the following:

$$\begin{aligned} \bar{\mathbb{H}}[x, y] &= 1 - F[x] + G[y] + \mathbb{H}[x, y] \\ &= 1 - F[x] - G[y] + \mathbb{C}[F[x], G[y]] \\ &= \mathbb{C}_S[F[x], G[y]] \\ &= \mathbb{C}_S[1 - \overline{F}[x], 1 - \overline{G}[y]] \end{aligned} \tag{73}$$

for $(x, y)^T \in \mathcal{F} \times \mathcal{G}$, where $\overline{F} = 1 - F$ and $\overline{G} = 1 - G$ are the marginal survival functions corresponding to F and G respectively. It is easy to show that the function

$$\mathbb{C}_S[u, v] = 1 - u - v + \mathbb{C}[u, v] \,, \tag{74}$$

in the above expression, is by virtue of $\mathbb{C}$ being a copula, a copula itself and will be referred to as the associated (with $\mathbb{C}$) survival copula.

Remark 6. The converse to Skylar's result stipulates that if $\mathbb{C}$ is a copula and F and G are continuous marginal distribution functions, the function defined as $\mathbb{H}[x, y] = \mathbb{C}[F[x], G[y]]$ is a bivariate distribution function. Analogously, it can be shown that for the same copula $\mathbb{C}$, the function defined as $\mathbb{S}[x, y] = \mathbb{C}[1 - F[x], 1 - G[y]]$ is a bivariate survival function. One should, however, note that $\bar{H}[x, y] = \mathbb{C}_S[1-\overline{F}[x], 1-\overline{G}[y]]$, as shown in (73), and not $\mathbb{S}$ is the survival distribution corresponding to $\mathbb{H}$.

What this also suggests is that when modeling bivariate survival data, it may be more convenient to express the joint survival function directly as a copula of its marginal survival functions. More specifically, we will assume that our generic pair of survival times, denoted by $(X^{'}, Y^{'})^T$ follows a joint, survival function $\bar{\mathbb{H}}$ expressed as

$$\bar{\mathbb{H}}[x, y] = \mathbb{C}[\overline{F}[x], \overline{G}[y], \theta], \tag{75}$$

for some copula $\mathbb{C}$. Assuming that the survival times for subject i, $(X_i^{'}, Y_i^{'})^T$, are subject to some independent right-censoring mechanism, what is observed is $Z_i = (X_i, Y_i)^T = (X_i^{'} \wedge C_i^X, Y_i^{'} \wedge C_i^Y)^T$, where $(C^X, C^Y)^T$ are the so called censoring variables and $\Delta_i = (\Delta_i^X, \Delta_i^Y)^T = (\mathbb{I}[X_i \leq C_i^X], \mathbb{I}[Y_i \leq C_i^Y])^T$ are the so called corresponding event variables. In particular, given that the marginal survival functions $\overline{F}$ and $\overline{G}$ are known, the log-likelihood function is given the data

$$Z^n = \{(Z_1, \Delta_1)^T, \ldots, (Z_n, \Delta_n)^T\}, \tag{76}$$

is given by

$$\begin{aligned}
\ell_n[\theta|\overline{F}, \overline{G}] = &\sum_{i=1}^{n} \log[c[\overline{F}[X_i], \overline{G}[Y_i], \theta]]^{\Delta_i^X \Delta_i^Y} \\
&+ \sum_{i=1}^{n} \log[\mathbb{C}_{10}[\overline{F}[X_i], \overline{G}[Y_i], \theta]]^{\Delta_i^X \bar{\Delta}_i^Y} \\
&+ \sum_{i=1}^{n} \log[\mathbb{C}_{01}[\overline{F}[X_i], \overline{G}[Y_i], \theta]]^{\bar{\Delta}_i^X \Delta_i^Y} \\
&+ \sum_{i=1}^{n} \log[\mathbb{C}[\overline{F}[X_i], \overline{G}[Y_i], \theta]]^{\bar{\Delta}_i^X \bar{\Delta}_i^Y},
\end{aligned} \tag{77}$$

where $\bar{\Delta}_i^X = 1-\Delta_i^1$ and $\bar{\Delta}_i^Y = 1-\Delta_i^Y$. Plug-in M-estimation methods similar to those described earlier for the uncensored case. may be employed to estimate θ.

Shih and Louis (1995) provide some arguments for the asymptotic law of the estimator in the special case where the marginal survival functions, $\overline{F}$ and $\overline{G}$, are estimated by the standard Kaplan-Meier marginal estimators and consider a related case study.

3.5. Semi-parametric bootstrap

To employ the simulation techniques outlined in the context of parametric copulas, one needs to know the parameter θ as well as the marginals F and G. This luxury of this knowledge is not afforded in data analysis problems and as such, we will be able to draw random pairs from neither the uniform copula $\mathbb{C}$ nor the target distribution function $\mathbb{H}$. As in most empirical simulation problems, we will need to settle for simulating from the estimated rather than actual distribution functions. Suppose that U^* and V^* are two independent uniform variates and that $\hat{\theta}_n, \hat{F}_n$ and $\hat{G}_n$ are estimators of θ, F and G, based on Z^n respectively. Then given Z^n, the pair

$$\begin{pmatrix} \tilde{U} \\ \tilde{V} \end{pmatrix} = \begin{pmatrix} \mathbb{C}_{01}^{-1}[U^*, V^*, \hat{\theta}_n] \\ V^* \end{pmatrix}, \tag{78}$$

is a distributed according to $\mathbb{C}[u, v, \hat{\theta}_n]$. If the marginal estimators are continuous, we may take this a step further by saying that given Z^n, the pair

$$\begin{pmatrix} \tilde{X} \\ \tilde{Y} \end{pmatrix} = \begin{pmatrix} \hat{F}_n^-[\tilde{U}] \\ \hat{G}_n^-[\tilde{V}] \end{pmatrix}, \tag{79}$$

is distributed according to $\hat{\mathbb{H}}[x, y] = \mathbb{C}[\hat{F}_n[x], \hat{G}_n[y], \hat{\theta}_n]$. Note that based on the observed sample Z^n we can now simulate a larger sample of size $N[n] > n$. More specifically, we can simulate a random sample

$$W^{*,N[n]} = \{(U_1^*, V_1^*)^T, \dots, (U_{N[n]}^*, V_{N[n]}^*)^T\}, \tag{80}$$

of size $N[n]$ from estimated uniform copula $\mathbb{C}[u, v, \hat{\theta}_n]$. Then given continuous estimators $\hat{F}_n$ and $\hat{G}_n$, of F and G, we can simulate a random sample

$$Z^{*,N[n]} = \{(X_1^*, Y_1^*)^T, \dots, (X_{N[n]}^*, Y_{N[n]}^*)^T\}, \tag{81}$$

where $(X_i^*, Y_i^*) = (\hat{F}_{U_i^*}^-[n], \hat{G}_{V_i^*}^-[n])$ for $i \in \{1, \dots, N[n]\}$, from the estimated distribution $\mathbb{C}[\hat{F}_n[x], \hat{G}_n[y], \hat{\theta}_n]$. This method can then be employed to simulate statistical functionals of $\mathbb{H}$. For example, $\mathbb{E}[h[X, Y]]$ based on Z^n can be simulated by

$$\mathbb{E}_{N[n]}^*[h[X, Y]] = \frac{1}{N[n]} \sum_{i=1}^{N[n]} h[X_i^*, Y_i^*]. \tag{82}$$

3.6. Minimum, maximum and extreme values

The distribution of the minimum and maximum, $\min[X, Y]$ and $\max[X, Y]$, under the copula model can be expressed as

$$\mathbb{P}[\min\{X, Y\} \leq z] = \mathbb{P}[X \leq z, Y \leq z] = \mathbb{C}[F[z], G[z]] , \tag{83}$$

and

$$\begin{aligned}\mathbb{P}[\max\{X, Y\} \leq z] &= 1 - \mathbb{P}[\max\{X, Y\} > z] \\ &= 1 - \mathbb{P}[X > z, Y > z] = 1 - \mathbb{C}_S[F[z], G[z]] .\end{aligned} \tag{84}$$

These expressions suggest that by estimating the copula, one obtains an estimate of the distributions of the minimum and maximum. There is a substantial amount of work in the area of bivariate extremes in the context of copulas. We will not elaborate on this issue but suggest Marshal and Olkin (1983), Capéraá *et al.* (1997) and Genest and Rivest (2001) as representative papers on this topic.

3.7. Goodness of fit

So far, it has been implicitly assumed that the copula function is known. In data analysis problems, one is faced with choosing the copula which generates the target distribution. The methods to be discussed are appropriate for selecting the best-fitting, based on some criteria, copula among a pool of potential candidate AC copulas. For a given distribution function $\mathbb{H}$, let use define

$$K[w] = \mathbb{P}[\mathbb{H}[x, y] \leq w]] = \int\int \mathbb{I}[\mathbb{H}[x, y] \leq u] \mathrm{d}\mathbb{H}[x, y] , \tag{85}$$

for each $w \in [0, 1]$. Furthermore, suppose that $\mathbb{H}$ is generated by a copula $\mathbb{C}$ parametrized by θ and note that

$$K_{\mathbb{C}}[w; \theta] = \int\int \mathbb{I}[\mathbb{C}[u, v, \theta] \leq u] \mathrm{d}u\mathrm{d}v = K[w] . \tag{86}$$

It can be shown, see for example Genest and Rivest (1993), that that if $\mathbb{C}$ is AC and generated by ϕ, that

$$K_{\mathbb{C}}[w; \theta] = w - \frac{\phi[w; \theta]}{\dot{\phi}[w; \theta]} , \tag{87}$$

where $\dot{\phi}[w] = \frac{\partial}{\partial w}\phi[w]$. We note that $K[w]$ in (85) can be estimated non-parametrically by say $\hat{K}_n[w]$ and that $K_{\mathbb{C}}[w; \theta]$, as the generator ϕ is assumed to be known, can be estimated by plugging in an estimator $\hat{\theta}_n$ of θ in (86). As such we can quantify the goodness of fit for the chosen copula by assessing the empirical discrepancy between $\hat{K}_n[w]$ and $K_{\mathbb{C}}[w;\hat{\theta}_n]$. Wang and Wells (2000), for example, suggest using the L^2 distance

$$\delta_2[\theta] = \int \{K[u] - K_{\mathbb{C},\theta}[u]\}^2 \mathrm{d}u \,. \tag{88}$$

The also generalize this idea in the context of bivariate censored variables. One can also employ a graphical approach by plotting for example $\hat{K}_n[w]$ versus $K_{\mathbb{C}}[w;\hat{\theta}_n]$. Also see Rivest and Wells (2001) for related issues.

3.8. Approximating copulas

In this section, we will present two methods for approximating copulas. For a given copula $\mathbb{C}$, its Bernstein approximation is given by

$$\mathbb{B}^n_{\mathbb{C}}[u,v] = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{C}\left[\frac{i}{n},\frac{j}{n}\right] b_n[u,i]b_n[v,j]\,, \tag{89}$$

where

$$b_n[w,k] = \binom{n}{k} w^k (1-w)^{n-k}\,, \tag{90}$$

for $k \in \{1,\ldots,n\}$ and $w \in [0,1]$, and its checkerboard approximation is given by

$$\mathbb{D}^n_{\mathbb{C}}[u,v] = n^2 \sum_{i=1}^{n}\sum_{j=1}^{n} \Delta^n_{\mathbb{C}}[i,j] d_n[u,i] d_n[v,j]\,, \tag{91}$$

where

$$\Delta^n_{\mathbb{C}}[i,j] = \mathbb{C}\left[\frac{i}{n},\frac{j}{n}\right] - \mathbb{C}\left[\frac{i-1}{n},\frac{j}{n}\right] - \mathbb{C}\left[\frac{i}{n},\frac{j-1}{n}\right] + \mathbb{C}\left[\frac{i-1}{n},\frac{j-1}{n}\right]\,, \tag{92}$$

for $i, j \in \{1,\ldots,n\}$ and

$$d_n[w,k] = \int_0^w \mathbb{I}\left[t \in \left[\frac{k-1}{n},\frac{k}{n}\right]\right] \mathrm{d}t\,, \tag{93}$$

for each $k \in \{1,\ldots,n\}$ and $w \in [0,1]$. It can be shown that $\mathbb{B}^n_{\mathbb{C}}$ and $\mathbb{D}^n_{\mathbb{C}}$ are copulas and that they converge uniformly to $\mathbb{C}$. For more detailed discussions on approximations of copulas, see Mikusinski *et al.* (1992), and Li *et al.* (1998) and Kulpa (1999).

3.9. *Efficient scoring*

The plug-in type estimators that were discussed are obviously, except in the case of independence, not fully efficient. Three references, which address the issue of efficient scoring for copulas, are Bickel *et al.* (1993) (Section 4.7), Maguluri (1993) and Klaassen and Wellner (1997). The first provides a comprehensive treatment on efficient estimation in the context of semi-parametric models. The latter derives the efficient score for the special case of the normal copula (1). The efficient score function is implicitly defined as a solution of coupled differential equations. Both references emphasize the point that although the solution to the corresponding differential equations may exist, in general it will be prohibitively difficult to derive in closed form. Two other comprehensive references on semi-parametrics, useful for studying theoretical properties of copulas, are the last chapter in van der Vaart (1998) and a series of lecture notes by van der Vaart on semi-parametric statistics in Bolthausen *et al.* (2002).

4. Applications

4.1. *Power and sample size considerations*

Power and sample size calculations are important tools in designing statistical experiments and trials. Often a small sample, say Z^n, of size n has made been available from a pilot-study to be utilized in designing the study. The idea is to estimate the distribution function $\mathbb{H}$ based on the observed sample Z^n and then simulate a larger sample, say $Z^{N[n]}$, of size $N[n] > n$ using the ideas discussed in Section 3.5. Note that this is similar, in nature, to the empirical power calculation methods discussed in Chapter 25 of Efron and Tibshirani (1993), where the joint distribution function $\mathbb{H}$ is estimated by the bivariate empirical distribution function based on the pilot data. The potential advantage of the copula based approach described here is that one is able to impose dependence structures on the data, rather than having the data completely impose the structure of the dependence. That may potentially be of great importance if for example the dependence is weak but due to the physical nature of the problem, the dependence structure must necessarily be non-negative. In this case, the imposition of a positive dependence structure may be more efficient.

4.2. *Microarrays*

The development of statistical methodology for the analysis of microarray data has gained great momentum in the last few years. We have already discussed the utility of copulas in quantification as well as estimation of dependent censored survival variables. The goal of this section is to illustrate the utility of copulas in detecting gene markers which are associated with a censored response such as a survival variable. In this setup, we have an $n \times M$ matrix $\mathbb{Z}^{n,M} = [Z^{n,1}|\ldots|Z^{n,M}]$ (n subjects and M markers) of gene array expressions. Gene i is said to be a prognostic marker if its corresponding gene expression is associated with the survival time T. Suppose that the joint survival function of T and Z^i, where Z^i is a generic gene expression for gene i, is given by

$$\bar{\mathbb{H}}_i[t, x] = \mathbb{C}[S_T[t], S_i[z], \theta_i]\,, \tag{94}$$

where S_T and S_i are the marginal survival functions of T and Z^i respectively. Given that the survival times are Y^n, we may estimate θ_i by Y^n and $Z^{n,i}$ using the likelihood (77). It should be noted that the likelihood will have a simpler form in this case as only the survival times and not the expressions are subject to censoring. For notational simplicity, we will assume that marker i is not associated with the survival time if $\theta_i = \theta_0$. That is if $\theta_i = \theta_0$, then $\mathbb{C}_i = \mathbb{C}_{\mathrm{I}}$. The hypotheses of interested can be canonically presented as testing $H_0 : \theta_i = \theta_0$ for all $i \in \{1, \ldots, M\}$ versus $H_1 : \theta_i \neq \theta_0$ for some $i \in \{1, \ldots, M\}$. Let $\hat{\theta}_n^i$ denote the estimator corresponding to θ_1 and let

$$\hat{\xi}^{M,n} = \max\{|\hat{\theta}_n^1|, \ldots, |\hat{\theta}_n^M|\}\,. \tag{95}$$

We will reject H_0 in favor of H_1 if $\hat{\xi}^{M,n}$ is large. More specifically, given a family-wise error rate of size $\alpha \in (0, 1)$, we want to find a critical value ξ_α such that $\mathbb{P}[\hat{\xi}^{M,n} > \xi_\alpha|H_0] \leq \alpha$. Note that under the null hypothesis, none of the markers are associated with the survival variable, whence the distribution of $\hat{\xi}^{M,n}$, under H_0, can be generated via permutation resampling by simply permuting the rows on Y^n. As $n!$, the number of possible permutations, is in most practical applications, prohibitively large, we will typically settle for B permutations. Let ξ_α^B denote the approximation of the critical value ξ_α based on B permutations. Then gene i is declared to be associated with the survival endpoint, if $|\hat{\theta}_n^i| > \xi_\alpha^B$. In this setup, we have assumed that the generating copula is the identical for all the gene markers. We may generalize this by considering having different copulas.

4.3. Multivariate receiver operator curves

Consider a classifier that assigns a subject as type A if some corresponding measurement, say X, exceeds a certain threshold say τ and as type B otherwise. Receiver Operator Curves (ROC) have shown to be very useful and have enjoyed broad popularity in assessing the performance of such classifiers. An illustrative example is that of a biomarker such as PSA (Prostate-Specific Antigen) whose large values are often thought to be indicative of prostate cancer. Swets and Pickett (1982) lists over 100 potential applications while Pepe (2000) provides a nice review on this topic. Given that S^A and S^B denote the survival function of X under populations A and B respectively, ROC curve is formally defined as

$$\rho_1[u] = S_1^A[\zeta_1^B[u]]\,, \tag{96}$$

for $u \in [0, 1]$, where $\zeta_1^B[u]$ is the inverse function of S_1^B An aggregate measure for the assessment of the performance of the test is the Area under the ROC (AROC) defined as

$$\alpha_1 = \int_0^1 \rho_1[u]\mathrm{d}u\,. \tag{97}$$

It is easily seen, the details are omitted here, that this quantity is equivalent to $\mathbb{P}[X^A > X^B]$, where X^A and X^B are generic measurements from populations A and B respectively. A standard non-parametric estimator, for this quantity, is the Mann-Whitney statistic (see Sidak *et al.* (1999) for more details). As such, the AROC is often useful in addressing questions about stochastic ordering between two distributions. Given that we have two random samples, $X_1^A, \ldots, X_{n_A}^A$ and $X_1^B, \ldots, X_{n_B}^B$, of sizes n_A and n_B, from populations A and B respectively, the non-parametric estimator of $\rho_1[\tau_1]$ is given by

$$\rho_1[u] = S_1^A[\zeta_1^B[u]]\,, \tag{98}$$

$$\hat{\rho}_1[\tau_1] = S_n^1[\zeta_n^1[\tau_1]]\,, \tag{99}$$

where S_n^1 is the standard empirical estimator of S^1 based on $X_1^A, \ldots, X_{n_A}^A$ and ζ_n^1 is the marginal quantile estimator based on $X_1^B, \ldots, X_{n_B}^B$. α_1 is given by

$$\hat{\alpha}_1 = \frac{1}{n_A}\frac{1}{n_B}\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\mathbb{I}[X_i^A > X_j^B]\,. \tag{100}$$

The pair of measurements $(X, Y)^T$ is assumed to follow the survival function

$$\bar{\mathbb{H}}^A[x, y] = \mathbb{C}[S_1^A[x], S_2^A[y], \theta]\,, \tag{101}$$

under population A while following the survival function

$$\bar{\mathbb{H}}^B[x, y] = \mathbb{C}[S_1^B[x], S_2^B[y], \theta]\,, \tag{102}$$

under population B. What should be noted, under the specified model, is that the measurements $(X, Y)^T$ have potentially different marginal effects while having identical dependence structures under the two populations. This can be generalized by using different dependence parameters or by using different copulas altogether. The ROC vector is then defined as

$$\rho[u, v] = \begin{pmatrix} S_1^A[\zeta_1^B[u]] \\ S_2^A[\zeta_2^B[v]] \end{pmatrix} , \tag{103}$$

for $(u, v) \in [0, 1]^2$ and its corresponding vector of AROCs is given by

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \displaystyle\int_0^1 S_1^A[\zeta_1^B[u]]\mathrm{d}u \\ \displaystyle\int_0^1 S_2^A[\zeta_2^B[v]]\mathrm{d}v \end{pmatrix} . \tag{104}$$

The covariance matrix of $\hat{\alpha}$ under (101) and (102) is now easily obtained. By estimating the parameters of (101) and (102), one obtains a plug-in estimates of these covariance matrices. The method based on multivariate U-statistic outlined in DeLong *et al.* (1988) can be considered as a non-parametric counterpart to the semi-parametric copula based approached presented In some applications with two or more markers, one marker is of primary interest while the remaining markers are of secondary interest or peripheral. In this case, we may consider the performance of the primary marker conditional on the secondary marker(s). The conditional distribution, in the context of the copula model, can be easily estimated to produce Conditional ROC curves (CROC) as well as Conditional AROCs (CAROC).

4.4. Applications to problems is finance, insurance and risk management

The applications presented thus far may give the erroneous impression that the utility of copulas is limited to biostatistics. Frees and Valdez (1998), in addition to providing a review of copulas, present applications to insurance problems. Wang (1998), Li (2000) and Lauprete *et al.* (2002) discuss applications of copulas to financial and insurance problems. A series of working papers by researcher at Credit Lyonnais, with applications of copulas to problems in finance and credit risk management, are available for download (http://gro.creditlyonnais.fr/content/rd/home_copulas.htm).

4.5. Software

An open-source library, for use within GNU R (www.r-project.org), is currently in development by the authors. This library will offer general facilities for simulation and estimation for a number of standard families of bivariate copulas. More specifically, this library will also be useful in the context of the biostatistical applications discussed in this paper. Other non-commercial copula related software is available for download from the internet. In the commercial domain, S+FinMetrics®, an addon module for S-PLUS® (www.insightful.com), according to company literature, provides functionality for the simulation and estimation of parametric and empirical copulas. This add-on module, as its name suggests is primarily intended for financial applications. The MODEL® procedure in SAS® (www.sas.com), according to its online user manual, utilizes copulas for simulating from multivariate distributions.

5. Concluding remarks

As pointed out in the introduction and as suggested by the title of this paper, the goal set forth was to present a review of the concept of copulas vis-a-vis multivariate distribution functions and statistical dependence along with a list of the early as well as recent relevant literature. As this field has been growing and developing at a fast rate, the list of relevant literature is by no means exhaustive but is rather intended to be representative. To keep the presentation focused towards applications, we have often refrained, as much as possible, from providing technical details as most of these are well documented the literature. More importantly, our hope is that the number of applications mentioned in this presentation will demonstrate the utility of copulas in engaging problems in applications for which the appropriate modeling of the dependence structure is paramount.

Acknowledgments

The authors are grateful to a referee and an associate editor for helpful suggestions and comments.

References

Bickel, P. J., Klaassen, C. A. J., Ritov, Y., and Wellner, J. A. *Efficient and Adaptive Estimation for Semiparametric Models*, Johns Hopkins University Press.

Bolthausen, E., Perkins, E., and van der Vaart, A. W. (2002) *Lecture Notes in Probability and Statistics*, vol. 1781 *of Lecture Notes in Mathematics. Semiparametric Statistics*, Springer-Verlag.

Capéraá, P., Fougéres, A. L., and Genest, C. (1997) A nonparametric estimation procedure for bivariate extreme value copulas, *Biometrika*, 84, 567-577.

Cifarelli, D. M., Conti, P. L., and Regazzini, E. (1996) On the Asymptotic Distribution of a General Measure of Monotone Dependence, *The Annals of Statistics*, 24, 1386-1399, CLAYTON, D. G. (1978) A Model for Association in Bivariate Life Tables and Its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence, *Biometrika*, 65, 141-152.

DABROWSKA, D. M. (1988) Kaplan-Meier Estimate on the Plane, *The Annals of Statistics*, 16, 1475-1489.

DABROWSKA, D. M. (1989) Kaplan-Meier Estimate on the Plane: Weak Convergence, LIL, and the Bootstrap, *Journal of Multivariate Analysis*, 29, 308-325.

DEHEUVELS, P. (1981a) A Kolmogorov-Smirnov type test for independence and multivariate samples (STMA V24 1687), *Revue Roumaine de Mathematiques Pures et Appliquees*, 26, 213-226.

DEHEUVELS, P. (1981b) Test for multivariate independence, In *Topics in Applied Statistics-Proceedings of Statistics' 81 Canada Conference*, 569-580.

DELONG, E. R., DELONG, D. M., and CLARKE-PEARSON, D. L. (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach, *Biometrics*, 44, 837-845.

EFRON, B. and TIBSHIRANI, R. J. (1993) *An Introduction to the Bootstrap*, Chapman & Hall.

FALK, M. (1983) Relative efficiency and deficiency of kernel type estimators of smooth distribution functions, *Statistica Neerlandica*, 37, 73-83.

FERNHOLZ, L. T. (1991) Almost sure convergence of smoothed empirical distribution functions, *Scandinavian Journal of Statistics*, 18, 255-262.

FISHER, N. I. (1997) Copulas, In S. Kotz, C. B. Read, and D. L. Banks, eds. *Encyclopedia of Statistical Sciences*, John Wiley & Sons.

FISHER, N. I. and SEN, P. K. (ED) *The Collected Works of Wassily Hoeffding*, Springer-Verlag.

FISHER, N. I. and SWITZER, P. (2001) Graphical Assessment of Dependence: Is a Picture Worth 100 Tests?, *The American Statistician*, 55, 3, 233-239.

FRANK, M. J. (1979) On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$, *Aeqationes Mathematicae*, 19.

FREES, E. W. and VALDEZ, E. A. (1998) Understanding relationships using copulas, *North American Acturial Journal*, 2 (1), 1-25

GENEST, C., GHOUDI, K., and RIVEST, L. P. (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika*, 82, 543-552.

GENEST, C. and RIVEST, L. P. (1993) Statistical Inference Procedures for Bivariate Archimedean Copulas, *Journal of the American Statistical Association*, 88, 1034-1043.

GENEST, C. and RIVEST, L. P. (2001) On the multivariate probability integral transformation, *Statistics and Probability Letters*, 53 (4), 391-399.

GILL, R. D., VAN DER LAAN, M. J., and WELLNER, J. A. (1995) Inefficient Estimators of the Bivariate Survival Function for Three Models, *Annales de l'Institut Henri Poincare, Section B, Calcul des Probabilities et Statistique*, 31, 545-597.

GUMBEL, E. J. (1960) Bivariate Exponential Distributions, *Journal of the American Statistical Association*, 55, 698-707.

GUMBEL, E. J. (1961) Bivariate Logistic Distributions, *Journal of the American Statistical Association*, 56, 335-349.

HOUGAARD, P. (2000) *Analysis of Multivariate Survival Data*, Springer-Verlag.

JOE, H. (1997) *Multivariate Models and Dependence Concepts*, Chapman & Hall.

JOHNSON, M. E. (1987) *Multivariate Statistical Simulation*, Wiley.

JUNG, S. H., WIEAND, S., and CHA, S. S. (1995) A statistic for comparing two correlated markers which are prognostic for time to an event, *Statistics in Medicine*, 14, 2217-2225.

KLAASSEN, C. A. J. and WELLNER, J. A. (1997) Efficient Estimation in the Bivariate Normal Copula Model: Normal Margins Are Least Favourable, *Bernoulli*, 3, 55-77.

KUPLA, T. (1999) On approximation of copulas, *International Journal of Mathematics and Mathematical Sciences*, 22, 259-269.

LAUPRETE, G. J., SAMAROV, A. M., and WELSCH, R. E. (2002) Robust Portfolio Optimization, *Metrika*, 55 (2), 139-149.

LEE, A. J. (1993) Generating random binary deviates having fixed marginal distributions and specified degrees of association, *The American Statistician*, 47, 209-215.

LI, D. X. (2000) On default correlation: a copula function approach, *Journal of Fixed Income*, 9, 43-54.

LI, X., MIKUSIŃSKI, P., and TAYLOR, M. D. (1998) Strong Approximations of Copulas, *Journal of Mathematical Analysis and Applications*, 225, 608-623.

MAGULURI, G. (1993) Semiparametric estimation of association in a bivariate survival Function, *The Annals of Statistics*, 21, 1648-1662.

MARDIA, K. V. (1970) *Families of Bivariate Distributions.* In *Griffin's Statistical Monographs and Courses* No. 27, Charles Griffin.

MARSHALL, A. W. and OLKIN, I. (1983) Domains of Attraction of Multivariate Extreme Value Distributions, *The Annals of Probability*, 11, 168-177.

MARSHALL, A. W. and OLKIN, I. (1988) Families of multivariate distributions, *Journal of the American Statistical Association*, 83, 834-841.

MARSHALL, A. W. and OLKIN, I. (1991) Multivariate Distributions Generated from Mixtures of Convolution and Product Families, In Henry W. Block, Allan R., Sampson, and Savits, Thomas H. eds. *Topics in Statistical Dependence*, Institute of Mathematical Statistics.

MIKUSINSKI, P., SHERWOOD, H., and TAYLOR, M. D. (1992) Shuffles of min, *Stochastica*, 13, 61-74.

NELSEN, R. (1998) *An Introduction to Copulas*, Springer-Verlag.

OAKES, D. (1982) A Model for Association in Bivariate Survival Data, *Journal of the Royal Statistical Society, Series B, Methodological*, 44, 414-422.

OAKES, D. (1986) Semiparametric Inference in a Model for Association in Bivariate Survival Data, *Biometrika*, 73, 353-361.

OAKES, D. (1994) Multivariate Survival Distributions, *Journal of Nonparametric Statistics*, 3, 343-354.

O'QUIGLEY, J. and PRENTICE, R. L. (1991) Nonparametric tests of association between survival time and continuously measured covariates: The logit-rank and associated procedures, *Biometrics*, 47, 117-127.

PEPE, M. S. (2000) Receiver operating characteristic methodology, *Journal of the American Statistical Association*, 95, 449, 308-311.

PRENTICE, R. L. and CAI, J. (1992) Covariance and Survivor Function Estimation Using Censored Multivariate Failure Time Data, *Biometrika*, 80, 711-712.

REISS, R. D. (1981) Nonparametric estimation of smooth distribution functions, *Scandinavian Journal of Stratistics*, 8, 116-119.

RIVEST, L. P. and WELLS, M. T. (2001) A Martingale Approach to the Copula-graphic Estimator for the Survival Function Under Dependent Censoring, *Journal of Multivariate Analysis*, 79 (1), 138-155.

SCARSINI, M. (1984) On measure of concordance, *Stochatica*, 8, 201-218.

SCHWEIZER, B. (1991) *Advances in Probability Distributions With Given Marginals,* Vol. 67 of *Mathematics and its applications, Thirty years of copula*, Kluwer Academic Publishers.

SCHWEIZER, B. and SKLAR, A. (1983) *Probabilistic metric spaces*, Elsevier, North-Holland.

SCHWEIZER, B. and WOLFF, E. F. (1981) On nonparametric measures of dependence for random variables, *The Annals of Statistics*, 9, 879-885.

SHIH, J. H. AND LOUIS, T. A. (1995) Inferences on the association parameter in copula models for bivariate survival data, *Biometrics*, 51, 1384-1399.

SHIH, J. H. and LOUIS, T. A. (1996) Tests of Independence for Bivariate Survival Data, *Biometrics*, 52, 1440-1449.

SIDAK, Z., SEN, P. K., and HAJEK, J. (1999) *Theory of Rank Tests*, Academic Press, second edition.

SKLAR, A. (1959) Fonctions de repartition a n dimensions et leures marges, *Publ. Inst. Statist. Univ. Paris*, 8, 229-231.

SKLAR, A. (1996) *Random Variables, Distribution Functions, and Copulas – A Personal Look Backward and Forward.* In *Distributions with Fixed Marginals and Related Topics (IMS Lecture Notes Monograph Series*, Volume 28*)*, Institute of Mathematical Statistics (Hayward).

SWETS, J. A. and PICKETT, R. M. (1982) *Evaluation of diagnostic systems: Methods from signal detection theory*, Academic Press.

VAN DER VAART, A. W. (1998) *Asymptotic Statistics*, Cambridge University Press.

VAN DER VAART, A. W. and WELLNER, J. A. (1996) *Weak convergence and empirical processes: with applications to statistics*, Springer-Verlag.

WANG, S. (1998) Aggregation of correlated risk portfolios: model and algorithms, *Proceedings of the Casualty Actuarial Society*, LXXXV, 848-938.

WANG, W. and WELLS, M. T. (2000) Model selection and semiparametric inference for bivariate failure-time data, *Journal of the American Statistical Association*, 95, 449, 73-76.

WINTER, B. B. (1979) Convergence Rate of Perturbed Empirical Distribution Functions, *Journal of Applied Probability*, 16, 163-173.

YAMATO, H. (1973) Uniform Convergence of an Estimator of a Distribution Function, *Bull. Math. Statistics*, 15, 69-78.

ZHENG, M. and KLEIN, J. P. (1994) A Self-consistent Estimator of Marginal Survival Functions Based on Dependent Competing Risk Data and An Assumed Copula, *Communications in Statistics, Part A – Theory and Methods*, 23, 2299-2311.

ZHENG, M. and KLEIN, J. P. (1995) Estimates of Marginal Survival for Dependent Competing Risks Based on An Assumed Copula, *Biometrika*, 82, 127-138.

ZHENG, M. and KLEIN, J. P. (1996) *Identifiability and Estimation of Marginal Survival Functions for Dependent Competing Risks Assuming the Copula Is Known.* In *Lifetime Data Models in Reliability and Survival Analysis*, Kluwer Academic (Dordrecht; Norwell, MA).

KOUROS OWZAR
Department of Biostatistics and Bioinformatics
Duke University Medical Center
Durham, NC (USA)
kouros.owzar@duke.edu

PRANAB KUMAR SEN
Departments of Statistics and Biostatistics
University of North Carolina at Chapel Hill
Chapel Hill, NC (USA)
pksen@bios.unc.edu