Multivariate survival modelling: a unified approach with copulas

P. Georges, A-G. Lamy, E. Nicolas, G. Quibel & T. Roncalli*

Groupe de Recherche Opérationnelle
Crédit Lyonnais
France

May 28, 2001

Abstract

In this paper, we review the use of copulas for multivariate survival modelling. In particular, we study properties of survival copulas and discuss the dependence measures associated to this construction. Then, we consider the problem of competing risks. We derive the distribution of the failure time and order statistics. After having presented statistical inference, we finally provide financial applications which concern the life time value, the link between default, prepayment and credit lifetime, the measure of risk for a credit portfolio and the pricing of credit derivatives.

Contents

* *Corresponding author:* Groupe de Recherche Opérationnelle, Bercy-Expo — Immeuble Bercy Sud — 4$^{\text{è}}$ étage, 90 quai de Bercy — 75613 Paris Cedex 12 — France; *E-mail adress:* thierry.roncalli@creditlyonnais.fr

1 Introduction

The overall purpose of this paper is to present a copula approach to multivariate survival modelling. This is not the first time that copulas are used in survival analysis. They appear implicitly in Clayton [1978], Hougaard [1986a,1986b], Marshall and Olkin [1988] and Heckman and Honoré [1989] or more explicitly in Oakes [1989], Bagdonavicius, Malov and Nikulin [1998,1999], Shih and Louis [1995] and Bandeen-Roche and Liang [1996]. In what follows, we try to unify these different works and to propose a systematic approach based on survival copulas. We will see that copulas provide a general framework, that could encompass many models generally presented without links between them (Hougaard [1987]).

What is the fundamental difference between unvariate and multivariate survival data? As Hougaard [2000] says, "*the term multivariate survival data covers the field where independence between survival times cannot be assumed*". We have also to specify the joint distribution of the survival times or the corresponding multivariate survival function. In general, it is done in two steps. First, we consider the univariate data separately in order to characterize the specific properties of the survival times. Then, we search to describe the joint behaviour of the survival times by taking into account the properties exhibited in the first step. Copulas are also a natural tool for constructing families of multivariate *survival function with given margins*. For example, Clayton [1978] is one of the first to propose a bivariate association model for survival analysis. Without knowing the concept of copulas, his model assumes a copula.

The paper is organized as follows. In section two, we study properties of survival copulas and discuss the dependence measures associated to this construction. Then, we consider the modelling of competing risks. In particular, we derive the distribution of the failure time and other order statistics. Section fourth presents statistical inference. Finally in section five, we provide different applications in finance.

2 Multivariate survival models and copulas

2.1 Some definitions

Let T denote a *survival time* with distribution $\mathbf{F}$. The *survival function* is given by $\mathbf{S}(t) = \Pr\{T > t\} = 1 - \mathbf{F}(t)$. It is also known as the *survival rate*. The density f of T is then equal to $\partial_1 \mathbf{F}$ or equivalently $-\partial_1 \mathbf{S}(t)$. Generally in survival modelling, one of the main concepts is the *hazard rate* or *risk function* $\lambda(t)$ defined as follows (Lancaster [1990])

$$\lambda(t) = \lim_{\Delta \to 0^+} \frac{1}{\Delta} \Pr\{t \leq T \leq t + \Delta \mid T \geq t\} \tag{1}$$

This function can be interpreted as the instantaneous failure rate assuming the system has survived to time t. Another expressions of λ are

$$\begin{aligned} \lambda(t) &= -\frac{\partial_1 \mathbf{S}(t)}{\mathbf{S}(t)} \\ &= \frac{f(t)}{\mathbf{S}(t)} \end{aligned} \tag{2}$$

Let us define the *hazard function* $\Lambda(t)$ as the integral of λ between 0 and t

$$\Lambda(t) = \int_0^t \lambda(s)\, \mathrm{d}s \tag{3}$$

Λ is known under different names: the *cumultative hazard function* (Andersen, Borgan, Gill and Keiding [1993]), the *integrated hazard function* (Hougaard [1999]), etc. The link between Λ and $\mathbf{S}$ is done by the following relationship

$$\mathbf{S}(t) = \exp(-\Lambda(t)) \tag{4}$$

Another important concept is the "baseline" hazard function $\lambda_0(t)$ (Frees and Valdez [1998]). Cox proportional hazard model is generally used to incorporate explanatory variables X in survival distributions (Cox [1972]). The hazard rate takes then the following expression

$$\lambda(t) = \exp\left(\beta^\top X\right) \lambda_0(t) \tag{5}$$

Moreover, we note $\Lambda_0(t) = \int_0^t \lambda_0(s)\, \mathrm{d}s$ and $\boldsymbol{\chi}(t) = \exp(-\Lambda_0(t))$.

We extend now the previous definitions to the multivariate case. The multivariate survival function $\mathbf{S}(t)$ is defined by

$$\mathbf{S}(t_1, \ldots, t_N) = \Pr\{T_1 > t_1, \ldots, T_N > t_N\} \tag{6}$$

where $T_1, \ldots, T_N$ are N survival times with univariate survival functions $\mathbf{S}_n(t_n)$. Of course, we have[1]

$$\begin{aligned} \mathbf{S}_n(t_n) &= \Pr\{T_n > t_n\} \\ &= \Pr\{T_1 \geq 0, \ldots, T_{n-1} \geq 0, T_n > t_n, T_{n+1} \geq 0, \ldots, T_N \geq 0\} \\ &= \mathbf{S}(0, \ldots, 0, t_n, 0, \ldots, 0) \end{aligned} \tag{8}$$

[1] In this paper, we assume that the survival times are continuous and take their values in $\mathbb{R}^+$. Generally, a distribution function is defined by

$$\mathbf{F}(t) = \Pr\{T \leq t\} \tag{7}$$

This explains the definition adopted here for the survival function $\mathbf{S}(t) = 1 - \mathbf{F}(t) = \Pr\{T > t\}$. However, we can also adopt the following definition $\mathbf{S}(t) = \Pr\{T \geq t\}$.

We note that the relationship between the multivariate survival function $\mathbf{S}$ and the multivariate distribution function $\mathbf{F}$ is not direct as for the univariate case

$$\mathbf{S}(t_1,\ldots,t_N) \neq 1 - \mathbf{F}(t_1,\ldots,t_N) \tag{9}$$

If the survival function $\mathbf{S}$ is absolutely continuous, the joint density has the following expression

$$f(t_1,\ldots,t_N) = \partial_{1,\ldots,N}\mathbf{F}(t_1,\ldots,t_N) = (-1)^N \partial_{1,\ldots,N}\mathbf{S}(t_1,\ldots,t_N) \tag{10}$$

Multivariate expressions of the hazard rate and the hazard function are given by

$$\begin{aligned}\lambda(t_1,\ldots,t_N) &= \lim_{\max\Delta_n\to 0^+}\frac{1}{\Delta_1\cdots\Delta_N}\Pr\{t_1\le T_1\le t_1+\Delta_1,\ldots,t_N\le T_N\le t_N+\Delta_N \mid T_1\ge t_1,\ldots,T_N\ge t_N\}\\ &= \frac{f(t_1,\ldots,t_N)}{\mathbf{S}(t_1,\ldots,t_N)}\\ &= (-1)^N\frac{\partial_{1,\ldots,N}\mathbf{S}(t_1,\ldots,t_N)}{\mathbf{S}(t_1,\ldots,t_N)}\end{aligned} \tag{11}$$

and

$$\Lambda(t_1,\ldots,t_N) = \int_0^{t_1}\cdots\int_0^{t_N}\lambda(s_1,\ldots,s_N)\,\mathrm{d}s_1\cdots\mathrm{d}s_N \tag{12}$$

The relationship between $\mathbf{S}$ and Λ can not be simply formulated as in the univariate case[2]. For example, we obtain in the bivariate case

$$\mathbf{S}(t_1,t_2) = \mathbf{S}_1(t_1)\,\mathbf{S}_2(t_2)\,e^{\Lambda(t_1,t_2)} \tag{13}$$

Using hazard functions, construction of multivariate survival function is also not easy. Moreover, it is generally based on the conditional hazard rate (Shaked and Shanthikumar [1987]). In this article, we focus on another construction called the copula (or marginal) modelling which is more natural.

2.2 An example with the Clayton model

Clayton [1978] considers a bivariate 'association' model for an ordered pair of individuals. Let us denote T_1 and T_2 the age at failure of the first and second members of the pair. Clayton introduces also a cross-ratio function $\vartheta(t_1,t_2)$ defined in the following way

$$\vartheta(t_1,t_2) = \frac{\lambda(t_1 \mid T_2 = t_2)}{\lambda(t_1 \mid T_2 \ge t_2)} \tag{14}$$

"*This function may be interpreted as the ratio of the hazard rate of the conditional distribution of T_1, given $T_2 = t_2$, to that of T_1, given $T_2 \ge t_2$*" (Oakes [1989], page 488). We have

$$\begin{aligned}\lambda(t_1 \mid T_2 = t_2) &= -\frac{\partial_1\mathbf{S}_1(t_1 \mid T_2 = t_2)}{\mathbf{S}_1(t_1 \mid T_2 = t_2)}\\ &= -\frac{\partial_{1,2}\mathbf{S}(t_1,t_2)}{\partial_2\mathbf{S}(t_1,t_2)}\end{aligned} \tag{15}$$

and

$$\lambda(t_1 \mid T_2 \ge t_2) = -\frac{\partial_1\mathbf{S}(t_1,t_2)}{\mathbf{S}(t_1,t_2)} \tag{16}$$

[2] see Dabrowska [1996].

A new expression of $\vartheta(t_1,t_2)$ is then

$$\begin{aligned}\vartheta(t_1,t_2) &= \frac{\partial_{1,2}\mathbf{S}(t_1,t_2)\times\mathbf{S}(t_1,t_2)}{\partial_1\mathbf{S}(t_1,t_2)\times\partial_2\mathbf{S}(t_1,t_2)} \\ &= \frac{f(t_1,t_2)\times\mathbf{S}(t_1,t_2)}{\partial_1\mathbf{S}(t_1,t_2)\times\partial_2\mathbf{S}(t_1,t_2)}\end{aligned} \tag{17}$$

because $f(t_1,t_2)=\partial_{1,2}\mathbf{F}(t_1,t_2)=\partial_{1,2}\mathbf{S}(t_1,t_2)$. This representation of $\vartheta(t_1,t_2)$ corresponds to the one used by Oakes [1989] and we will see later that it plays an important role in survival modelling. Clayton assumes that $\vartheta(t_1,t_2)$ is constant and is equal to a parameter ϑ with $\vartheta\geq 0$. We have also

$$\frac{\partial_{1,2}\mathbf{S}(t_1,t_2)}{\mathbf{S}(t_1,t_2)}-\vartheta\frac{\partial_1\mathbf{S}(t_1,t_2)}{\mathbf{S}(t_1,t_2)}\times\frac{\partial_2\mathbf{S}(t_1,t_2)}{\mathbf{S}(t_1,t_2)}=0 \tag{18}$$

It comes that the survival function $\mathbf{S}(t_1,t_2)$ is the solution of the non-linear second-order partial differential equation[3]:

$$\partial_{1,2}\pi(t_1,t_2)+(\vartheta-1)\times\partial_1\pi(t_1,t_2)\times\partial_2\pi(t_1,t_2)=0 \tag{20}$$

with

$$\pi(t_1,t_2)=-\ln\mathbf{S}(t_1,t_2) \tag{21}$$

Clayton shows that the solution has the following form:

$$\mathbf{S}(t_1,t_2)=\left[1+(\vartheta-1)(a_1(t_1)+a_2(t_2))\right]^{-\frac{1}{\vartheta-1}} \tag{22}$$

where a_1 and a_2 are two nondecreasing functions with $a_1(0)=a_1(0)=0$.

We will now give the canonical representation of the survival function (22). The univariate margins of $\mathbf{S}(t_1,t_2)$ are respectively

$$\mathbf{S}_1(t_1):=\Pr\{T_1\geq t_1,T_2\geq 0\}=\mathbf{S}(t_1,0)=\left[1+(\vartheta-1)a_1(t_1)\right]^{-\frac{1}{\vartheta-1}} \tag{23}$$

and

$$\mathbf{S}_2(t_2)=\left[1+(\vartheta-1)a_2(t_2)\right]^{-\frac{1}{\vartheta-1}} \tag{24}$$

The survival copula associated to the Clayton model is then[4]

$$\begin{aligned}\breve{\mathbf{C}}(u_1,u_2) &= \mathbf{S}\left(\mathbf{S}_1^{-1}(u_1),\mathbf{S}_2^{-1}(u_2)\right) \\ &= \left[1+(\vartheta-1)\left(\frac{u_1^{1-\vartheta}-1}{\vartheta-1}+\frac{u_2^{1-\vartheta}-1}{\vartheta-1}\right)\right]^{-\frac{1}{\vartheta-1}} \\ &= \left(u_1^{1-\vartheta}+u_2^{1-\vartheta}-1\right)^{-\frac{1}{\vartheta-1}}\end{aligned} \tag{26}$$

[3]We remark that

$$\partial_{1,2}\left[-\ln\mathbf{S}(t_1,t_2)\right]=\frac{\partial_1\mathbf{S}(t_1,t_2)\times\partial_2\mathbf{S}(t_1,t_2)}{\mathbf{S}^2(t_1,t_2)}-\frac{\partial_{1,2}\mathbf{S}(t_1,t_2)}{\mathbf{S}(t_1,t_2)} \tag{19}$$

The result is then straightforward.

[4]We have

$$\mathbf{S}_n^{(-1)}(u_n)=a_n^{(-1)}\left(\frac{u_n^{1-\vartheta}-1}{\vartheta-1}\right) \tag{25}$$

If we assume that a_n is strictly increasing, the existence of the copula follows immediately. Otherwise, we have to involve the "Extension Theorem for Copulas" of Sklar [1996].

This copula function is a special case of the Cook-Johnson copula which is defined by the following function

$$\mathbf{C}\left(u_1,\ldots,u_N;\theta\right)=\left(\sum_{n=1}^{N}u_n^{-\theta}-N+1\right)^{-\frac{1}{\theta}} \tag{27}$$

We have the correspondence $\vartheta=\theta+1$.

2.3 Survival copulas

2.3.1 Definition

Let $\breve{\mathbf{C}}$ be a copula. A multivariate survival function $\mathbf{S}$ can be defined as follows

$$\mathbf{S}\left(t_1,\ldots,t_N\right)=\breve{\mathbf{C}}\left(\mathbf{S}_1\left(t_1\right),\ldots,\mathbf{S}_N\left(t_N\right)\right) \tag{28}$$

where $(\mathbf{S}_1,\ldots,\mathbf{S}_N)$ are the marginal survival functions. NELSEN [1999] notices that "$\breve{\mathbf{C}}$ *couples the joint survival function to its univariate margins in a manner completely analogous to the way in which a copula connects the joint distribution function to its margins*". We can then demonstrate the following theorem for survival distributions which is equivalent as this given by SKLAR [1959] for distributions.

Theorem 1 (Sklar's canonical representation) *Let $\mathbf{S}$ be an N-dimensional survival function with margins $\mathbf{S}_1,\ldots,\mathbf{S}_N$. Then $\mathbf{S}$ has a copula representation:*

$$\mathbf{S}\left(t_1,\ldots,t_N\right)=\breve{\mathbf{C}}\left(\mathbf{S}_1\left(t_1\right),\ldots,\mathbf{S}_N\left(t_N\right)\right) \tag{29}$$

The copula $\breve{\mathbf{C}}$ is unique if the margins are continuous. Otherwise, only the subcopula is uniquely determined on $\operatorname{Ran}\mathbf{S}_1\times\operatorname{Ran}\mathbf{S}_2\times\ldots\times\operatorname{Ran}\mathbf{S}_N$.

Proof. The proof is the same as for distribution functions (see SKLAR [1959], DEHEUVELS [1978] and SKLAR [1996]). ■

In the case $N=2$, NELSEN [1999] proves the theorem using distribution functions. Let $\mathbf{C}$ be the copula function of the bivariate distribution of (T_1,T_2). We have

$$\begin{aligned}\mathbf{S}\left(t_1,t_2\right)&=\Pr\left\{T_1>t_1,T_2>t_2\right\}\\&=1-\mathbf{F}_1\left(t_1\right)-\mathbf{F}_2\left(t_2\right)+\mathbf{F}\left(t_1,t_2\right)\\&=\mathbf{S}_1\left(t_1\right)+\mathbf{S}_2\left(t_2\right)-1+\mathbf{C}\left(1-\mathbf{S}_1\left(t_1\right),1-\mathbf{S}_2\left(t_2\right)\right)\\&=\breve{\mathbf{C}}\left(\mathbf{S}_1\left(t_1\right),\mathbf{S}_2\left(t_2\right)\right)\end{aligned} \tag{30}$$

with

$$\breve{\mathbf{C}}\left(u_1,u_2\right)=u_1+u_2-1+\mathbf{C}\left(1-u_1,1-u_2\right) \tag{31}$$

To prove the previous theorem, we have to verify that $\breve{\mathbf{C}}$ is a copula function.

1. The margins of $\breve{\mathbf{C}}$ are uniform:

$$\begin{aligned}\breve{\mathbf{C}}\left(u_1,1\right)&=u_1+\mathbf{C}\left(1-u_1,0\right)\\&=u_1\end{aligned} \tag{32}$$

and

$$\breve{\mathbf{C}}\left(1,u_2\right)=u_2 \tag{33}$$

2. $\breve{\mathbf{C}}$ is a grounded function:

$$\begin{aligned}\breve{\mathbf{C}}(u,0) &= \breve{\mathbf{C}}(0,u) \\ &= u-1+\mathbf{C}(1,1-u) \\ &= u-1+1-u \\ &= 0\end{aligned} \tag{34}$$

3. $\breve{\mathbf{C}}$ is 2-increasing:

$$V_{\breve{\mathbf{C}}}([u_1,v_1]\times[u_2,v_2]) = \breve{\mathbf{C}}(v_1,v_2) - \breve{\mathbf{C}}(v_1,u_2) - \breve{\mathbf{C}}(u_1,v_2) + \breve{\mathbf{C}}(u_1,u_2) \geq 0 \tag{35}$$

whenever $(u_1,u_2) \in [0,1]^2$, $(v_1,v_2) \in [0,1]^2$ such $0 \leq u_1 \leq v_1 \leq 1$ and $0 \leq u_2 \leq v_2 \leq 1$. To show this property, we compute

$$V_{\breve{\mathbf{C}}}([u_1,v_1]\times[u_2,v_2]) = \mathbf{C}(1-v_1,1-v_2) - \mathbf{C}(1-v_1,1-u_2) - \mathbf{C}(1-u_1,1-v_2) + \mathbf{C}(1-u_1,1-u_2) \tag{36}$$

Using the notations $\acute{u}_i = 1-u_i$ and $\acute{v}_i = 1-v_i$, we remark that

$$\begin{aligned}V_{\breve{\mathbf{C}}}([u_1,v_1]\times[u_2,v_2]) &= V_{\mathbf{C}}([\acute{v}_1,\acute{u}_1]\times[\acute{v}_2,\acute{u}_2]) \\ &\geq 0\end{aligned} \tag{37}$$

because $\mathbf{C}$ is 2-increasing and $0 \leq \acute{v}_1 \leq \acute{u}_1 \leq 1$ and $0 \leq \acute{v}_2 \leq \acute{u}_2 \leq 1$.

In the general case, we obtain similar results.

Theorem 2 *The relationship between the copula* $\mathbf{C}$ *and the survival copula* $\breve{\mathbf{C}}$ *is given by*

$$\breve{\mathbf{C}}(u_1,\ldots,u_n,\ldots,u_N) = \bar{\mathbf{C}}(1-u_1,\ldots,1-u_n,\ldots,1-u_N) \tag{38}$$

with

$$\bar{\mathbf{C}}(u_1,\ldots,u_n,\ldots,u_N) = \sum_{n=0}^{N}\left[(-1)^n \sum_{\mathbf{v}(u_1,\ldots,u_n,\ldots,u_N)\in\mathcal{Z}(N-n,N,1)} \mathbf{C}(v_1,\ldots,v_n,\ldots,v_N)\right] \tag{39}$$

where $\mathcal{Z}(M,N,\epsilon)$ *denotes the set* $\left\{\mathbf{v}\in[0,1]^N \mid v_n \in \{u_n,\epsilon\}, \sum_{n=1}^{N}\mathcal{X}_{\{\epsilon\}}(v_n) = M\right\}$.

Proof. see Appendix A.1. ■

Theorem 3 *The relationship between the copula* $\breve{\mathbf{C}}$ *and the survival copula* $\mathbf{C}$ *is given by*

$$\mathbf{C}(u_1,\ldots,u_n,\ldots,u_N) = \sum_{n=0}^{N}\left[(-1)^n \sum_{\mathbf{v}(u_1,\ldots,u_n,\ldots,u_N)\in\mathcal{Z}(N-n,N,0)} \breve{\mathbf{C}}(1-v_1,\ldots,1-v_n,\ldots,1-v_N)\right] \tag{40}$$

Proof. see Appendix A.2. ■

2.3.2 Properties

We collect here some properties that are useful for survival analysis. We first note that because $\breve{\mathbf{C}}$ is a copula function, results on copula apply to survival copulas. For example, we have

$$\mathbf{C}^{-} \prec \breve{\mathbf{C}} \prec \mathbf{C}^{+} \tag{41}$$

We refer to Nelsen [1999] for other interesting properties. In what follows, we give some new results which exploit the relationship between copulas and their survival copulas.

Property 1 *The copula is radially symmetric if and only if*

$$\breve{\mathbf{C}} = \mathbf{C} \tag{42}$$

Proof. By definition, a random vector $(T_1, \ldots, T_N)$ is said to be radially symmetric about $(t_1^\star, \ldots, t_N^\star)$ if and only if (Nelsen [1999], p. 31)

$$\mathbf{F}\left(t_1^\star + t_1, \ldots, t_N^\star + t_N\right) = \mathbf{S}\left(t_1^\star - t_1, \ldots, t_N^\star - t_N\right) \quad \text{for all } \mathbf{t} \in \mathbb{R}^N \tag{43}$$

The previous equation is equivalent to

$$\begin{aligned}
& \mathbf{C}\left(\mathbf{F}_1\left(t_1^\star + t_1,\right), \ldots, \mathbf{F}_N\left(t_N^\star + t_N\right)\right) = \breve{\mathbf{C}}\left(\mathbf{S}_1\left(t_1^\star - t_1,\right), \ldots, \mathbf{S}_N\left(t_N^\star - t_N\right)\right) \quad \text{for all } \mathbf{t} \in \mathbb{R}^N \\
\iff\ & \mathbf{C}\left(\mathbf{F}_1\left(t_1^\star + t_1,\right), \ldots, \mathbf{F}_N\left(t_N^\star + t_N\right)\right) = \breve{\mathbf{C}}\left(\mathbf{S}_1\left(t_1^\star - t_1,\right), \ldots, \mathbf{S}_N\left(t_N^\star - t_N\right)\right) \quad \text{for all } \mathbf{t} \in \mathbb{R}^N \\
\iff\ & \mathbf{C}\left(\mathbf{F}_1\left(t_1^\star + t_1,\right), \ldots, \mathbf{F}_N\left(t_N^\star + t_N\right)\right) = \breve{\mathbf{C}}\left(\mathbf{F}_1\left(t_1^\star + t_1,\right), \ldots, \mathbf{F}_N\left(t_N^\star + t_N\right)\right) \quad \text{for all } \mathbf{t} \in \mathbb{R}^N \\
\iff\ & \mathbf{C}\left(u_1, \ldots, u_N\right) = \breve{\mathbf{C}}\left(u_1, , \ldots, u_N\right) \quad \text{for all } \mathbf{u} \in [0,1]^N
\end{aligned} \tag{44}$$

because

$$\begin{aligned}
\mathbf{F}_n\left(t_n^\star + t_n\right) &= \mathbf{F}\left(+\infty, \ldots, +\infty, t_n^\star + t_n, +\infty, \ldots, +\infty\right) \\
&= \mathbf{S}\left(-\infty, \ldots, -\infty, t_n^\star - t_n, -\infty, \ldots, -\infty\right) \\
&= \mathbf{S}_n\left(t_n^\star - t_n\right)
\end{aligned} \tag{45}$$

■

We remark that the previous property is very interesting for computational purposes, because it is equivalent to work with the copula or to work with the survival copula. This is for example the case of the Normal copula[5]. Another examples are the following:

- the product copula $\mathbf{C}^{\perp}$
  For example, we verify that we have in the bivariate case

$$\begin{aligned}
\breve{\mathbf{C}}^{\perp}\left(u_1, u_2\right) &= u_1 + u_2 - 1 + \left(1 - u_1\right)\left(1 - u_2\right) \\
&= u_1 u_2 \\
&= \mathbf{C}^{\perp}\left(u_1, u_2\right)
\end{aligned} \tag{46}$$

- The upper Fréchet copula $\mathbf{C}^{+}$
  In the case $N = 2$, we have

$$\begin{aligned}
\breve{\mathbf{C}}^{+}\left(u_1, u_2\right) &= u_1 + u_2 - 1 + \min\left(1 - u_1, 1 - u_2\right) \\
&= u_1 + u_2 - \max\left(u_1, u_2\right) \\
&= \min\left(u_1, u_2\right) \\
&= \mathbf{C}^{+}\left(u_1, u_2\right)
\end{aligned} \tag{47}$$

[5] because the Normal distribution is radially symmetric about $(0, \ldots, 0)$.

- The bivariate lower Fréchet copula $\mathbf{C}^{-}$
  We have

$$\begin{aligned}
\breve{\mathbf{C}}^{-}(u_1,u_2) &= u_1+u_2-1+\max(1-u_1-u_2,0) \\
&= -\min(1-u_1-u_2,0) \\
&= \max(u_1+u_2-1,0) \\
&= \mathbf{C}^{-}(u_1,u_2)
\end{aligned} \tag{48}$$

Property 2 *In the bivariate case, if $\mathbf{C}_1 \succ \mathbf{C}_2$ then $\breve{\mathbf{C}}_1 \succ \breve{\mathbf{C}}_2$.*

Proof. If $\mathbf{C}_1 \succ \mathbf{C}_2$, we have

$$\begin{aligned}
& \mathbf{C}_1(1-u_1,1-u_2) \geq \mathbf{C}_2(1-u_1,1-u_2) \\
\Longleftrightarrow\quad & u_1+u_2-1+\mathbf{C}_1(1-u_1,1-u_2) \geq u_1+u_2-1+\mathbf{C}_2(1-u_1,1-u_2) \\
\Longleftrightarrow\quad & \breve{\mathbf{C}}_1(u_1,u_2) \geq \breve{\mathbf{C}}_2(u_1,u_2)
\end{aligned} \tag{49}$$

■

Remark 4 *This property is not verified for $N > 2$. For example, let consider special cases of the Farlie-Gumbel-Morgenstern copula (*Johnson *and* Kotz *[1975]). If we define the function $\mathbf{C}_{(\theta)}$ as follows*

$$\mathbf{C}_{(\theta)}(u_1,u_2,u_3) = u_1u_2u_3(1+\theta(1-u_1)(1-u_2)) \tag{50}$$

We verify that it is a copula[6] *for $|\theta| \leq 1$. We remark that*

$$\breve{\mathbf{C}}_{(\theta_1)}(u_1,u_2,u_3) - \breve{\mathbf{C}}_{(\theta_2)}(u_1,u_2,u_3) = (\theta_1-\theta_2)\cdot u_1u_2u_3(1-u_1)(1-u_2) \tag{52}$$

So, if $\theta_1 \leq \theta_2$, we have $\mathbf{C}_{(\theta_1)} \prec \mathbf{C}_{(\theta_2)}$ and $\breve{\mathbf{C}}_{(\theta_1)} \prec \breve{\mathbf{C}}_{(\theta_2)}$. We define now $\mathbf{C}_{(\theta)}$ in a more complicated way

$$\mathbf{C}_{(\theta)}(u_1,u_2,u_3) = u_1u_2u_3(1+\theta(1-u_1)(1-u_2)(1-u_3)) \tag{53}$$

$\mathbf{C}_{(\theta)}$ is a copula function[7] *for $|\theta| \leq 1$. If $\theta_1 \leq \theta_2$, it implies that $\mathbf{C}_{(\theta_1)} \prec \mathbf{C}_{(\theta_2)}$. Nevertheless, we have $\breve{\mathbf{C}}_{(\theta_1)} \succ \breve{\mathbf{C}}_{(\theta_2)}$ because*

$$\begin{aligned}
\breve{\mathbf{C}}_{(\theta_1)}(u_1,u_2,u_3) - \breve{\mathbf{C}}_{(\theta_2)}(u_1,u_2,u_3) &= -(\theta_1-\theta_2)\cdot u_1u_2u_3(1-u_1)(1-u_2)(1-u_3) \\
&= \mathbf{C}_{(\theta_2)}(u_1,u_2,u_3) - \mathbf{C}_{(\theta_1)}(u_1,u_2,u_3) \\
&\geq 0
\end{aligned} \tag{55}$$

The following result can be useful for simulation issues.

Property 3 *Let $X_1,\ldots,X_N$ be N random variables with continuous distributions $\mathbf{F}_1,\ldots,\mathbf{F}_N$ and copula $\mathbf{C}$. We consider N continous distributions $\mathbf{G}_1,\ldots,\mathbf{G}_N$ and we denote T_n the random variable $T_n = \mathbf{G}_n^{(-1)}(1-\mathbf{F}_n(X_n))$. It comes that the margins and the copula of the random vector $(T_1,\ldots,T_N)$ are respectively $\mathbf{G}_1,\ldots,\mathbf{G}_N$ and the survival copula $\breve{\mathbf{C}}$ of $\mathbf{C}$.*

---

[6] $\mathbf{C}_{(\theta)}$ is grounded and $\mathbf{C}_{(\theta)}(\mathbf{u}) = u_n$ if all coordinates of $\mathbf{u}$ are 1 except u_n. Moreover, the density is

$$c_{(\theta)}(u_1,u_2,u_3) = 1+\theta(1-2(u_1+u_2)+4u_1u_2) \tag{51}$$

If $|\theta| \leq 1$, $c_{(\theta)}(u_1,u_2,u_3) \geq 0$.

[7] $\mathbf{C}_{(\theta)}$ is grounded and $\mathbf{C}_{(\theta)}(\mathbf{u}) = u_n$ if all coordinates of $\mathbf{u}$ are 1 except u_n. Moreover, the density is

$$c_{(\theta)}(u_1,u_2,u_3) = 1+\theta(1-2(u_1+u_2+u_3)+4(u_1u_2+u_1u_3+u_2u_3)-8u_1u_2u_3) \tag{54}$$

If $|\theta| \leq 1$, $c_{(\theta)}(u_1,u_2,u_3) \geq 0$.

Proof. The fact that $\mathbf{C}\langle T_1,\ldots,T_N\rangle$ is $\breve{\mathbf{C}}\langle X_1,\ldots,X_N\rangle$ is easy to show because $\mathbf{G}_n^{(-1)}$ and $1-\mathbf{F}_n$ are respectively strictly increasing and strictly decreasing functions. Generalisation of the theorem 2.4.4 of Nelsen [1999] can then be applied. To prove that the margins are $\mathbf{G}_1,\ldots,\mathbf{G}_N$, we use the following statements:

$$\begin{aligned}\Pr\{T_n\le t_n\} &= \Pr\left\{\mathbf{G}_n^{(-1)}\left(1-\mathbf{F}_n\left(X_n\right)\right)\le t_n\right\} \\ &= 1-\Pr\left\{\mathbf{F}_n\left(X_n\right)\le 1-\mathbf{G}_n\left(t_n\right)\right\} \\ &= \mathbf{G}_n\left(t_n\right)\end{aligned} \tag{56}$$

■

The density of $\breve{\mathbf{C}}$ is related to the density of $\mathbf{C}$ by the next property.

Property 4 *The absolutely continuous component of the survival copula* $\breve{\mathbf{C}}$ *at the point* $(u_1,\ldots,u_N)$ *is equal to the absolutely continuous component of the copula* $\mathbf{C}$ *at the reflexive point* $(1-u_1,\ldots,1-u_N)$.

Proof. see equation (239) page 66. ■

We obtain a related result for radially symmetric copulas. The absolutely continuous component of the survival copula $\breve{\mathbf{C}}$ is equal to these of the copula $\mathbf{C}$. Moreover, it is symmetric about the point $\left(\frac{1}{2},\ldots,\frac{1}{2}\right)$.

2.4 Special cases

In this paragraph, we present some well-known copula functions and give some references for their applications in survival modelling and reliability theory. Because a survival copula is a copula function too, we notice that survival functions may be defined in two ways. The first approach uses a survival copula, that is a survival copula based on a copula derived from a specific distribution

$$\mathbf{S}\left(t_1,\ldots,t_N\right)=\breve{\mathbf{C}}\left(\mathbf{S}_1\left(t_1\right),\ldots,\mathbf{S}_N\left(t_N\right)\right) \tag{57}$$

The second approach specifies directly a copula function

$$\mathbf{S}\left(t_1,\ldots,t_N\right)=\mathbf{C}\left(\mathbf{S}_1\left(t_1\right),\ldots,\mathbf{S}_N\left(t_N\right)\right) \tag{58}$$

Let us consider the previous Clayton model. We have

$$\mathbf{S}\left(t_1,t_2\right)=\left(\mathbf{S}_1^{-\theta}\left(t_1\right)+\mathbf{S}_2^{-\theta}\left(t_2\right)-1\right)^{-\frac{1}{\theta}} \tag{59}$$

The copula function is then $\mathbf{C}\left(u_1,u_2\right)=\left(u_1^{-\theta}+u_2^{-\theta}-1\right)^{-\frac{1}{\theta}}$. We may also build another model using the survival copula of the Cook-Johnson copula. We have

$$\breve{\mathbf{C}}\left(u_1,u_2\right)=u_1+u_2-1+\left(\left(1-u_1\right)^{-\theta}+\left(1-u_2\right)^{-\theta}-1\right)^{-\frac{1}{\theta}} \tag{60}$$

and

$$\mathbf{S}\left(t_1,t_2\right)=\mathbf{S}_1\left(t_1\right)+\mathbf{S}_2\left(t_2\right)-1+\left(\left(1-\mathbf{S}_1\left(t_1\right)\right)^{-\theta}+\left(1-\mathbf{S}_2\left(t_2\right)\right)^{-\theta}-1\right)^{-\frac{1}{\theta}} \tag{61}$$

The two survival functions (59) and (61) give two different models.

For the margins $\mathbf{S}_n$, we could use the univariate survival functions in Table 1 with explanatory variables. In finance, we generally consider piecewise constant hazard rates

$$\lambda\left(t\right)=\prod_{m=1}^{M}\lambda_m\mathbf{1}_{\left[t_{m-1}^{\star}<t\le t_m^{\star}\right]} \tag{62}$$

Model	$\lambda(t)$	$\mathbf{S}(t)$
Exponential	λ	$\exp(-\lambda t)$
Weibull	$\lambda\gamma t^{\gamma-1}$	$\exp(-\lambda t^{\gamma})$
Lognormal	$\gamma t^{-1}\phi(\gamma\ln(\lambda t))/(1-\Phi(\gamma\ln(\lambda t)))$	$1-\Phi(\gamma\ln(\lambda t))$
Loglogit	$\lambda\gamma^{-1}t^{\frac{1}{\gamma}}/\left(t+\lambda t^{1+\frac{1}{\gamma}}\right)$	$1/\left(1+\lambda t^{\frac{1}{\gamma}}\right)$
Gompertz	$\lambda\gamma\exp(\gamma t)$	$\exp(\lambda(1-e^{\gamma t}))$

Table 1: Some univariate survival functions

with $t_m^{\star}$ the knots of the function and $t_M^{\star}=\infty$. The survival function is then

$$\mathbf{S}(t)=\exp\left(\sum_{i=1}^{m}\lambda_i\left(t_{i-1}^{\star}-t_i^{\star}\right)-\lambda_{m+1}\left(t-t_m^{\star}\right)\right)\qquad\text{if } t\in\left]t_m^{\star},t_{m+1}^{\star}\right] \tag{63}$$

2.4.1 Copulas related to exponential distributions

The exponential distribution plays a prominent role in statistics. The reason is that it has a lot of interesting properties, and moreover it could be justified by many different mathematical constructions (or equivently it is the law of physical mecanisms). For some statistical problems, the choice of one distribution is not only a question of statistical fitting, because the properties of the distribution could not be adapted to the problem:

> *The application of probability theory and mathematical statistics to real life situations involves two steps. First, a model is set up, and then probabilistic or statistical methods are applied within the adopted model. In many instances the first step is arbitrary. [...] The way of avoiding an arbitrary distributional assumption is to develop a characterization theorem from some basic assumptions of the real life situation we face* (GALAMBOS [1982]).

In the same article, Galambos considers the following problem:

> Assume that a system consists of N identical components which are connected in series. This means that the systems fails as soon as one of the components fails. One can assume that the components function independently. Assume further that the random time interval until the failure of the system is one Nth of the time interval of component failure.

We have

$$\Pr\{\min(T_1,\ldots,T_N)\leq t\}=\Pr\{T_1\leq N\cdot t\} \tag{64}$$

Hence

$$\mathbf{S}(t)=\mathbf{S}^N\left(\frac{t}{N}\right) \tag{65}$$

with $\mathbf{S}(t)=\Pr\{T_1>t\}$. The only solution of this functional equation for all intergers $N\geq 1$ is the exponential distribution:

$$\mathbf{S}(t)=\exp(-\lambda t) \tag{66}$$

We obtain here a 'physical' justification of the raison d'être of the exponential distribution. Other examples could be found in different fields of probability: Poisson processes, renewal theory or order statistics. GALAMBOS and KOTZ [1978] have tried to unify the different characterization of the exponential distribution and have shown the equivalence of the four following properties:

1. the hazard rate of an exponential random variable T is constant;

2. its expected residual life $\mathbb{E}[T \mid T \geq t]$ is constant;

3. it has the "lack of memory" (LMP) property, that is

$$\Pr\{T \geq t_1 + t_2 \mid T \geq t_1\} = \Pr\{T \geq t_2\} \tag{67}$$

or equivently if T is absolutely continuous

$$\mathbf{S}(t_1 + t_2) = \mathbf{S}(t_1)\,\mathbf{S}(t_2) \tag{68}$$

4. the distribution of $N \cdot T_{1:N}$ is the same as T.

These different properties explain the importance of the exponential distribution in reliability and survival modelling. Different authors have also searched to propose extensions to the multidimensional case. The first one is Gumbel [1960] who proposes

$$\mathbf{F}(t_1, t_2) = 1 - e^{-t_1} - e^{-t_2} + e^{-(t_1 + t_2 + \theta t_1 t_2)} \tag{69}$$

and

$$\mathbf{F}(t_1, t_2) = \left(1 - e^{-t_1}\right)\left(1 - e^{-t_2}\right)\left(1 + \theta e^{-(t_1 + t_2)}\right) \tag{70}$$

Its two bivariate exponential distributions are in fact two bivariate distributions with exponential margins. As Freund [1961] remarks, Gumbel does "*not discuss the appropriateness of these models to particular physical situations*". Using a model for the lifetimes of two components, Freund find "*a bivariate extension of the exponential distribution*":

$$\mathbf{S}(t_1, t_2) = \begin{cases} \frac{\alpha_1}{\alpha_1 + \beta_1 - \beta_2} e^{-\beta_2 t_2 - (\alpha_1 + \beta_1 - \beta_2) t_1} + \frac{\beta_1 - \beta_2}{\alpha_1 + \beta_1 - \beta_2} e^{-(\alpha_1 + \beta_1) t_2} & \text{for} \quad t_1 < t_2 \\ \frac{\beta_1}{\alpha_1 + \beta_1 - \alpha_2} e^{-\alpha_2 t_1 - (\alpha_1 + \beta_1 - \alpha_2) t_2} + \frac{\alpha_1 - \alpha_2}{\alpha_1 + \beta_1 - \alpha_2} e^{-(\alpha_1 + \beta_1) t_1} & \text{for} \quad t_1 \geq t_2 \end{cases} \tag{71}$$

However, we remark that the margins are not exponential, but mixtures of exponentials:

$$\mathbf{S}(t_1, 0) = \frac{\beta_1}{\alpha_1 + \beta_1 - \alpha_2} e^{-\alpha_2 t_1} + \frac{\alpha_1 - \alpha_2}{\alpha_1 + \beta_1 - \alpha_2} e^{-(\alpha_1 + \beta_1) t_1} \tag{72}$$

The difference between the Gumbel and Freund exponential distributions poses the following problem:

How to characterize a bivariate exponential distribution?

The answer given by Hutchinson and Lai [1990] is to consider bivariate distributions which verify bivariate extensions of properties of the univariate exponential distribution. For example, we could consider the following properties:

1. $\mathbf{F}(t_1, t_2)$ is absolutely continuous;

2. the hazard rate $\lambda(t_1, t_2)$ is constant;

3. $\mathbf{F}$ has the "lack of memory" (LMP) property

$$\mathbf{S}(t_1 + t, t_2 + t) = \mathbf{S}(t_1, t_2)\,\mathbf{S}(t, t) \tag{73}$$

4. the order statistic $\min(T_1, T_2)$ is exponential distributed;

5. $\min(T_1, T_2)$ and $T_1 - T_2$ are independently distributed;

6. etc.

Nevertheless, it is not possible to obtain a bivariate distribution which has all these properties[8] (Block and Basu [1974]). So, we could only obtain a distribution "close" to the idea of an exponential bivarate distribution.

Marshall and Olkin [1967] consider a two-component system subject to "fatal" shocks governed by three independent Poisson processes $N_1(t)$, $N_2(t)$ and $N_{12}(t)$ with intensities λ_1, λ_2 and λ_{12}. The two first processes N_1 and N_2 control shocks to individual components, whereas N_{12} control shocks to both components. We have

$$\begin{aligned}\mathbf{S}(t_1, t_2) &= \Pr\{N_1(t_1) = 0\} \cdot \Pr\{N_2(t_2) = 0\} \cdot \Pr\{N_{12}(\max(t_1, t_2)) = 0\} \\ &= \exp(-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)) \end{aligned} \tag{74}$$

This distribution is certainly the most known bivariate exponential distribution and is sometimes denoted BVE. Nevertheless, it is not absolutely continuous. By omitting the singular part of BVE, Block and Basu [1974] have defined the ACBVE distribution in the following way:

$$\mathbf{S}(t_1, t_2) = \frac{\lambda_1 + \lambda_2 + \lambda_{12}}{\lambda_1 + \lambda_2} \exp(-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)) - \frac{\lambda_{12}}{\lambda_1 + \lambda_2} \exp(-(\lambda_1 + \lambda_2 + \lambda_{12}) \max(t_1, t_2)) \tag{75}$$

The three distributions Freund, BVE and ACBVE have the properties (3), (4) and (5). BVE is not absolutely continuous whereas the margins of Freund and ACBVE are not exponentials, but mixture of two exponentials.

Thanks to Sklar's theorem, we could deduce the copula of the previous distributions. Let us consider the Gumbel's bivariate exponential distribution (69). We have

$$\mathbf{C}(u_1, u_2) = u_1 + u_2 - 1 + (1 - u_1)(1 - u_2) e^{-\theta \ln(1-u_1) \ln(1-u_2)} \tag{76}$$

Barnett [1980] has shown that the corresponding survival copula is

$$\breve{\mathbf{C}}(u_1, u_2) = u_1 u_2 \exp(-\theta \ln u_1 \ln u_2) \tag{77}$$

It is called the Gumbel-Barnett copula (Hutchinson and Lai [1990], page 94). Note also that the second Gumbel's bivariate exponential distribution (70) corresponds in fact to a FGM copula. Nelsen [1999] shows that the survival copula of the Marshall-Olkin model is

$$\breve{\mathbf{C}}(u_1, u_2) = \min\left(u_1^{1-\theta_1} u_2, u_1 u_2^{1-\theta_2}\right) \tag{78}$$

with $\theta_1 = \lambda_{12} / (\lambda_1 + \lambda_{12})$ and $\theta_2 = \lambda_{12} / (\lambda_2 + \lambda_{12})$. In the case of the Freund and ACBVE distributions, there does not exist an analytical expression of the copula (but it could be computed numerically). For other copulas related to exponential distributions, we refer to Joe and Hu [1996], Joe [1997] and Joe and Ma [2000].

Remark 5 *The previous bivariate exponential distributions has a multivariate extension. This is for example the case of Marshall-Olkin family (*Basu *[1988]).*

2.4.2 Archimedean copulas

Definition 6 (Schweizer and Sklar [1983, Theorem 6.3.6, p. 88], Nelsen [1999, Theorem 4.6.2, p. 122]) *Let φ be a continous strictly decreasing function from $[0, 1]$ to $[0, \infty)$ such that $\varphi(0) = \infty$, $\varphi(1) = 0$, and*

$$(-1)^n \frac{\mathrm{d}^n}{\mathrm{d}u^n} \varphi^{-1}(u) \geq 0 \quad \text{for } n = 1, \ldots, N \tag{79}$$

The function defined by

$$\mathbf{C}(u_1, \ldots, u_N) = \max\left(\varphi^{-1}(\varphi(u_1) + \ldots + \varphi(u_N)), 0\right) \tag{80}$$

is an Archimedean copula. φ is called the generator of the copula.

[8] except the product of two exponential distributions, that is the distribution of two independent exponential random variables.

The name 'archimedean' is explained in details in LING [1965] (see also GENEST and MACKAY [1986a,1986b] and NELSEN [1999] page 98). Archimedean copulas play an important role because of their mathematical properties and computational facilities. Moreover, many standard 2-copulas are Archimedean[9] (see Table 4.1 of NELSEN [1999]).

This is for example the case of the Frank copula, which is defined by the function

$$\mathbf{C}(u_1, u_2; \theta) = -\frac{1}{\theta} \ln \left(1 + \frac{[\exp(-\theta u_1) - 1][\exp(-\theta u_2) - 1]}{\exp(-\theta) - 1} \right) \tag{81}$$

The expression of the generator φ is

$$\varphi(u) = -\ln \left(\frac{\exp(-\theta u) - 1}{\exp(-\theta) - 1} \right) \tag{82}$$

We can show that this copula is *comprehensive*: $\mathbf{C}$ corresponds respectively to $\mathbf{C}^-$, $\mathbf{C}^\perp$ and $\mathbf{C}^+$ when θ tends to $-\infty$, 0 and $+\infty$ — but only $\mathbf{C}^-$ and $\mathbf{C}^\perp$ are Archimedean. We remark that

$$\begin{aligned}
\breve{\mathbf{C}}(u_1, u_2) &= u_1 + u_2 - 1 - \frac{1}{\theta} \ln \left(1 + \frac{[\exp(-\theta(1-u_1)) - 1][\exp(-\theta(1-u_2)) - 1]}{\exp(-\theta) - 1} \right) \\
&= -\frac{1}{\theta} \ln \left(\left(1 + \frac{[\exp(-\theta(1-u_1)) - 1][\exp(-\theta(1-u_2)) - 1]}{\exp(-\theta) - 1} \right) \exp(-\theta(u_1 + u_2 - 1)) \right) \\
&= -\frac{1}{\theta} \ln \left(1 + \frac{[\exp(-\theta u_1) - 1][\exp(-\theta u_2) - 1]}{\exp(-\theta) - 1} \right)
\end{aligned} \tag{83}$$

It comes that the survival copula is equal to the copula (GENEST [1987]). Moreover, we can show that the only Archimedean copula that verify this property is the Frank copula (see FRANK [1979,1991]). So, for other Archimedean copulas, the corresponding survival copula is not Archimedean. However, we may find copulas which are not Archimedean such that the survival copula is Archimedean. For example, GENEST and MACKAY [1986a] remark that the Gumbel-Barnett copula (77) is Archimedean — $\varphi(u) = \ln(1 - \theta \ln t)$ — whereas the Gumbel copula (76) is not Archimedean.

Note that the Clayton model corresponds to an Archimedean copula: $\varphi(u) = \ln(1 - \theta \ln t)$. Another interesting model has been suggested by HOUGAARD [1986a,1986b]:

$$\mathbf{C}(u_1, u_2; \theta) = \exp\left(-\left((-\ln u_1)^\theta + (-\ln u_2)^\theta \right)^{\frac{1}{\theta}} \right) \tag{84}$$

It is called the Gumbel-Hougaard copula[10] (HUTCHINSON and LAI [1990], page 84). The main difference between these two copulas is that the Clayton (or Cook-Johnson) copula could be extended to present a negative dependence (GENEST and MACKAY [1986a]):

$$\mathbf{C}(u_1, u_2; \theta) = \max\left(\left(u_1^{-\theta} + u_2^{-\theta} - 1 \right)^{-\frac{1}{\theta}}, 0 \right) \tag{85}$$

If $\theta \in [-1, 0[$, $\mathbf{C} \prec \mathbf{C}^\perp$. For the Gumbel-Hougaard copula, extension to negative dependence is not possible.

2.4.3 Frailty models

Frailty models have been introduced by LANCASTER [1979] and VAUPEL, MANTON and STALLARD [1979] and have been popularized by OAKES [1989]. The main idea is to introduce dependence between survival times

[9]Note that the conditions (79) and $\varphi(0) = \infty$ are replaced by φ is convex in the bivariate case.

[10]GUMBEL [1961] is one of the first to use this copula.

$T_1, \ldots, T_N$ by using an unobserved random variable W, called the frailty. Given the frailty W with distribution $\mathbf{G}$, the survival times are assumed to be independent:

$$\Pr\{T_1 > t_1, \ldots, T_N > t_N \mid W = w\} = \prod_{n=1}^{N} \Pr\{T_n > t_n \mid W = w\} \tag{86}$$

We have

$$\begin{aligned} \mathbf{S}(t_1, \ldots, t_N \mid w) &= \prod_{n=1}^{N} \mathbf{S}_n(t_1 \mid w) \\ &= \boldsymbol{\chi}_1^w(t_1) \times \boldsymbol{\chi}_2^w(t_2) \times \cdots \times \boldsymbol{\chi}_N^w(t_N) \end{aligned} \tag{87}$$

where $\boldsymbol{\chi}_n(t_n)$ is the baseline survival function. The unconditional joint survival function is then defined by

$$\mathbf{S}(t_1, \ldots, t_N) = \mathbb{E}[\mathbf{S}(t_1, \ldots, t_N \mid w)] \tag{88}$$

where the expectation is taken with respect to the random variable W. We have

$$\mathbf{S}(t_1, \ldots, t_N) = \int \prod_{n=1}^{N} [\boldsymbol{\chi}_n(t_n)]^w \, \mathrm{d}\mathbf{G}(w) \tag{89}$$

In order to have a more interesting representation of frailty models, we need the following theorem due to Marshall and Olkin [1988].

Theorem 7 (Marshall and Olkin [1988, Theorem 2.1, p. 835]) *Let $\mathbf{F}_1, \ldots, \mathbf{F}_N$ be univariate distribution functions, and let $\mathbf{G}$ be an N-variate distribution function such that $\bar{\mathbf{G}}(0, \ldots, 0) = 1$, with univariate marginals $\mathbf{G}_n$. Denote the Laplace transform of $\mathbf{G}$ and $\mathbf{G}_n$, respectively, by ψ and ψ_n. Let $\mathbf{C}$ be an N-variate distribution function with all univariate marginals uniform on $[0,1]$. If $\mathbf{H}_n(x) = \exp\left(-\psi_n^{-1}(\mathbf{F}_n(x))\right)$, then*

$$\mathbf{F}(x_1, \ldots, x_N) = \int \cdots \int \mathbf{C}\left([\mathbf{H}_1(x_1)]^{w_1}, \ldots, [\mathbf{H}_N(x_N)]^{w_N}\right) \mathrm{d}\mathbf{G}(w_1, \ldots, w_N) \tag{90}$$

is an N-variate distribution function with marginals $\mathbf{F}_1, \ldots, \mathbf{F}_N$.

Marshall and Olkin [1988] study then "*a particularly interesting and simple case of (90)*". Let us assume that the univariate marginal distributions $\mathbf{G}_n$ are the same — we note them $\mathbf{G}_1$ — $\mathbf{G}$ is the upper Fréchet bound and $\mathbf{C}$ is the product copula $\mathbf{C}^{\perp}$. Expression (90) becomes

$$\begin{aligned} \mathbf{F}(x_1, \ldots, x_N) &= \int \prod_{n=1}^{N} [\mathbf{H}_n(x_n)]^{w_1} \, \mathrm{d}\mathbf{G}_1(w_1) \\ &= \int \exp\left(-w_1 \sum_{n=1}^{N} \psi_1^{-1} \mathbf{F}_n(x_n)\right) \mathrm{d}\mathbf{G}_1(w_1) \\ &= \psi_1\left(\psi_1^{-1}(\mathbf{F}_1(x_1)) + \ldots + \psi_1^{-1}(\mathbf{F}_N(x_N))\right) \end{aligned} \tag{91}$$

The corresponding copula is then given by

$$\mathbf{C}(u_1, \ldots, u_N) = \psi_1\left(\psi_1^{-1}(u_1) + \ldots + \psi_1^{-1}(u_N)\right) \tag{92}$$

It is a special case of an Archimedean copula where the generator φ is the inverse of a Laplace transform. We can now state the definition of frailty survival functions.

Definition 8 *A frailty survival function is a special case of the construction based on copulas*

$$\mathbf{S}(t_1,\ldots,t_N)=\breve{\mathbf{C}}(\mathbf{S}_1(t_1),\ldots,\mathbf{S}_N(t_N)) \tag{93}$$

where $\breve{\mathbf{C}}$ is an Archimedean copula with a generator corresponding to the inverse of the Laplace transform of the distribution of the frailty variable W. More generally, the generator is the inverse of a Laplace transform[11].

We can now give some examples of frailty distributions. Let us consider the Clayton model. The copula function is frailty with the Laplace transform of a Gamma variate: $\psi(x)=(1+x)^{-\frac{1}{\theta}}$. The Gumbel-Hougaard copula is frailty too and we have $\psi(x)=\exp\left(-x^{\frac{1}{\theta}}\right)$. This is the Laplace transfom of a positive stable distribution.

We end this paragraph by two remarks.

Remark 10 *Frailty models exhibit only positive dependence (in the PQD sense). To show this property, we use theorem 3.1 of* Marshall *and* Olkin *[1988]. They show that if the distributions $\mathbf{C}$ and $\mathbf{G}$ in (90) are associated, then $\mathbf{F}$ given by (90) is associated. Because $\mathbf{C}^{\perp}$ and the upper Fréchet bound are associated, frailty distributions are necessarily associated. It comes that the survival times are PQD. Note that the PQD property is not satisfied for 'extensions' of copulas derived from frailty models. For example, if we consider the Cook-Johnson copula (85), $\mathbf{C}\prec\mathbf{C}^{\perp}$ if $\theta\in[-1,0[$. In this case, $\psi(x)=(1+x)^{-\frac{1}{\theta}}$ is not a Laplace transform, because the completely monotone property does not hold. Another example is the Frank copula which corresponds to a frailty model* ***if and only if*** *$\theta\geq 0$.*

Remark 11 *In this article, frailty models are in fact* ***proportional*** *hazards frailty models:*

$$\Lambda(t\mid W=w)=w\Lambda(t) \tag{94}$$

We do not consider others frailty models, for example the multiplicative hazards frailty models (Anderson *and* Louis *[1995]):*

$$\Lambda(t\mid W=w)=\Lambda(wt) \tag{95}$$

In this last case, we have

$$\mathbf{S}(t_1,\ldots,t_N)=\int\prod_{n=1}^{N}\boldsymbol{\chi}_n(wt_n)\,\mathrm{d}\mathbf{G}(w) \tag{96}$$

2.4.4 Miscellaneous copulas

Other families of copula functions may be used for survival modelling. In particular, we refer to Joe [1997] and Nelsen [1999] for additional materials.

For example, we may consider the Normal copula defined as follows

$$\mathbf{C}(u_1,\ldots,u_N;\boldsymbol{\rho})=\Phi_{\boldsymbol{\rho}}\left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_N)\right) \tag{97}$$

[11] We may use the following characterization.

Definition 9 (Feller [1971, theorem 1, p. 439]) *A function ψ on $[0,\infty)$ is the Laplace transform of a probability distribution iff it is completely monotone, and $\psi(0)=1$.*

where $\boldsymbol{\rho}$ is a symmetric, positive definite matrix with diag $\boldsymbol{\rho} = \mathbf{1}$ and $\Phi_{\boldsymbol{\rho}}$ is the standardized multivariate Normal cdf with correlation matrix $\boldsymbol{\rho}$. The corresponding density is

$$\mathbf{c}\left(u_1, \ldots, u_N; \boldsymbol{\rho}\right) = \frac{1}{\left|\boldsymbol{\rho}\right|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\boldsymbol{\varsigma}^\top \left(\boldsymbol{\rho}^{-1} - \mathbb{I}\right)\boldsymbol{\varsigma}\right) \tag{98}$$

with $\varsigma_n = \Phi^{-1}\left(u_n\right)$. Except the work of SONG [2000], this copula function has been hardly studied. However, it presents very interesting properties for financial applications (BOUYÉ, DURRLEMAN, NIKEGHBALI, RIBOULET and RONCALLI [2000a,2000b]). Moreover, there exists an analytical expression of the ML estimator of the parameters (DURRLEMAN, NIKEGHBALI and RONCALLI [2000a]) and simulation is straightforward.

2.5 Dependence measures and related concepts

2.5.1 General considerations about dependence

Since the works of DEHEUVELS [1978,1979a] and SCHWEIZER and WOLFF [1981], we know that the dependence between random variables is characterized entirely by the copula of the corresponding multivariate distribution. However, it could be interesting to use a dependence measure (a single number) to compare different survival functions, because the direct comparison between survival copulas may not be obvious[12]. It is common to use the correlation measure and to speak about "*correlated survival times*". In this paragraph, we will show that it could make no sense to use this measure.

"*The traditional way of evaluating dependence in a bivariate distribution is by means of the correlation coefficient (Pearson correlation)*" (HOUGAARD [2000], page 129). It is defined by

$$\rho\left(T_1, T_2\right) = \frac{\operatorname{cov}\left(T_1, T_2\right)}{\sqrt{\operatorname{var}\left(T_1\right)\operatorname{var}\left(T_2\right)}} \tag{99}$$

with

$$\operatorname{cov}\left(T_1, T_2\right) = \int_0^\infty \int_0^\infty \left(\mathbf{S}\left(t_1, t_2\right) - \mathbf{S}_1\left(t_1\right)\mathbf{S}_2\left(t_2\right)\right) \mathrm{d}t_1 \, \mathrm{d}t_2 \tag{100}$$

and

$$\operatorname{var}\left(T_n\right) = 2\int_0^\infty t\mathbf{S}_n\left(t\right) \mathrm{d}t - \left[\int_0^\infty \mathbf{S}_n\left(t\right) \mathrm{d}t\right]^2 \tag{101}$$

The Pearson correlation is an appropriate measure of the dependence when the random variables have **jointly** a multivariate normal distribution. EMBRECHTS, MCNEIL and STRAUMANN [1999,2000] show that the standard correlation approach to dependency remains natural and unproblematic in the class of elliptical distributions. When the distribution is not elliptical, the use of the Pearson correlation may be problematic. This is generally the case in survival modelling.

Here are some misinterpretations of the Pearson correlation:

1. T_1 and T_2 are independent **if and only if** $\rho\left(T_1, T_2\right) = 0$;
2. $\rho\left(T_1, T_2\right) = 0$ means that there are no perfect dependence between T_1 and T_2;
3. for given margins, the permissible range of $\rho\left(T_1, T_2\right)$ is $[-1, 1]$;

[12] Moreover, the concordance ordering is only a *partial* ordering of the set of copulas (NELSEN [1999], page 34).

To show that the first statement is false[13], we consider the cubic copula of DURRLEMAN, NIKEGHBALI and RONCALLI [2000b]

$$\mathbf{C}(u_1,u_2) = u_1u_2 + \alpha\left[u_1(u_1-1)(2u_1-1)\right]\left[u_2(u_2-1)(2u_2-1)\right] \tag{102}$$

with $\alpha \in [-1,2]$. If the margins of T_1 and T_2 are continous and symmetric, the authors show that the Pearson correlation is zero. Moreover, if $\alpha \neq 0$, the random variables T_1 and T_2 are not independent. For the second statement[14], we consider the following copula:

$$\mathbf{C}(u_1,u_2) = \begin{cases} u_1 & 0 \leq u_1 \leq \frac{1}{2}u_2 \leq \frac{1}{2} \\ \frac{1}{2}u_2 & 0 \leq \frac{1}{2}u_2 \leq u_1 \leq 1 - \frac{1}{2}u_2 \\ u_1 + u_2 - 1 & \frac{1}{2} \leq 1 - \frac{1}{2}u_2 \leq u_1 \leq 1 \end{cases} \tag{103}$$

NELSEN [1999] shows that $\rho(U_1,U_2) = 0$, but $\Pr\{U_2 = 1 - |2U_1 - 1|\} = 1$, i.e. "*the two random variables can be uncorrelated although one can be predicted perfectly from the other*" (NELSEN [1999], page 57). For the last statement, we need works of TCHEN [1980], who shows that ρ is increasing with respect to the concordance order[15]

$$\mathbf{C}_1 \succ \mathbf{C}_2 \Rightarrow \rho(T_1,T_2;\mathbf{C}_1) \geq \rho(T_1,T_2;\mathbf{C}_2) \tag{104}$$

It comes that $\rho(T_1,T_2)$ is bounded

$$\rho^-(T_1,T_2) \leq \rho(T_1,T_2) \leq \rho^+(T_1,T_2) \tag{105}$$

and the bounds are attained for the Fréchet copulas $\mathbf{C}^-$ and $\mathbf{C}^+$: $\rho^-(T_1,T_2) = \rho(T_1,T_2;\mathbf{C}^-)$ and $\rho^+(T_1,T_2) = \rho(T_1,T_2;\mathbf{C}^+)$. Moreover, we note that

$$\mathbf{C} \succ \mathbf{C}^\perp \Longrightarrow \rho(T_1,T_2;\mathbf{C}) \geq 0 \tag{106}$$

and

$$\mathbf{C} \prec \mathbf{C}^\perp \Longrightarrow \rho(T_1,T_2;\mathbf{C}) \leq 0 \tag{107}$$

So, if the dependence between the survival times is positive (in the PQD sense), the Pearson correlation is positive. In the same way, it is negative in the case of negative dependence (in the NQD sense). We would know characterize more precisely the bounds $\rho^-(T_1,T_2)$ and $\rho^+(T_1,T_2)$. For that, we consider the following interpretation of the copulas $\mathbf{C}^-$ and $\mathbf{C}^+$ given by MIKUSIŃSKI, SHERWOOD and TAYLOR [1991]:

- two random variables T_1 and T_2 are **countermonotonic** — or $\mathbf{C} = \mathbf{C}^-$ — if there exists a r.v. T such that $T_1 = f_1(T)$ and $T_2 = f_2(T)$ with f_1 non-increasing and f_2 non-decreasing;
- two random variables T_1 and T_2 are **comonotonic** — or $\mathbf{C} = \mathbf{C}^+$ — if there exists a random variable T such that $T_1 = f_1(T)$ and $T_2 = f_2(T)$ where the functions f_1 and f_2 are non-decreasing.

We obtain also

$$\rho^-(T_1,T_2) = \rho^+(T_1,T_2) = \frac{\mathbb{E}[f_1(T)f_2(T)] - \mathbb{E}[f_1(T)]\,\mathbb{E}[f_2(T)]}{\sigma[f_1(T)]\,\sigma[f_2(T)]} \tag{108}$$

It is well-known that the solution of the equation $\rho^- = -1$ (or $\rho^+ = 1$) is $f_1(t) = a_1t + b$ and $f_2(t) = a_2t + b$ with $a_1a_2 < 0$ ($a_1a_2 > 0$ for $\rho^+ = 1$). The bounds $\rho^- = -1$ and $\rho^+ = 1$ are then always attained if T_1 and T_2 are gaussian. In other cases, $\rho \in [\rho^-,\rho^+] \subset [-1,1]$. In particular, when there is a non-linear relationship between T_1 and T_2, $[\rho^-,\rho^+] \subset \,]-1,1[$. Let us consider the example due to WANG [1998]. We assume that $T_1 \sim \mathcal{LN}(\mu_1,\sigma_1)$

[13] see also example 4 of NELSEN [1995].
[14] see also example 6 of NELSEN [1995].
[15] see also CAMBANIS, SIMMONS and STOUT [1976] and WHITT [1976].

and $T_2 \sim \mathcal{LN}(\mu_2, \sigma_2)$. We can show that the minimum correlation ρ^- is given when $T_2 = e^{\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1} T_1^{-\frac{\sigma_2}{\sigma_1}}$ and the maximum correlation ρ^+ is given when $T_2 = e^{\mu_2 - \frac{\sigma_2}{\sigma_1}\mu_1} T_1^{\frac{\sigma_2}{\sigma_1}}$. It comes that (see Appendix A.4.2 of WANG [1998])

$$\begin{aligned} \rho^- &= \frac{e^{-\sigma_1\sigma_2} - 1}{\sqrt{e^{\sigma_1^2} - 1}\sqrt{e^{\sigma_2^2} - 1}} \\ \rho^+ &= \frac{e^{\sigma_1\sigma_2} - 1}{\sqrt{e^{\sigma_1^2} - 1}\sqrt{e^{\sigma_2^2} - 1}} \end{aligned} \tag{109}$$

Note that ρ^+ is equal to 1 if and only if $\sigma_1 = \sigma_2$ and ρ^- tends to -1 if $\sigma_1 \vee \sigma_2 \longrightarrow 0^+$. So, in the case where $\sigma_1 \neq \sigma_2$, the permissible range of $\rho(T_1, T_2)$ is not $[-1, 1]$ because $\rho^- > -1$ and $\rho^+ < 1$. Moreover, we have

$$\lim_{\sigma_1 \vee \sigma_2 \longrightarrow \infty} \rho^- = 0 \tag{110}$$

and

$$\lim_{|\sigma_1 - \sigma_2| \longrightarrow \infty} \rho^+ = 0 \tag{111}$$

In Figure 1, we have reported the range of $[\rho^-, \rho^+]$ for different values of σ_1 and σ_2. In some cases, we remark that $\rho^- \gg -1$ and $\rho^+ \ll 1$. For example, Table 2 presents numerical results when $\sigma_1 = 1$ and $\sigma_2 = 3$. We note that the minimum correlation is close to zero even if the dependence between T_1 and T_2 is perfectly negative and the maximum correlation is 0.16! For Kendall's tau $\tau(T_1, T_2)$ and Spearman's rho $\varrho(T_1, T_2)$, the bounds are attained.

Copula	$\rho(T_1, T_2)$	$\tau(T_1, T_2)$	$\varrho(T_1, T_2)$
$\mathbf{C}^-$	-0.008	-1	-1
Normal copula with parameter -0.7	$\simeq 0$	-0.49	-0.68
$\mathbf{C}^\perp$	0	0	0
Normal copula with parameter 0.7	$\simeq 0.10$	0.49	0.68
$\mathbf{C}^+$	0.16	1	1

Table 2: Value of the dependence measures

Remark 12 *In survival modelling, the random variables are positive. That's imply that the lower bound -1 can never been reached:*

$$\rho^- \neq -1 \tag{112}$$

Moreover, because of the distributions generally used (Weibull, Gompertz, etc.), it is very difficult to obtain a large range of correlation. For example, VAN DEN BERG *[1997] shows that $\rho \in \left[-\frac{1}{3}, \frac{1}{2}\right]$ when the baseline hazards are constant.*

2.5.2 Concordance measures

The previous illustrations show that the correlation is a relevant measure of dependence in a few special cases. More appropriate are measures of concordance defined by SCARSINI [1984] (see also NELSEN [1999], page 136). This is the case of the Kendall's tau or the Spearman's rho. Because these two measures have been intensively studied in Chapter 5 of NELSEN [1999], we just give here a result which permits to compute them easily.

Theorem 13 *The Kendall's tau and the Spearman's rho of the survival copula $\breve{\mathbf{C}}$ are equal to the Kendall's tau and the Spearman's rho of the associated copula $\mathbf{C}$.*

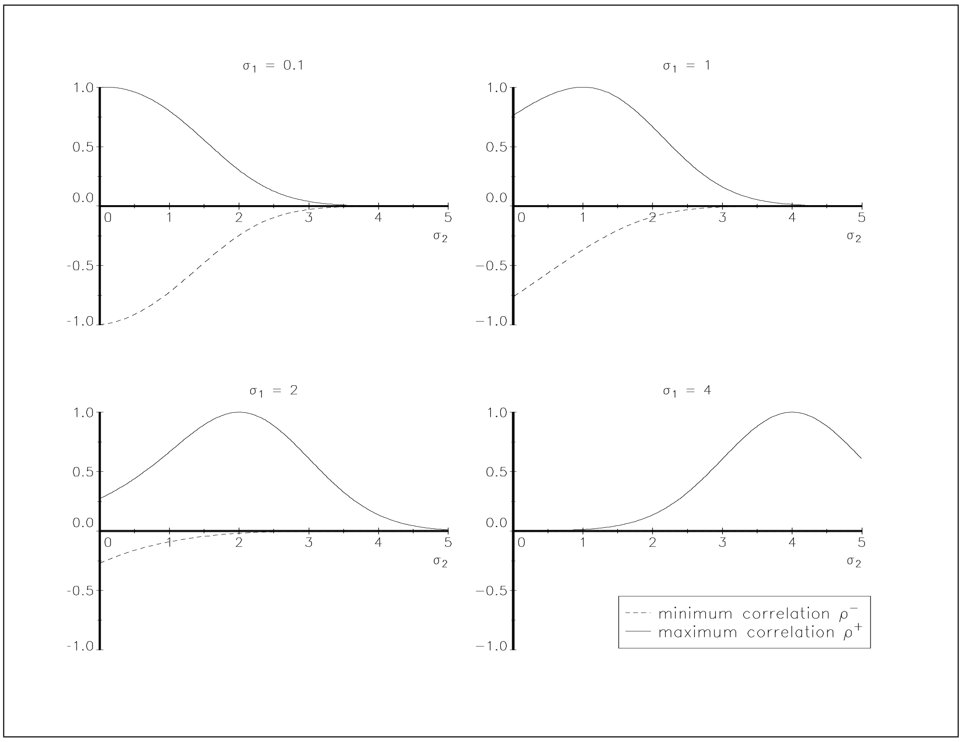

Figure 1: Permissible range of $\rho(T_1, T_2)$ when T_1 and T_2 are two Lognormal random variables

Proof. *We have*

$$\begin{aligned}
\varrho\left\langle\breve{\mathbf{C}}\right\rangle &= 12\iint_{[0,1]^2}\breve{\mathbf{C}}(u_1,u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 - 3 \\
&= 12\left[u_1^2u_2 + u_1u_2^2 - u_1u_2\right]_0^1 + 12\iint_{[0,1]^2}\mathbf{C}(1-u_1,1-u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 - 3 \\
&= \varrho\langle\mathbf{C}\rangle
\end{aligned} \tag{113}$$

and

$$\begin{aligned}
\tau\left\langle\breve{\mathbf{C}}\right\rangle &= 1 - 4\iint_{[0,1]^2}\partial_1\breve{\mathbf{C}}(u_1,u_2)\,\partial_2\breve{\mathbf{C}}(u_1,u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 \\
&= 1 - 4\iint_{[0,1]^2}\left[1-\partial_1\breve{\mathbf{C}}(1-u_1,u_2)\right]\left[1-\partial_2\breve{\mathbf{C}}(1-u_1,u_2)\right]\mathrm{d}u_1\,\mathrm{d}u_2 \\
&= \tau\langle\mathbf{C}\rangle
\end{aligned} \tag{114}$$

■

2.5.3 Tail dependence

τ and ϱ are two global measures of association. Sometimes, it could be useful to characterize the dependence more locally. In particular, we could be interested in tail dependence.

Definition 14 (Joe [1997, p. 33]) *If a bivariate copula* $\mathbf{C}$ *is such that*

$$\lim_{u\to 1}\frac{1-2u+\mathbf{C}(u,u)}{1-u} = \lambda_U \tag{115}$$

exists, then $\mathbf{C}$ *has upper tail dependence if* $\lambda_U \in (0,1]$ *and no upper tail dependence if* $\lambda_U = 0$. *Similarly, if*

$$\lim_{u \to 0} \frac{\mathbf{C}(u,u)}{u} = \lambda_L \tag{116}$$

exists, $\mathbf{C}$ *has lower tail dependence if* $\lambda_L \in (0,1]$ *and no lower tail dependence if* $\lambda_L = 0$.

We could interpret these definitions as follows. Let T_1 and T_2 be two survival times. We remark that λ_U is the limit in one of the following function

$$\begin{aligned}
\lambda_U(u) &= \Pr\left\{T_2 > \mathbf{F}_2^{-1}(u) \mid T_1 > \mathbf{F}_1^{-1}(u)\right\} \\
&= \Pr\left\{T_2 > \mathbf{S}_2^{-1}(1-u) \mid T_1 > \mathbf{S}_1^{-1}(1-u)\right\} \\
&= \frac{1 - 2u + \mathbf{C}(u,u)}{1-u}
\end{aligned} \tag{117}$$

For λ_L, it is the limit in zero of $\lambda_L(u)$ with

$$\begin{aligned}
\lambda_L(u) &= \Pr\left\{T_2 < \mathbf{F}_2^{-1}(u) \mid T_1 < \mathbf{F}_1^{-1}(u)\right\} \\
&= \Pr\left\{T_2 < \mathbf{S}_2^{-1}(1-u) \mid T_1 < \mathbf{S}_1^{-1}(1-u)\right\} \\
&= \frac{\mathbf{C}(u,u)}{u}
\end{aligned} \tag{118}$$

$\lambda_U(u)$ (resp. $\lambda_L(u)$) indicates then the probability that T_2 takes values greater (resp. less) than $t_2 = \mathbf{S}_2^{-1}(1-u)$ given that T_1 is already greater (resp. less) than $t_1 = \mathbf{S}_1^{-1}(1-u)$. $\lambda_U(u)$ and $\lambda_L(u)$ are called by Coles, Currie and Tawn [1999] the quantile-dependent measures of dependence. We note that they depend only on the copula function, not on the margins. Let $\breve{\lambda}_U(u)$ and $\breve{\lambda}_L(u)$ denote the corresponding measures when the copula between T_1 and T_2 is the survival copula $\breve{\mathbf{C}}$. We have

$$\begin{aligned}
\breve{\lambda}_U(u) &= \frac{1 - 2u + \breve{\mathbf{C}}(u,u)}{1-u} \\
&= \frac{\mathbf{C}(1-u, 1-u)}{1-u} \\
&= \lambda_L(1-u)
\end{aligned} \tag{119}$$

and

$$\begin{aligned}
\breve{\lambda}_L(u) &= \frac{2u - 1 + \mathbf{C}(1-u, 1-u)}{u} \\
&= \frac{1 - 2(1-u) + \mathbf{C}(1-u, 1-u)}{1-(1-u)} \\
&= \lambda_U(1-u)
\end{aligned} \tag{120}$$

We also deduce the following theorem.

Theorem 15 *Let* $\breve{\mathbf{C}}$ *be the survival copula associated to* $\mathbf{C}$. *We have*

$$\lambda_U\left\langle \breve{\mathbf{C}} \right\rangle = \lambda_L\left\langle \mathbf{C} \right\rangle \tag{121}$$

and

$$\lambda_L\left\langle \breve{\mathbf{C}} \right\rangle = \lambda_U\left\langle \mathbf{C} \right\rangle \tag{122}$$

The concept of tail dependence is important in survival modelling, because it indicates the behaviour of the joint survival times (T_1, T_2) at the limit cases $(0,0)$ and (∞,∞). They correspond to the cases of "immediate joint death" and "long-term joint survival". In Figure 2, we have represented the two functions $\lambda_U(u)$ and $\lambda_L(u)$ for Gumbel-Hougaard, Cook-Johnson and Normal copulas with Kendall's tau[16] equal to 0.5. We remark that the Gumbel-Hougaard copula has upper tail dependence with $\lambda_U = 2 - 2^{\frac{1}{\theta}}$, whereas the Cook-Johnson copula has lower tail dependence with $\lambda_L = 2^{-\frac{1}{\theta}}$. We verify also that the Normal copula has no upper or lower tail dependence.

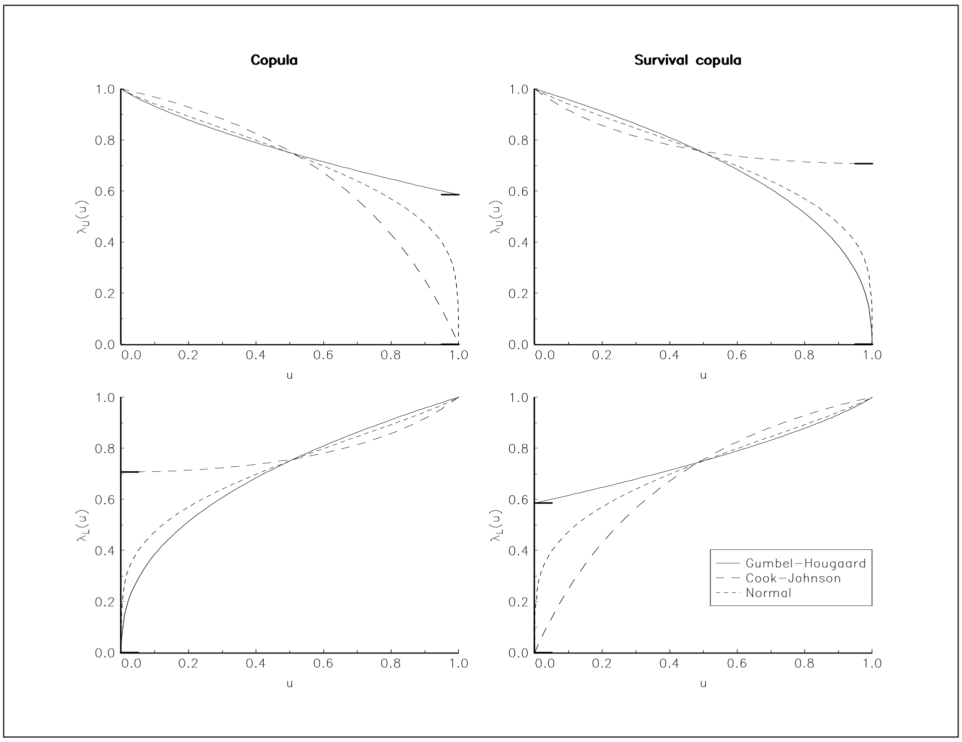

Figure 2: Tail dependence and survival copulas

2.5.4 Time-dependent association measures

In the previous paragraph, we have studied two local dependence measures. In what follows, we consider more specifically time-dependent measures.

The cross-ratio function has been introduced by Clayton [1978] and extensively exploited by David Oakes in several papers (see for example [121] and [122]). We remind that the cross-ratio ϑ is defined in the following way

$$\vartheta(t_1, t_2) = \frac{\lambda(t_1 \mid T_2 = t_2)}{\lambda(t_1 \mid T_2 \geq t_2)} \tag{123}$$

[16] τ is equal respectively to $1-\theta^{-1}$, $(\theta+2)^{-1}\theta$ and $2\pi^{-1}\arcsin(\theta)$ with θ the parameter of the copula.

In the case of a frailty model $\mathbf{S}(t_1,t_2)=\psi\left(\psi^{-1}(\mathbf{S}_1(t_1))+\psi^{-1}(\mathbf{S}_2(t_2))\right)$, OAKES [1989] shows that

$$\begin{aligned}\vartheta(t_1,t_2) &= \frac{\partial_{1,2}\mathbf{S}(t_1,t_2)\times\mathbf{S}(t_1,t_2)}{\partial_1\mathbf{S}(t_1,t_2)\times\partial_2\mathbf{S}(t_1,t_2)} \\ &= \frac{\psi''\left(\psi^{-1}(\mathbf{S}_1(t_1))+\psi^{-1}(\mathbf{S}_2(t_2))\right)}{\left[\psi'\left(\psi^{-1}(\mathbf{S}_1(t_1))+\psi^{-1}(\mathbf{S}_2(t_2))\right)\right]^2}\mathbf{S}(t_1,t_2) \\ &= -\frac{\partial^2\psi^{-1}(\mathbf{S}(t_1,t_2))}{\partial\psi^{-1}(\mathbf{S}(t_1,t_2))}\mathbf{S}(t_1,t_2) \end{aligned} \tag{124}$$

It comes that $\vartheta(t_1,t_2)$ depends on t_1 and t_2 only trough the survival function $\mathbf{S}(t_1,t_2)$:

$$\vartheta(t_1,t_2)=\vartheta(\mathbf{S}(t_1,t_2)) \tag{125}$$

Oakes shows then a stronger result.

Theorem 16 (Oakes [1989, Theorem 1, p. 488]) *Suppose that $\mathbf{S}(t_1,t_2)$ is an absolutely continuous bivariate survival function whose cross-ratio function $\vartheta(t_1,t_2)$ is expressible as $\vartheta(\mathbf{S}(t_1,t_2))$. Then, $\mathbf{S}(t_1,t_2)$ satisfies the Archimedean representation $\mathbf{S}(t_1,t_2)=\psi\left(\psi^{-1}(\mathbf{S}_1(t_1))+\psi^{-1}(\mathbf{S}_2(t_2))\right)$.*

With this theorem, we have implicitely a bijection between the set of frailty models and the set of functions $\vartheta(\mathbf{S}(t_1,t_2))$. Moreover, if we denote $\vartheta(\mathbf{S}(t_1,t_2))$ by $\vartheta(s)$, Oakes shows that the Laplace transform ψ is uniquely determined in terms of $\vartheta(s)$ by

$$\psi^{-1}(y)=\int_y^1\exp\left(\int_t^c s^{-1}\vartheta(s)\,\mathrm{d}s\right)\mathrm{d}t \tag{126}$$

up to a constant multiple $c<1$. For example, in the case of the Cook-Johnson copula, we have

$$\begin{aligned}\vartheta(t_1,t_2) &= -\frac{\theta(\theta+1)[\mathbf{S}(t_1,t_2)]^{-(\theta+2)}}{-\theta[\mathbf{S}(t_1,t_2)]^{-(\theta+1)}}\mathbf{S}(t_1,t_2) \\ &= \theta+1 \\ &= \vartheta(s)\end{aligned} \tag{127}$$

We verify that the inverse of the Laplace transform is given by the expression (126):

$$\begin{aligned}\psi^{-1}(y) &= \int_y^1\exp\left((\theta+1)\int_t^c s^{-1}\,\mathrm{d}s\right)\mathrm{d}t \\ &= \int_y^1 c^{\theta+1}t^{-(\theta+1)}\,\mathrm{d}t \\ &= \frac{c^{\theta+1}}{\theta}\left(y^{-\theta}-1\right) \\ &\propto \left(y^{-\theta}-1\right)\end{aligned} \tag{128}$$

For the Gumbel-Hougaard copula, we have

$$\begin{aligned}\vartheta(t_1,t_2) &= -\frac{\left(\theta(\theta-1)[-\ln\mathbf{S}(t_1,t_2)]^{\theta-2}+\theta[-\ln\mathbf{S}(t_1,t_2)]^{\theta-1}\right)\mathbf{S}^{-2}(t_1,t_2)}{-\theta\mathbf{S}^{-1}(t_1,t_2)[-\ln\mathbf{S}(t_1,t_2)]^{\theta-1}}\mathbf{S}(t_1,t_2) \\ &= 1-\frac{\theta-1}{\ln\mathbf{S}(t_1,t_2)}\end{aligned} \tag{129}$$

and

$$\vartheta(s) = 1 - \frac{\theta - 1}{\ln s} \tag{130}$$

Using the expression (126), we obtain

$$\begin{aligned}
\psi^{-1}(y) &= \int_y^1 \exp\left(\int_t^c \left(1 - \frac{\theta - 1}{\ln s}\right) s^{-1}\,\mathrm{d}s\right)\mathrm{d}t \\
&= \int_y^1 c(-\ln c)^{1-\theta}\, t^{-1}(-\ln t)^{\theta-1}\,\mathrm{d}t \\
&= \frac{c(-\ln c)^{1-\theta}}{\theta}(-\ln y)^{\theta} \\
&\propto (-\ln y)^{\theta}
\end{aligned} \tag{131}$$

which is the inverse of the Laplace transform $\psi(x) = \exp\left(-x^{\frac{1}{\theta}}\right)$. In the case where the model is not frailty, the cross-ratio could not be expressed as a function of the joint survival function. For example, if we consider the FGM copula $\mathbf{C}(u_1, u_2) = u_1 u_2 (1 - (1 - u_1)(1 - u_2))$, the expression of $\vartheta(t_1, t_2)$ is

$$\begin{aligned}
\vartheta(t_1, t_2) &= \frac{\partial_{1,2}\breve{\mathbf{C}}(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2)) \times \breve{\mathbf{C}}(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2))}{\partial_1\breve{\mathbf{C}}(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2)) \times \partial_2\breve{\mathbf{C}}(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2))} \\
&= \frac{[2(S_1(t_1) + S_2(t_2)) - 4S_1(t_1)S_2(t_2)]\left[S_1^2(t_1)S_2(t_2) + S_1(t_1)S_2^2(t_2) - S_1^2(t_1)S_2^2(t_2)\right]}{[2S_1(t_1)S_2(t_2) + S_2^2(t_2) - 2S_1(t_1)S_2^2(t_2)][2S_1(t_1)S_2(t_2) + S_1^2(t_1) - 2S_1^2(t_1)S_2(t_2)]}
\end{aligned} \tag{132}$$

At the points $(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2)) = (0.5, 0.5)$ and $(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2)) = (0.1875, 1)$, we have $\breve{\mathbf{C}}(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2)) = 0.1875$. But $\vartheta(t_1, t_2)$ takes two different values 0.75 and 0.8966.

Oakes [1989] shows that the cross-ratio $\vartheta(t_1, t_2)$ is related to a *conditional* version of Kendall's tau:

$$\begin{aligned}
\tau^\star(t_1, t_2) = {} & \Pr\{(T_1 - T_1')(T_2 - T_2') > 0 \mid \min(T_1, T_1') = t_1, \min(T_2, T_2') = t_2\} - \\
& \Pr\{(T_1 - T_1')(T_2 - T_2') < 0 \mid \min(T_1, T_1') = t_1, \min(T_2, T_2') = t_2\}
\end{aligned} \tag{133}$$

where (T_1, T_2) and (T_1', T_2') are two independent and identically distributed or *i.i.d.* random vector, each with joint survival function $\mathbf{S}$. Because we have

$$\vartheta(t_1, t_2) = \frac{\Pr\{(T_1 - T_1')(T_2 - T_2') > 0 \mid \min(T_1, T_1') = t_1, \min(T_2, T_2') = t_2\}}{\Pr\{(T_1 - T_1')(T_2 - T_2') < 0 \mid \min(T_1, T_1') = t_1, \min(T_2, T_2') = t_2\}} \tag{134}$$

it comes that

$$\vartheta(t_1, t_2) = \frac{1 + \tau^\star(t_1, t_2)}{1 - \tau^\star(t_1, t_2)} \tag{135}$$

In the case where the survival times are PQD, we know that $\tau \geq 0$. "*Thus we should expect $\vartheta(t_1, t_2) \geq 1$ for many (t_1, t_2). But $\vartheta(t_1, t_2)$ can be less than 1*" (Anderson and Louis [1995], page 672). However, in most cases, we verify that $\vartheta(t_1, t_2) \geq 1$ if $\breve{\mathbf{C}} \succ \mathbf{C}^{\perp}$. The principal interest of $\vartheta(t_1, t_2)$ is to give an information about the strength of the local dependence at the point (t_1, t_2). For example, Figure 3 shows that two copulas with same Kendall's tau may have different values of $\vartheta(t_1, t_2)$. Moreover, we remark the particular behaviour of the Gumbel-Hougaard copula. Even for θ close to one — the independent case, $\vartheta(\infty, \infty) = \infty$. Figure 4 corresponds to the Normal copula with Kendall's tau equal to 0.5. We see that $\vartheta(t_1, t_2)$ can not be expressible as $\vartheta(\mathbf{S}(t_1, t_2))$: given a level curve $s = \tilde{\mathbf{C}}(u_1, u_2)$, we verify that $\vartheta(s)$ is not constant.

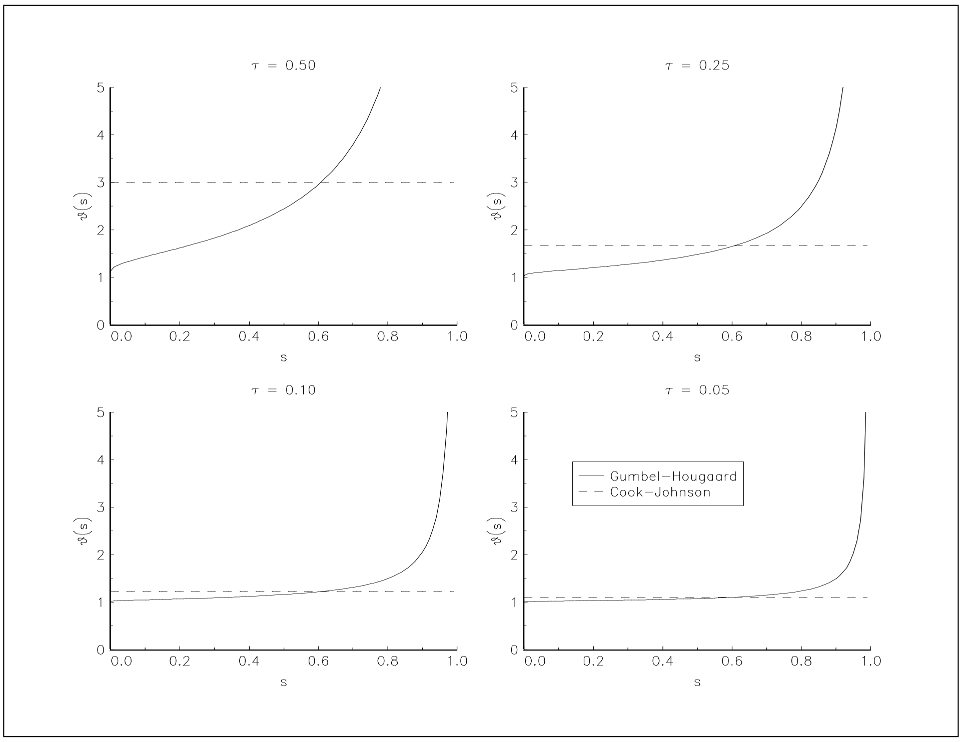

Figure 3: Cross-ratio of the Gumbel-Hougaard and Cook-Johnson copulas

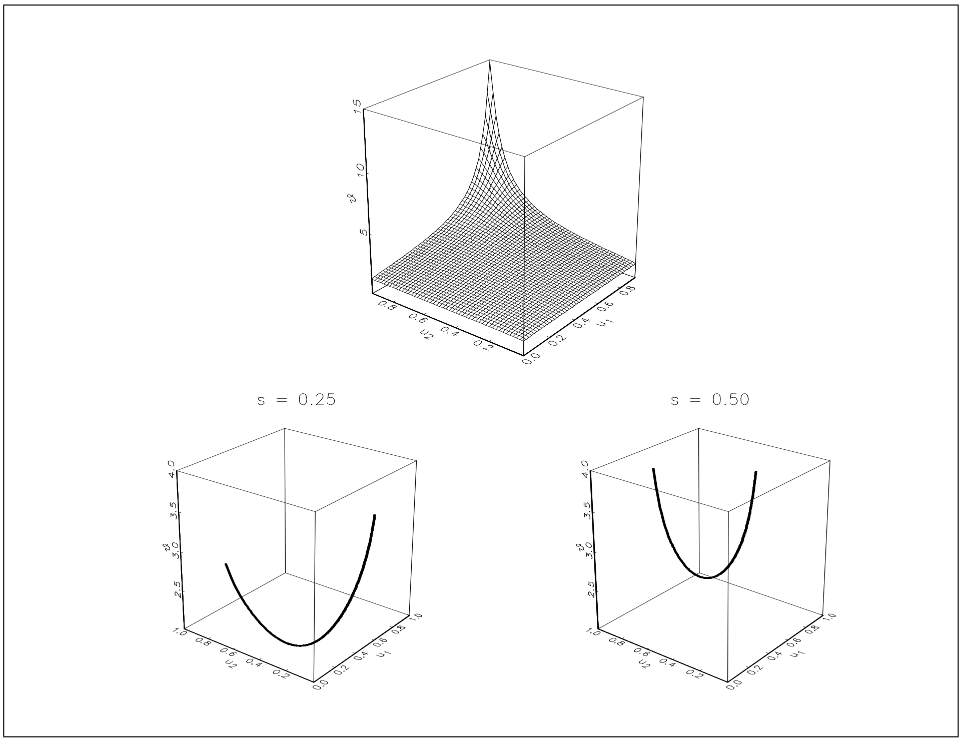

Figure 4: Cross-ratio of the Normal copula

There are a lot of other time-dependent measures of dependence. For example, Anderson, Louis, Holm and Harvald [1992] proposes to use the conditional expected residual life and the conditional probability. Manatunga and Oakes [1996] define a *truncated* Kendall's tau $\tau^*(t_1,t_2)$ which corresponds to the Kendall's tau of the conditional survival function $\mathbf{S}$ given that $T_1 > t_1$ and $T_2 > t_2$. Fan, Prentice and Hsu [2000] consider a class of measures that are weighted averages of local dependence measures:

$$D(t_1,t_2) = \int_0^{t_1}\int_0^{t_2} \frac{\kappa(\vartheta(x_1,x_2))}{\int_0^{t_1}\int_0^{t_2}\partial_{1,2}\varpi(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_1}\partial_{1,2}\varpi(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_1 \tag{136}$$

where κ is a function of ϑ and ϖ is a weight function. For example, if ϖ is the survival function $\mathbf{S}$ and κ is the identity function, we obtain

$$D(t_1,t_2) = \int_0^{t_1}\int_0^{t_2} \frac{\vartheta(x_1,x_2)}{\int_0^{t_1}\int_0^{t_2} f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_1} f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_1 \tag{137}$$

2.5.5 Concepts of ageing

In this paragraph, we consider some notions of ageing, which are useful to build reliability classes. In particular, we show how they are (or not) related to the properties of copulas. The idea is the same as for the class of bivariate extreme value distributions: their margins are of the three-types (Gumbel, Fréchet or Weibull) and the copula is necessarily an extreme value copula — $\mathbf{C}(u_1^t,u_2^t) = \mathbf{C}^t(u_1,u_2)$ for $t > 0$.

As remarked by Hutchinson and Lai [1990,1991], several definitions are possible to extend univariate ageing properties. For example, a distribution $\mathbf{F}$ is said to be IFR (Increasing Failure Rate) if $\mathbf{S}(t \mid x)$ is nonincreasing in x for each $t \geq 0$. In the bivariate case, there are several variants of the IFR property. The more general extension is the following: $\mathbf{S}(x_1+t_1,x_2+t_2)/\mathbf{S}(x_1,x_2)$ decreases in $x_1,x_2 \geq 0$ for all $t_1,t_2 \geq 0$. However, this condition is too restrictive to obtain a "large" reliability class. One may prefer to define the bivariate IFR property by the condition:

$$\mathbf{S}(x_1+t,x_2+t)/\mathbf{S}(x_1,x_2) \text{ decreases in } x_1,x_2 \geq 0 \text{ for all } t \geq 0 \tag{138}$$

We note that it is equivalent to $\breve{\mathbf{C}}(\mathbf{S}_1(x_1+t),\mathbf{S}_2(x_2+t))/\breve{\mathbf{C}}(\mathbf{S}_1(x_1),\mathbf{S}_2(x_2))$ decreases in $x_1,x_2 \geq 0$ for all $t \geq 0$. Because we expect that the margins belongs to the corresponding univariate reliability class, we assume that $\mathbf{S}_n(x+t)/\mathbf{S}_n(x)$ decreases in $x \geq 0$ for all $t \geq 0$. It comes that

$$\begin{aligned}\frac{\breve{\mathbf{C}}(\mathbf{S}_1(x_1+t),\mathbf{S}_2(x_2+t))}{\breve{\mathbf{C}}(\mathbf{S}_1(x_1),\mathbf{S}_2(x_2))} &= \frac{\breve{\mathbf{C}}(\mathbf{S}_1(t \mid x_1)\mathbf{S}_1(x_1),\mathbf{S}_2(t \mid x_2)\mathbf{S}_2(x_2))}{\breve{\mathbf{C}}(\mathbf{S}_1(x_1),\mathbf{S}_2(x_2))} \\ &= \frac{\breve{\mathbf{C}}(v_1(u_1)\cdot u_1, v_2(u_2)\cdot u_2)}{\breve{\mathbf{C}}(u_1,u_2)}\end{aligned} \tag{139}$$

with $v_n(u_n) = \mathbf{S}_n(t \mid x_n)$ and $u_n = \mathbf{S}_n(x_n)$. The idea is to define an IFR reliability class thanks to IFR copulas. In this special case, it seems very difficult to characterize all IFR copulas even if we use the fact that $v_n(u_n)$ decreases in u_n.

In general, we encounter these difficulties for almost all reliability classes. For example, Basu [1971] shows that the only absolutely continuous survival function with constant hazard rate $\lambda(t_1,t_2) = \lambda$ is the survival function of two independent exponential random variables. Let us now consider the LMP class:

$$\mathbf{S}(x_1+t_1,x_2+t_2) = \mathbf{S}(x_1,x_2)\,\mathbf{S}(t_1,t_2) \text{ for all } x_1,x_2 \geq 0 \text{ and } t_1,t_2 \geq 0 \tag{140}$$

Marshall and Olkin [1967] show that the solution of this functional equation is the product of two exponential survival functions. So, this property is too strong, and one may prefer to define the LMP class by the following assumption:

$$\mathbf{S}(x_1+t,x_2+t) = \mathbf{S}(x_1,x_2)\,\mathbf{S}(t,t) \text{ for all } x_1,x_2 \geq 0 \text{ and } t \geq 0 \tag{141}$$

By using the fact that the margins are LMP, we have $\breve{\mathbf{C}}(u_1v_1, u_2v_2) = \breve{\mathbf{C}}(u_1, u_2)\,\breve{\mathbf{C}}(v_1, v_2)$ with $u_n = \mathbf{S}_n(x_n)$ and $v_n = \mathbf{S}_n(t)$. If we define a LMP copula as a copula function which verifies $\breve{\mathbf{C}}(u_1v, u_2v) = \breve{\mathbf{C}}(u_1, u_2)\,\breve{\mathbf{C}}(v_1, v_2)$ for all u_1, u_2, v_1, v_2 in $[0,1]$, it comes that if the copula and the margins are LMP, the distribution $\mathbf{F}$ is LMP too. For example, $\mathbf{C}^{\perp}$ is a LMP copula. However, the previous characterization appears too strong: if we solve the functional equation $\breve{\mathbf{C}}(u_1v, u_2v) = \breve{\mathbf{C}}(u_1, u_2)\,\breve{\mathbf{C}}(v_1, v_2)$, the only solution is $\mathbf{C}^{\perp}$. We could then use the relationships between u_1, u_2, v_1 and v_2 to obtain a better characterization.

Definition 17 *A copula function is LMP if it verifies*

$$\breve{\mathbf{C}}\left(u^{1+\alpha_1}, u^{\alpha_2(\alpha_1+\alpha_3)}\right) = \breve{\mathbf{C}}(u, u^{\alpha_2\alpha_3})\,\breve{\mathbf{C}}(u^{\alpha_1}, u^{\alpha_1\alpha_2}) \tag{142}$$

for all $\alpha_1, \alpha_2, \alpha_3 \geq 0$.

We remark that the survival copula of a LMP copula is necessarily an extreme value copula.

Another interesting property is the IFRA (Increasing Failure Rate on Average) notion. Block and Savits [1980] define it as follows: $\mathbf{F}$ is said to be IFRA if $\mathbb{E}^{\alpha}[h(T_1, T_2)] \leq \mathbb{E}[h^{\alpha}(T_1, T_2)/\alpha]$ for all continuous nonnegative increasing functions h and $\alpha \in [0,1]$. In this case, we could show that if $\mathbf{F}$ is IFRA, then $\mathbf{S}^{\alpha}(t_1, t_2) \leq \mathbf{S}(\alpha t_1, \alpha t_2)$.

Proposition 18 *If the survival copula verifies $\breve{\mathbf{C}}^{\alpha}(u_1, u_2) \leq \breve{\mathbf{C}}(u_1^{\alpha}, u_2^{\alpha})$ for all u_1, u_2 in $[0,1]$ and $\alpha \in [0,1]$, and if the margins are IFRA, then the distribution $\mathbf{F}$ is IFRA.*

Proof. We have $\breve{\mathbf{C}}^{\alpha}(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2)) \leq \breve{\mathbf{C}}(\mathbf{S}_1^{\alpha}(t_1), \mathbf{S}_2^{\alpha}(t_2))$. But $\breve{\mathbf{C}}(\mathbf{S}_1^{\alpha}(t_1), \mathbf{S}_2^{\alpha}(t_2)) \leq \breve{\mathbf{C}}(\mathbf{S}_1(\alpha t_1), \mathbf{S}_2(\alpha t_2))$ because the margins are IFRA. It comes that

$$\breve{\mathbf{C}}^{\alpha}(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2)) \leq \breve{\mathbf{C}}(\mathbf{S}_1^{\alpha}(t_1), \mathbf{S}_2^{\alpha}(t_2)) \leq \breve{\mathbf{C}}(\mathbf{S}_1(\alpha t_1), \mathbf{S}_2(\alpha t_2)) \tag{143}$$

■

We consider a last reliability class. If we define the bivariate notion of NBU (New Better than Used) as follows:

$$\mathbf{S}(x_1 + t_1, x_2 + t_2) \leq \mathbf{S}(x_1, x_2)\,\mathbf{S}(t_1, t_2) \text{ for all } x_1, x_2 \geq 0 \text{ and } t_1, t_2 \geq 0 \tag{144}$$

we will say that a NBU copula is a copula function such that

$$\breve{\mathbf{C}}(u_1v_1, u_2v_2) \leq \breve{\mathbf{C}}(u_1, u_2)\,\breve{\mathbf{C}}(v_1, v_2) \tag{145}$$

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

Proposition 19 *If the copula and the margins are NBU, then the distribution $\mathbf{F}$ is NBU.*

Proof. We have

$$\breve{\mathbf{C}}(\mathbf{S}_1(x_1)\,\mathbf{S}_1(t_1), \mathbf{S}_2(x_2)\,\mathbf{S}_2(t_2)) \leq \breve{\mathbf{C}}(\mathbf{S}_1(x_1), \mathbf{S}_2(x_2))\,\breve{\mathbf{C}}(\mathbf{S}_1(t_1), \mathbf{S}_2(x_2)) \tag{146}$$

and

$$\breve{\mathbf{C}}(\mathbf{S}_1(x_1 + t_1), \mathbf{S}_2(x_2 + t_2)) \leq \breve{\mathbf{C}}(\mathbf{S}_1(x_1)\,\mathbf{S}_1(t_1), \mathbf{S}_2(x_2)\,\mathbf{S}_2(t_2)) \tag{147}$$

■

Example 20 *The bivariate exponential distribution (69) given by* Gumbel *[1960] is NBU. To show that, we remark that the margins are NBU and the copula verifies*

$$\begin{aligned}
\breve{\mathbf{C}}(u_1v_1, u_2v_2) &= u_1v_1u_2v_2 \exp(-\theta \ln(u_1v_1)\ln(u_2v_2)) \\
&= u_1u_2 \exp(-\theta \ln u_1 \ln u_2) \cdot v_1v_2 \exp(-\theta \ln v_1 \ln v_2) \cdot \exp(-\theta(\ln u_1 \ln v_2 + \ln v_1 \ln u_2)) \\
&\leq \breve{\mathbf{C}}(u_1, u_2)\,\breve{\mathbf{C}}(v_1, v_2)
\end{aligned} \tag{148}$$

2.6 Deheuvels copulas and empirical processes of survival times

In this paragraph, we consider empirical processes of survival times. We just give here the main idea, because it may be too long to define the notations and the mathematical tools. Moreover, a forthcoming working paper considers the more general problem of empirical processes and copulas. In this working paper, a section is dedicated to empirical processes of survival times.

If we consider classic textbooks on empirical processes (Csörgő [1983], Shorack and Wellner [1986], Pollard [1990]), we remark that they do not use copulas, or more precisely the term "copula". Nevertheless, they are briefly mentioned in some books (Gaenssler and Stute [1987], van der Vaart and Wellner [1996]). However, if one read between the lines, copulas are almost everywhere. We may then study empirical processes in the point of view of copulas. One of the main tool is also the empirical dependence (copula) function[17] $\hat{\mathbf{C}}_n$ defined by Deheuvels [1978,1979b]. For example, Deheuvels [1980] shows that under general assumptions, we have

$$\max_{\mathbf{u}} \left| \hat{\mathbf{C}}_n(\mathbf{u}) - \mathbb{E}\left[\hat{\mathbf{C}}_n(\mathbf{u})\right] \right| = O\left(\left(\frac{n}{\ln(\ln n)} \right)^{-1/2} \right) \tag{149}$$

with probability one. Let $\mathbf{W}_n(\mathbf{u}) = n^{1/2}\left(\hat{\mathbf{C}}_n(\mathbf{u}) - \mathbf{C}^{\perp}(\mathbf{u})\right)$ be the empirical dependence process. Deheuvels [1981] shows another important result. In the space $\mathbf{C}[0,1]$ (see Billingsley [1999]), the empirical dependence process converges weakly to a limiting Gaussian process on $[0,1]^N$:

$$\mathbf{W}_n(\mathbf{u}) \longrightarrow \mathbf{W}_\infty(\mathbf{u}) \tag{150}$$

with

$$\mathbb{E}\left[\mathbf{W}_\infty(\mathbf{u})\,\mathbf{W}_\infty(\mathbf{v})\right] = \prod_{i=1}^N \min(u_i, v_i) - \sum_{i=1}^N \min(u_i, v_i) \prod_{i \neq j}^N u_i v_i + (N-1) \prod_{i=1}^N u_i v_i \tag{151}$$

The problem is now to define similar canonical decomposition of empirical processes of survival times. This problem has been solved by Dabrowska [1996] in the general case. Let $\prod\!\!\!\!\!\!\;\,\prod$ be the product-integral (see Gill and Johansen [1990]). In the univariate case, we have

$$\mathbf{S}(t) = \prod\!\!\!\!\!\!\!\!\;\prod_{(0,t]} (1 - \mathrm{d}\Lambda) \tag{152}$$

where Λ is a measure defined by $\Lambda(\mathrm{d}t) = \mathbf{F}(\mathrm{d}t)/\mathbf{S}(t)$. In the multivariate case, this relationship becomes $\mathbf{S}(\mathbf{t}) = \prod\!\!\!\!\!\!\;\,\prod (1 + \mathbf{L}(\mathrm{d}\mathbf{t}))$ where $\mathbf{L}$ is called the *iterated odds ratio measure* or *cumulant measure* (Gill [1994]). In the bivariate case, the previous expression takes a simple form because $\mathbf{S}(t_1, t_2) = \mathbf{S}_1(t_1)\,\mathbf{S}_2(t_2)\,e^{\Lambda(t_1,t_2)}$. In the general multivariate case, the expression of $\mathbf{L}$ is very complicated (see Gill and Johansen [1990] or Gill [1992b]). However, $\mathbf{L}$ could be viewed as a dependence measure. Using the previous framework, Dabrowska [1996] provides a decomposition of survival functions in terms of signed interaction measures. Moreover, she generalizes to censored data Deheuvels's decomposition of empirical copula functions.

3 Competing risks models

Competing risks (CR) models arise in several fields: reliability, biometrics, finance, etc. They correspond to the study of any failure process in which they are different causes of failure. To present the concepts used in CR models, we follows the seminal paper of Prentice, Kalbfleisch, Peterson, Flournoy, Farewell and Breslow [1978] (referred to as 'PK' in the sequel).

[17] In this paragraph, n denotes the size of the sample.

3.1 Some definitions

Let τ be the failure time with survival function $\mathbf{S}_\tau$ and hazard rate λ_τ. The type of failure is denoted by $n_\tau \in \{1, \ldots, N\}$. The *cause-specific* hazard function λ_n^* is defined by

$$\lambda_n^*(t) = \lim_{\Delta \to 0^+} \frac{1}{\Delta} \Pr\{t \leq \tau \leq t + \Delta \mid \tau \geq t, n_\tau = n\} \tag{153}$$

If we assume that the types of failure are distinct, we obtain

$$\lambda_\tau(t) = \sum_{n=1}^{N} \lambda_n^*(t) \tag{154}$$

We have also

$$\begin{aligned} \mathbf{S}_\tau(t) &= \exp\left(-\int_0^t \sum_{n=1}^{N} \lambda_n^*(s)\, \mathrm{d}s\right) \\ &= \prod_{n=1}^{N} \mathbf{S}_n^*(t) \end{aligned} \tag{155}$$

with $\mathbf{S}_n^*(t) = \exp\left(-\int_0^t \lambda_n^*(s)\, \mathrm{d}s\right)$. Note however that $\mathbf{S}_n^*$ is not a survival function. With this framework, the model could be identified, that is, "*the cause-specific hazard functions have the potential to be directly estimated from data*" (PK [1978]).

Another framework is to formulate the CR model in terms of latent failure times $T_1, \ldots, T_N$. In this case, the failure time corresponds to

$$\tau = \min(T_1, \ldots, T_N) \tag{156}$$

We have

$$\begin{aligned} \mathbf{S}_\tau(t) &= \Pr\{\min(T_1, \ldots, T_N) \geq t\} \\ &= \breve{\mathbf{C}}(\mathbf{S}_1(t), \ldots, \mathbf{S}_N(t)) \end{aligned} \tag{157}$$

with $\breve{\mathbf{C}}$ the survival copula of the latent times $T_1, \ldots, T_N$. The model with latent failure times have been criticized by PK [1978] because of identifiability problems. For example, TSIATIS [1975] shows that for every joint survival function of latent failure times, a joint survival function of **independent** latent failure times gives the same observable data. However, HECKMAN and HONORÉ [1989] show that the model could be (non-parametrically) identified under some conditions. If one could write the joint survival function $\mathbf{S}$ as

$$\mathbf{S}(t_1, \ldots, t_N) = \breve{\mathbf{C}}\left(e^{-Z_1(t_1)\phi_1(X)}, \ldots, e^{-Z_N(t_N)\phi_N(X)}\right) \tag{158}$$

some regulatory conditions on $\breve{\mathbf{C}}$, Z_n and ϕ ensure that these functions are identified (HECKMAN and HONORÉ [1989], HONORÉ [1993]). This result has been extended by ABBRING and VAN DEN BERG [2000], which prove that the conditions can be weakened for frailty models. Moreover, if the copula is known, "*competing risks data are sufficient to identify the marginal survival functions*" (ZHENG and KLEIN [1994,1995]).

In this section, we use the representation in terms of latent failure times. It is the most familiar in finance, and **non-parametrically** identification problems could be ignored in most of financial problems.

3.2 Distribution of the failure time

The failure time is characterized by the following theorem.

Theorem 21 *The survival function of the failure time τ is given by the diagonal section of the survival distribution:*

$$\mathbf{S}_{\tau}(t) = \breve{\mathbf{C}}(\mathbf{S}_1(t), \ldots, \mathbf{S}_N(t)) \tag{159}$$

Proof. We have

$$\begin{aligned} \mathbf{S}_{\tau}(t) &= \Pr\{\min(T_1, \ldots, T_N) \geq t\} \\ &= \Pr\{T_1 \geq t, \ldots, T_N \geq t\} \\ &= \breve{\mathbf{C}}(\mathbf{S}_1(t), \ldots, \mathbf{S}_N(t)) \end{aligned} \tag{160}$$

■

We can then establish related results about the distribution and the density. It comes that

$$\begin{aligned} \mathbf{F}_{\tau}(t) &= 1 - \breve{\mathbf{C}}(\mathbf{S}_1(t), \ldots, \mathbf{S}_N(t)) \\ &= 1 - \breve{\mathbf{C}}(1 - \mathbf{F}_1(t), \ldots, 1 - \mathbf{F}_N(t)) \end{aligned} \tag{161}$$

and

$$f_{\tau}(t) = \sum_{n=1}^{N} \partial_n \breve{\mathbf{C}}(\mathbf{S}_1(t), \ldots, \mathbf{S}_N(t)) \times f_n(t) \tag{162}$$

In the case where the survival times are *i.i.d.*, we retreive the well-known results

$$\mathbf{F}_{\tau}(t) = 1 - [1 - \mathbf{F}_1(t)]^N \tag{163}$$

and

$$\begin{aligned} f_{\tau}(t) &= \sum_{n=1}^{N} \left(\underbrace{[1 - \mathbf{F}_1(t)] \times \ldots \times [1 - \mathbf{F}_1(t)]}_{N-1 \text{ times}} \right) f_1(t) \\ &= N[1 - \mathbf{F}_1(t)]^{N-1} f_1(t) \end{aligned} \tag{164}$$

Example 22 *Let us consider the case of the* Normal *copula with matrix $\boldsymbol{\rho}$ of parameters. We define $P_{(n)}$ as the $N \times N$ permutation matrix with encoding $p = \begin{pmatrix} 1 & \cdots & n-1 & n+1 & \cdots & N & n \end{pmatrix}$ and $\bar{\boldsymbol{\rho}}_{(n)}$ as follows*

$$\bar{\boldsymbol{\rho}}_{(n)} = P_{(n)} \boldsymbol{\rho} P_{(n)}^{\top} \tag{165}$$

The density of τ is then

$$f_{\tau}(t) = \sum_{n=1}^{N} \mathbf{C}\left(\mathbf{S}_1^{\star}(t), \ldots, \mathbf{S}_{n-1}^{\star}(t), \mathbf{S}_{n+1}^{\star}(t), \ldots, \mathbf{S}_N^{\star}(t); \boldsymbol{\rho}_{(n)}^{\star}\right) \times f_n(t) \tag{166}$$

with $\mathbf{C}$ the Normal *copula function of dimension $N-1$. The matrix $\boldsymbol{\rho}_{(n)}^{\star}$ is defined as follows*

$$\boldsymbol{\rho}_{(n)}^{\star} = \left[\bar{\boldsymbol{\rho}}_{11} - \frac{1}{\boldsymbol{\rho}_{n,n}} \bar{\boldsymbol{\rho}}_{12}^{\top} \bar{\boldsymbol{\rho}}_{12}^{\top}\right] \setminus \boldsymbol{\sigma} \setminus \boldsymbol{\sigma}^{\top} \tag{167}$$

with

$$\underset{N\times N}{\bar{\boldsymbol{\rho}}_{(n)}} = \begin{bmatrix} \underset{(N-1)\times(N-1)}{\bar{\boldsymbol{\rho}}_{11}} & \underset{(N-1)\times 1}{\bar{\boldsymbol{\rho}}_{12}} \\ & \underset{1\times 1}{\bar{\boldsymbol{\rho}}_{22}} \end{bmatrix} \tag{168}$$

and

$$\boldsymbol{\sigma}\odot\boldsymbol{\sigma} = \begin{bmatrix} \boldsymbol{\rho}_{1,1} - \frac{\boldsymbol{\rho}_{1,n}^2}{\boldsymbol{\rho}_{n,n}} \\ \vdots \\ \boldsymbol{\rho}_{n-1,-1} - \frac{\boldsymbol{\rho}_{n-1,n}^2}{\boldsymbol{\rho}_{n,n}} \\ \boldsymbol{\rho}_{n+1,n+1} - \frac{\boldsymbol{\rho}_{n+1,n}^2}{\boldsymbol{\rho}_{n,n}} \\ \vdots \\ \boldsymbol{\rho}_{N,N} - \frac{\boldsymbol{\rho}_{N,n}^2}{\boldsymbol{\rho}_{n,n}} \end{bmatrix} \tag{169}$$

Note also that

$$\mathbf{S}_i^\star(t) = \Phi\left(\frac{\Phi^{-1}(\mathbf{S}_i(t)) - \frac{\boldsymbol{\rho}_{i,n}}{\boldsymbol{\rho}_{n,n}}\Phi^{-1}(\mathbf{S}_n(t))}{\sqrt{\boldsymbol{\rho}_{i,i} - \frac{\boldsymbol{\rho}_{i,n}^2}{\boldsymbol{\rho}_{n,n}}}}\right) \tag{170}$$

We consider now some illustrations based on $\mathsf{Weibull}$ survival times $T_n \sim \mathsf{Weibull}\left(\lambda_n^0, \gamma_n\right)$. We have

$$\mathbf{S}_n(t) = \exp\left(-\lambda_n^0 t^{\gamma_n}\right) \tag{171}$$

The hazard rate $\lambda_n(t)$ is then $\lambda_n^0\gamma_n t^{\gamma_n-1}$ and the expression of the density is $f_n(t) = \lambda_n^0\gamma_n t^{\gamma_n-1}e^{-\lambda_n^0 t^{\gamma_n}}$. We remark that the hazard rate is increasing in the case $\gamma > 1$ and decreasing in the case $0 < \gamma < 1$ (see Figures 5 and 7). If γ is equal to one, the hazard rate is constant and we obtain an exponential survival time. If we assume that the survival copula is the $\mathsf{Gumbel\text{-}Hougaard}$ copula with parameter $\theta \geq 1$, we obtain

$$\begin{aligned} \mathbf{S}_\tau(t) &= \exp\left(-\left[(-\ln\mathbf{S}_1(t))^\theta + \ldots + (-\ln\mathbf{S}_N(t))^\theta\right]^{\frac{1}{\theta}}\right) \\ &= \exp\left(-\left[\sum_{n=1}^N \left(\lambda_n^0 t^{\gamma_n}\right)^\theta\right]^{\frac{1}{\theta}}\right) \end{aligned} \tag{172}$$

and

$$f_\tau(t) = \left[\sum_{n=1}^N \left(\lambda_n^0 t^{\gamma_n}\right)^\theta\right]^{\frac{1}{\theta}-1}\left[\sum_{n=1}^N \frac{\gamma_n}{t}\left(\lambda_n^0 t^{\gamma_n}\right)^\theta\right]\exp\left(-\left[\sum_{n=1}^N \left(\lambda_n^0 t^{\gamma_n}\right)^\theta\right]^{\frac{1}{\theta}}\right) \tag{173}$$

In the case where the survival times are identically distributed, the failure time is a $\mathsf{Weibull}$ survival time $\tau \sim \mathsf{Weibull}\left(N^{\frac{1}{\theta}}\lambda^0, \gamma\right)$ — $\mathbf{S}_\tau(t) = \exp\left(-N^{\frac{1}{\theta}}\lambda^0 t^\gamma\right)$. Moreover if $\theta = 1$ — the survival copula becomes the product copula — $\tau \sim \mathsf{Weibull}\left(N\lambda^0, \gamma\right)$ and we say that the $\mathsf{Weibull}$ distribution is a **min-stable** distribution (Resnick [1987]). This property explains that the survival distribution (172) have been studied in reliability theory (Lee [1979]). Figures 6 and 8 present $\mathbf{S}_\tau(t)$ and $f_\tau(t)$ when the survival times are identically distributed. We remark that the case $\gamma > 1$ is very different from the case $\gamma < 1$. Moreover, we see that the parameter θ has a big influence of the failure time. Note that $\theta \to \infty$ corresponds to the upper Fréchet copula. In this case, we have $\tau = T_1$.

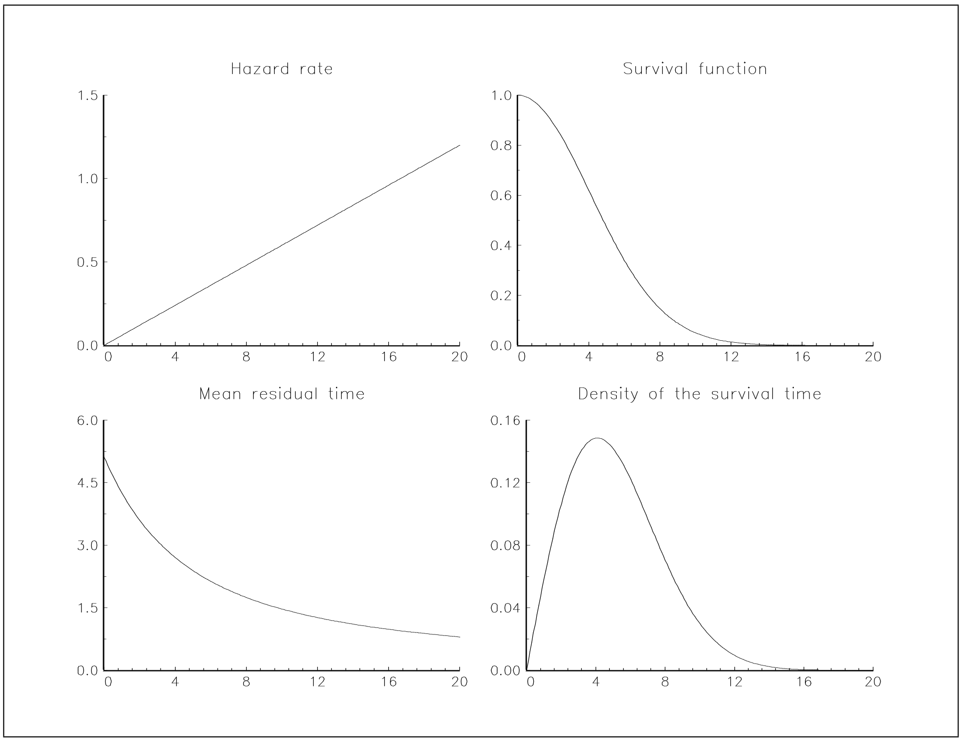

Figure 5: Weibull $(0.03, 2)$ survival time

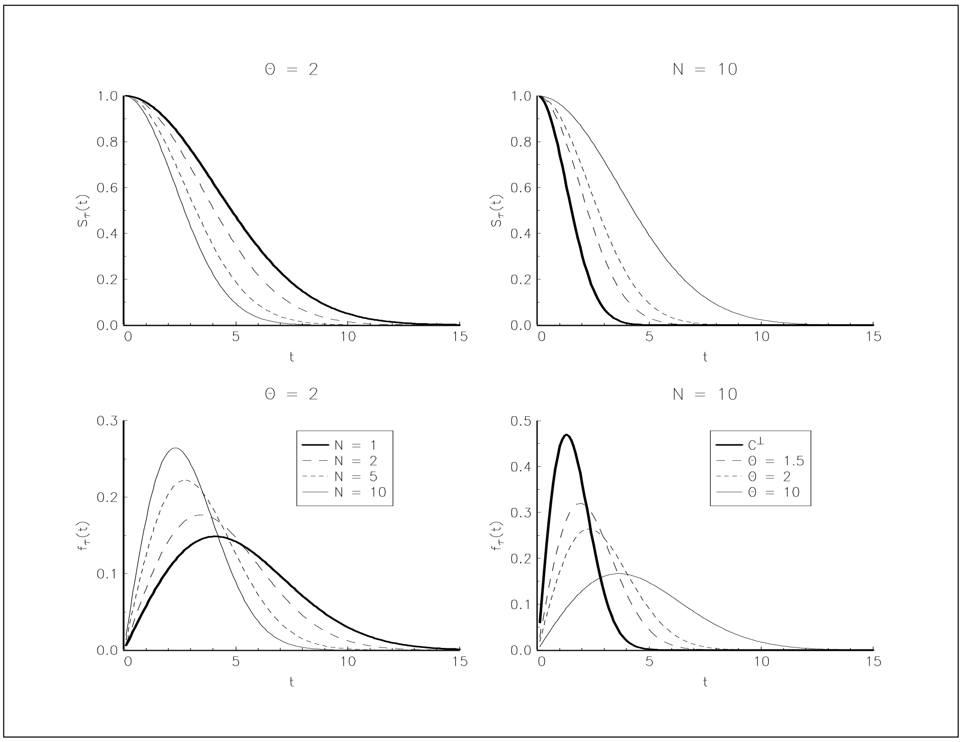

Figure 6: Failure time with Weibull $(0.03, 2)$ survival times

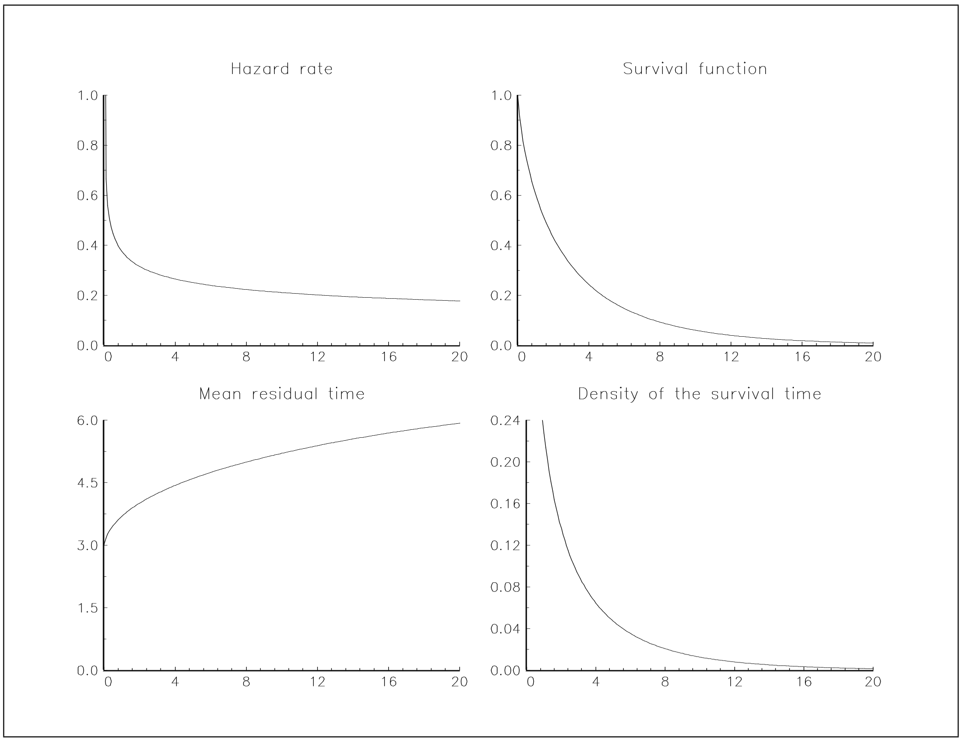

Figure 7: Weibull $(0.5, 0.75)$ survival time

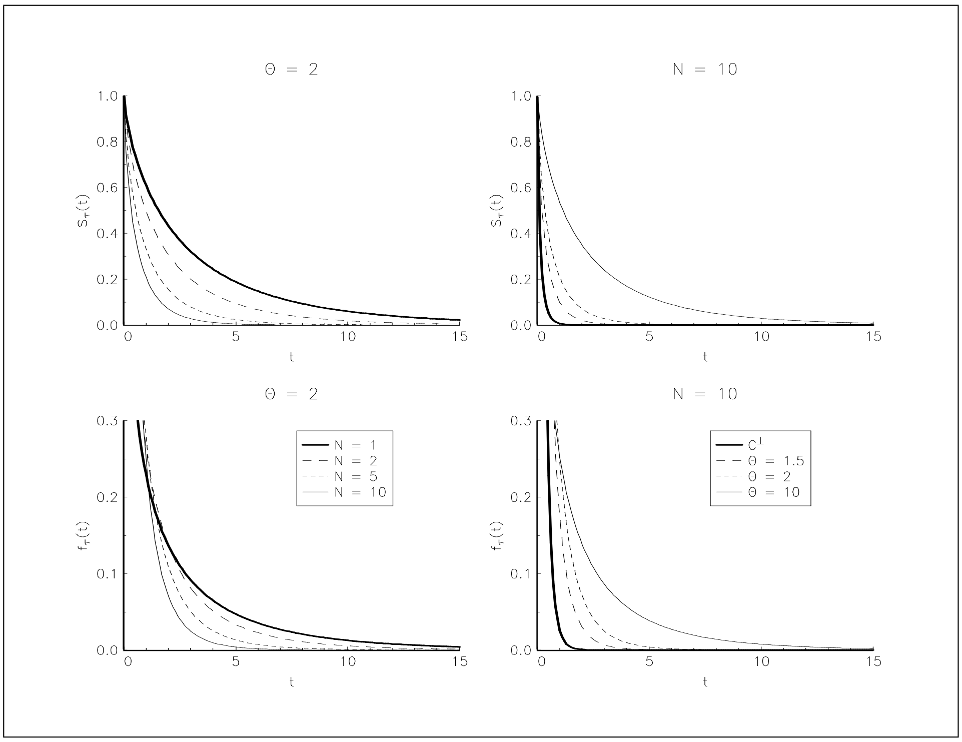

Figure 8: Failure time with Weibull $(0.5, 0.75)$ survival times

3.3 Order statistics

Theorem 23 *The survival function of the order statistic $T_{n:N}$ ($n \leq N$) is given by the following formula*

$$\mathbf{S}_{n:N}(t) = 1 - \mathbf{F}_{n:N}(t) \tag{174}$$

with

$$\mathbf{F}_{n:N}(t) = \sum_{k=n}^{N} \left[\sum_{l=n}^{k} (-1)^{k-l} \binom{k}{l} \sum_{\mathbf{v}(\mathbf{F}_1(t),\ldots,\mathbf{F}_N(t)) \in \mathcal{Z}(N-k,N,1)} \mathbf{C}(v_1,\ldots,v_N) \right] \tag{175}$$

Proof. see Appendix A.3. ■

We remark that the failure time τ corresponds to the order statistic $T_{1:N}$ and we verify that

$$\begin{aligned}
\mathbf{F}_{1:N}(t) &= \sum_{k=1}^{N} \left[\sum_{l=1}^{k} (-1)^{k-l} \binom{k}{l} \sum_{\mathbf{v}(\mathbf{F}_1(t),\ldots,\mathbf{F}_N(t)) \in \mathcal{Z}(N-k,N,1)} \mathbf{C}(v_1,\ldots,v_N) \right] \\
&= 1 - \left[1 + \sum_{k=1}^{N} \left[(-1)^k \sum_{\mathbf{v}(\mathbf{F}_1(t),\ldots,\mathbf{F}_N(t)) \in \mathcal{Z}(N-k,N,1)} \mathbf{C}(v_1,\ldots,v_N) \right] \right] \\
&= 1 - \breve{\mathbf{C}}(\mathbf{S}_1(t),\ldots,\mathbf{S}_N(t))
\end{aligned} \tag{176}$$

We also note that the last order statistic $T_{N:N}$ is the maximum of the survival times and its distribution is the diagonal section of the multivariate distribution:

$$\mathbf{F}_{N:N}(t) = \mathbf{C}(\mathbf{F}_1(t),\ldots,\mathbf{F}_N(t)) \tag{177}$$

Moreover, we retreive the well known results when the survival times are *i.i.d.*

$$\mathbf{F}_{N:N}(t) = \prod_{n=1}^{N} \mathbf{F}_n(t) = \mathbf{F}_1^N(t) \tag{178}$$

It is interesting to characterize the density function of $T_{n:N}$. We have

$$f_{n:N}(t) = \sum_{k=n}^{N} \left[\sum_{l=n}^{k} (-1)^{k-l} \binom{k}{l} \sum_{\mathbf{v}(\mathbf{F}_1(t),\ldots,\mathbf{F}_N(t)) \in \mathcal{Z}(N-k,N,1)} \left[\sum_{m=1}^{N} \mathbf{1}_{[v_m \neq 1]} \partial_m \mathbf{C}(v_1,\ldots,v_N) \times f_m(t) \right] \right] \tag{179}$$

We may also characterize other statistics which are relevant in reliability or life modelling. For example, we could be interested in the range $W = T_{N:N} - T_{1:N}$ or subranges $W_{n_1:n_2} = T_{n_2:N} - T_{n_1:N}$ ($n_1 < n_2$). However, to derive explicit formulas, we need the joint distribution of $T_{n_1:N}$ and $T_{n_2:N}$. In the case *i.i.d.*, Balakrishnan and Clifford Cohen [1991] give some tractable formula for the density. In the general case, the problem is open. One solution is then to use Monte Carlo methods. For the computation of $\mathbf{S}_{n:N}(t)$, Monte Carlo methods may be preferred too when N is large and n is small. We give in Table 3 an idea about the time needed to compute $\mathbf{S}_{n:N}(t)$ for $t \in \left\{-3 : \frac{8}{100} : 5\right\}$ with the exact formula in the case where the margins are gaussian and the copula is Gumbel-Hougaard. The computation has been done with the GAUSS software and a Pentium III 550 Mhz. Even if these times are given for indication, we remark that they dramatically increase with the dimension N.

In Figures 9 and 10, we have reported the survival function and the density function of some order statistics $T_{n:N}$ when the copula is Cook-Johnson[18] and the survival times are $\mathcal{LN}(0,1)$. We may compare the true

[18] The parameter θ is equal to 0.5 in Figure 9.

$N \diagdown n$	1	2	3	4	5	6	7	8	9	10	11	12	13
8	5	3	3	3	2	2	0	0					
9	9	9	9	6	6	3	2	2	0				
10	30	28	28	26	22	13	4	4	0	0			
11	98	98	97	94	84	61	32	11	3	1	0		
12	339	339	336	333	313	257	163	72	20	3	2	0	
13	1297	1297	1296	1290	1153	1125	840	449	156	34	5	2	0

Table 3: Computational time (in hundredths of a second) for the calculus of $\mathbf{S}_{n:N}(t)$

density of $T_{n:N}$ with the estimated density based on a Monte Carlo scheme. N_s represents the number of simulations. Figure 11 corresponds to the order statistic $T_{5:10}$ with $\theta = 1$. We remark that the estimated density[19] "converges" to the true one when the number of simulations increases. However, the rate of convergence depends on the dimension N and very high dimensions require a large number of simulations. In Figure 12, we have plotted the density of the range W.

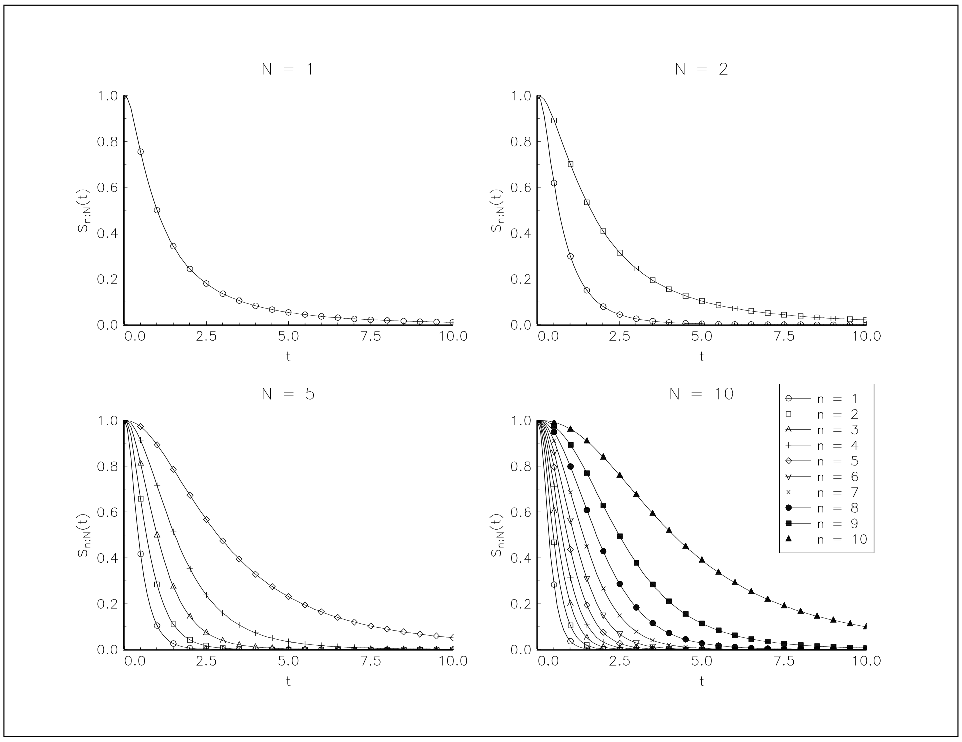

Figure 9: Survival function of order statistics $T_{n:N}$

4 Statistical inference

In this section, we present some methods to estimate multivariate survival functions. We do not provide an exhaustive review of the methods. For example, we do not consider non-parametric estimation (Hanley and Parnes [1983], Dabrowska [1988], Fermanian [1997]) or counting processes approach (Nielsen, Gill, Andersen and Sorensen [1992]).

[19] All the densities have been estimated using an Epanechnikov kernel with a left truncated point.

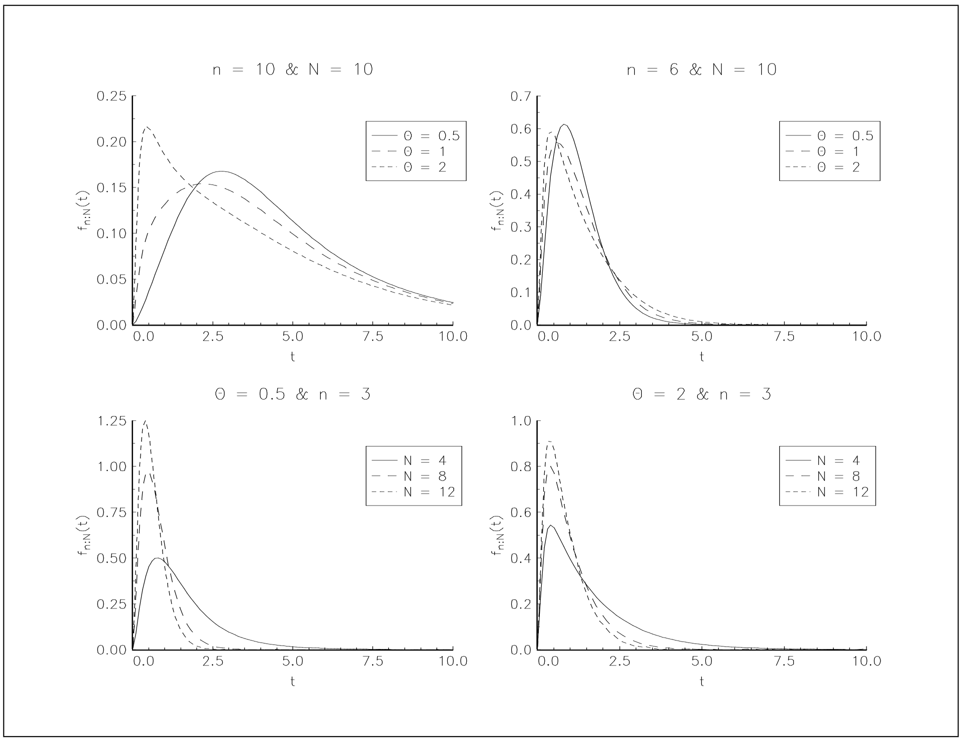

Figure 10: Density function of order statistics $T_{n:N}$

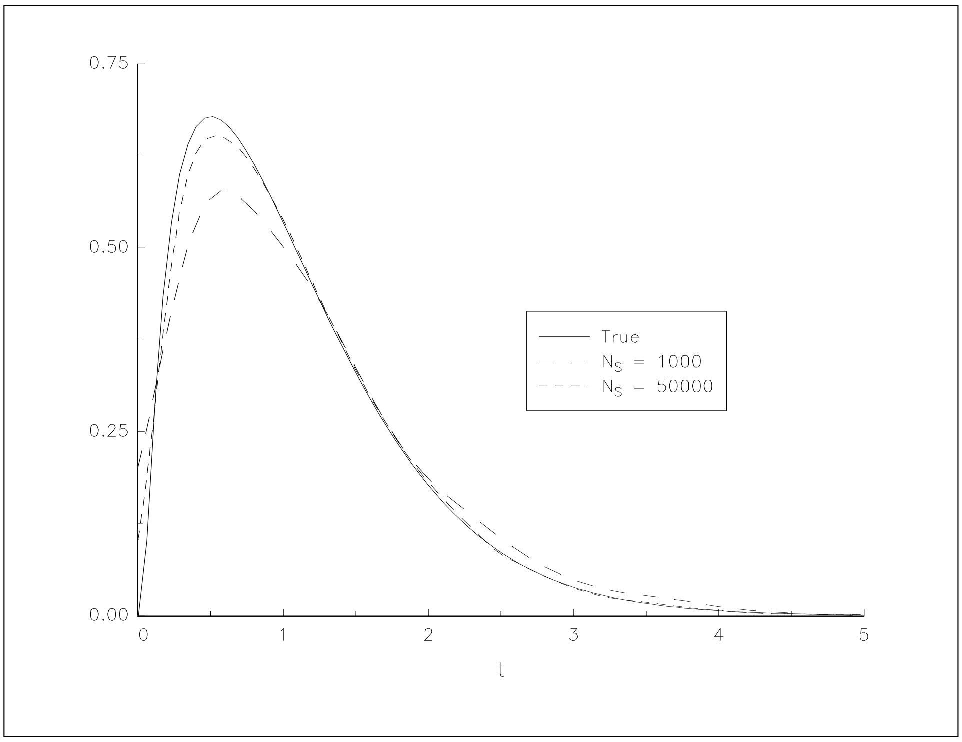

Figure 11: Density function of the order statistic $T_{5:10}$

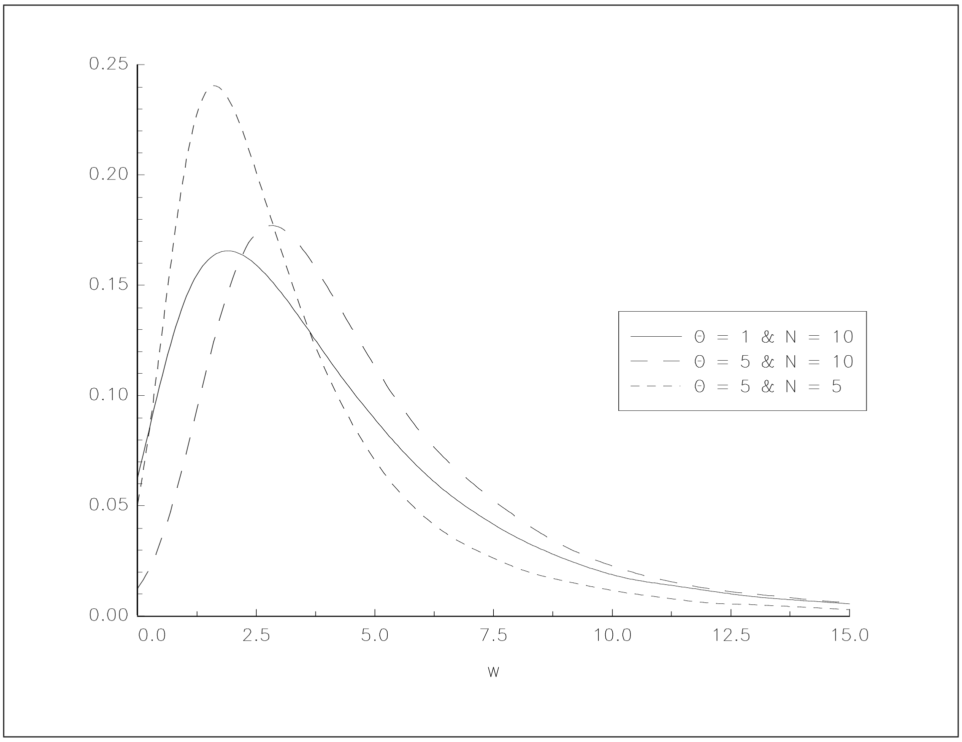

Figure 12: Density function of the range W

4.1 Maximum likelihood method

We consider the estimation problem of the vector of parameters $\boldsymbol{\theta}$ of the survival function $\mathbf{S}$. With a copula structure, we have

$$\mathbf{S}(t_1, t_2; \boldsymbol{\theta}) = \breve{\mathbf{C}}\left(\mathbf{S}_1\left(t_1; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(t_2; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right) \tag{180}$$

with $\boldsymbol{\theta} = \left(\boldsymbol{\theta}^1, \boldsymbol{\theta}^2, \boldsymbol{\theta}^{12}\right)$. In this case, $\boldsymbol{\theta}^1$ and $\boldsymbol{\theta}^2$ are the specific parameters of the univariate survival functions, whereas $\boldsymbol{\theta}^{12}$ is the parameter (possibly multidimensional) of the survival copula function.

Let $\mathbf{t} = \{(t_{1,i}, t_{2,i}), i = 1, \ldots, n\}$ denote a sample. The log-likelihood is

$$\ell(\mathbf{t}; \boldsymbol{\theta}) = \sum_{i=1}^{n} \ln f(t_{1,i}, t_{2,i}; \boldsymbol{\theta}) \tag{181}$$

The ML estimate (MLE) corresponds then to

$$\hat{\boldsymbol{\theta}}_{\mathsf{ML}} = \arg\max_{\boldsymbol{\theta}\in\Theta} \ell(\mathbf{t}; \boldsymbol{\theta}) \tag{182}$$

where Θ is the parameter space. However, dealing with survival times is not as simple, because records on survival traits are often incomplete: survival data are usually censored or/and truncated (Leung, Elashoff and Afifi [1997]). Censoring and truncation mechanims could be very difficult to take into account. In what follows, we consider the mechanims which could be generally found in finance.

We use the following notations:

- T is the survival time;
- C^- is the left-censoring time;
- C^+ is the right-censoring time;
- D is the observed time;

We observe the triplet $\left(D, \delta^-, \delta^+\right)$ with

$$D = \left(T \wedge C^+\right) \wedge \left(T \vee C^-\right) = \begin{cases} C^- & \text{if} \quad T \leq C^- \\ T & \text{if} \quad C^- < T \leq C^+ \\ C^+ & \text{if} \quad C^+ < T \end{cases} \tag{183}$$

and $\left(\delta^-, \delta^+\right) = \left(\mathbf{1}_{[T \leq C^-]}, \mathbf{1}_{[T > C^+]}\right)$. Moreover, we consider a left truncation variable Z. D is observed **only if** $T > Z$. Let $\delta^\star$ denote $\mathbf{1}_{[T>Z]}$. In what follows, bivariate survival data correspond to a sample of the form $\mathbf{y} = \left\{\mathbf{y}_i = \left(d_{1,i}, d_{2,i}, z_{1,i}, z_{2,i}, \delta^-_{1,i}, \delta^-_{2,i}, \delta^+_{1,i}, \delta^+_{2,i}, \delta^\star_{1,i}, \delta^\star_{2,i}\right), i = 1, \ldots, n\right\}$.

4.1.1 Parametric estimation

We assume that the censoring times are independent of the survival times and are not informative. Using results of Appendix C, we have then

$$\ell\left(\mathbf{y}; \boldsymbol{\theta}\right) = \sum_{i=1}^{n} \ell\left(\mathbf{y}_i; \boldsymbol{\theta}\right) \tag{184}$$

with

$$\begin{aligned}
\ell\left(\mathbf{y}_i; \boldsymbol{\theta}\right) \quad \propto \quad & \left(1 - \delta^-_{1,i}\right)\left(1 - \delta^-_{2,i}\right)\left(1 - \delta^+_{1,i}\right)\left(1 - \delta^+_{2,i}\right) \cdot \ln \breve{c}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right) + \\
& \left(1 - \delta^-_{1,i}\right)\left(1 - \delta^+_{1,i}\right)\left(1 - \delta^-_{2,i}\delta^+_{2,i}\right) \cdot \ln f_1\left(d_{1,i}; \boldsymbol{\theta}^1\right) + \\
& \left(1 - \delta^-_{2,i}\right)\left(1 - \delta^+_{2,i}\right)\left(1 - \delta^-_{1,i}\delta^+_{1,i}\right) \cdot \ln f_2\left(d_{2,i}; \boldsymbol{\theta}^2\right) + \\
& \delta^-_{1,i}\left(1 - \delta^-_{2,i}\right)\left(1 - \delta^+_{1,i}\right)\left(1 - \delta^+_{2,i}\right) \cdot \ln\left(1 - \partial_2 \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right)\right) + \\
& \left(1 - \delta^-_{1,i}\right)\delta^-_{2,i}\left(1 - \delta^+_{1,i}\right)\left(1 - \delta^+_{2,i}\right) \cdot \ln\left(1 - \partial_1 \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right)\right) + \\
& \left(1 - \delta^-_{1,i}\right)\left(1 - \delta^-_{2,i}\right)\delta^+_{1,i}\left(1 - \delta^+_{2,i}\right) \cdot \ln \partial_2 \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right) + \\
& \left(1 - \delta^-_{1,i}\right)\left(1 - \delta^-_{2,i}\right)\left(1 - \delta^+_{1,i}\right)\delta^+_{2,i} \cdot \ln \partial_1 \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right) + \\
& \delta^-_{1,i}\delta^-_{2,i}\left(1 - \delta^+_{1,i}\right)\left(1 - \delta^+_{2,i}\right) \cdot \ln\left(1 - \mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right) - \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right) + \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right)\right) + \\
& \delta^-_{1,i}\left(1 - \delta^-_{2,i}\right)\left(1 - \delta^+_{1,i}\right)\delta^+_{2,i} \cdot \ln\left(\mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right) - \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right)\right) + \\
& \left(1 - \delta^-_{1,i}\right)\delta^-_{2,i}\delta^+_{1,i}\left(1 - \delta^+_{2,i}\right) \cdot \ln\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right) - \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right)\right) + \\
& \left(1 - \delta^-_{1,i}\right)\left(1 - \delta^-_{2,i}\right)\delta^+_{1,i}\delta^+_{2,i} \cdot \ln \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(d_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right) - \\
& \delta^\star_{1,i}\delta^\star_{2,i} \cdot \ln \breve{\mathbf{C}}\left(\mathbf{S}_1\left(z_{1,i}; \boldsymbol{\theta}^1\right), \mathbf{S}_2\left(z_{2,i}; \boldsymbol{\theta}^2\right); \boldsymbol{\theta}^{12}\right) - \\
& \delta^\star_{1,i}\left(1 - \delta^\star_{2,i}\right) \cdot \ln \mathbf{S}_1\left(z_{1,i}; \boldsymbol{\theta}^1\right) - \\
& \left(1 - \delta^\star_{1,i}\right)\delta^\star_{2,i} \cdot \ln \mathbf{S}_2\left(z_{2,i}; \boldsymbol{\theta}^2\right)
\end{aligned} \tag{185}$$

MLE is defined as the solution of (182). If the log-likelihood function does attain an interior maximum, MLE is the solution of the score equation:

$$\frac{\partial}{\partial \boldsymbol{\theta}} \ell\left(\mathbf{y}; \boldsymbol{\theta}\right) = \mathbf{0} \tag{186}$$

Let $\boldsymbol{\theta}_0$ be the vector of true parameters. Under regularity conditions, we have

$$n^{1/2}\left(\hat{\boldsymbol{\theta}}_{\mathsf{ML}}-\boldsymbol{\theta}_0\right)\underset{\text{as}}{\sim}\mathcal{N}\left(0,\mathfrak{J}^{-1}\left(\boldsymbol{\theta}_0\right)\right) \tag{187}$$

where $\mathfrak{J}\left(\boldsymbol{\theta}_0\right)$ is the limit of the average information matrix.

In the previous approach (the full ML estimation), the parameters $\boldsymbol{\theta}$ are estimated simultaneously. Nevertheless, the copula approach suggests to perform the estimation in two steps (Shih and Louis [1995]):

1. we first estimate $\boldsymbol{\theta}^1$ and $\boldsymbol{\theta}^2$ separately by maximizing log-likelihoods $\ell\left(\mathbf{y}^1;\boldsymbol{\theta}^1\right)$ and $\ell\left(\mathbf{y}^2;\boldsymbol{\theta}^1\right)$ of the univariate survival data;
2. then, we estimate $\boldsymbol{\theta}^{12}$ by maximizing $\ell\left(\mathbf{y};\hat{\boldsymbol{\theta}}^1,\hat{\boldsymbol{\theta}}^2,\boldsymbol{\theta}^{12}\right)$ given the previous estimates $\hat{\boldsymbol{\theta}}^1$ and $\hat{\boldsymbol{\theta}}^2$.

This estimation method is called by Shih and Louis the *two-stage parametric ML method*. In Joe and Xu [1996], it is denoted the estimation method of *inference functions for margins* (IFM). The estimates $\hat{\boldsymbol{\theta}}^1_{\mathsf{IFM}}$ and $\hat{\boldsymbol{\theta}}^2_{\mathsf{IFM}}$ are then given by

$$\begin{aligned}\hat{\boldsymbol{\theta}}^j_{\mathsf{IFM}} \quad = \quad & \arg\max\sum_{i=1}^n\left(1-\delta^-_{j,i}\right)\left(1-\delta^+_{j,i}\right)\ln f_j\left(d_{j,i};\boldsymbol{\theta}^j\right)+\sum_{i=1}^n\delta^-_{j,i}\left(1-\delta^+_{j,i}\right)\ln\left(1-\mathbf{S}_j\left(d_{j,i};\boldsymbol{\theta}^j\right)\right)+ \\ & \sum_{i=1}^n\left(1-\delta^-_{j,i}\right)\delta^+_{j,i}\ln\mathbf{S}_j\left(d_{j,i};\boldsymbol{\theta}^j\right)-\sum_{i=1}^n\delta^\star_{j,i}\ln\mathbf{S}_j\left(z_{j,i};\boldsymbol{\theta}^j\right)\end{aligned} \tag{188}$$

for $j=1,2$. It comes that

$$\hat{\boldsymbol{\theta}}^{12}_{\mathsf{IFM}}=\arg\max\ell\left(\mathbf{y};\hat{\boldsymbol{\theta}}^1_{\mathsf{IFM}},\hat{\boldsymbol{\theta}}^2_{\mathsf{IFM}},\boldsymbol{\theta}^{12}\right) \tag{189}$$

"*This procedure is computationally simpler than estimating all parameters simultaneously. A numerical optimization with many parameters is much more time-consuming compared with several numerical optimizations, each with fewer parameters*" (Joe [1987], page 300). The main advantage of IFM is then the simplification of numerical computations. Nevertheless, the IFM estimator has other desirable properties. Like the MLE, it is asymptotically efficient (Joe and Xu [1996], Joe [1997]) and we have

$$n^{1/2}\left(\hat{\boldsymbol{\theta}}_{\mathsf{IFM}}-\boldsymbol{\theta}_0\right)\underset{\text{as}}{\sim}\mathcal{N}\left(0,\mathfrak{G}^{-1}\left(\boldsymbol{\theta}_0\right)\right) \tag{190}$$

with $\mathfrak{G}\left(\boldsymbol{\theta}_0\right)$ the **information matrix** of Godambe. Let us define the score function $\mathbf{g}\left(\mathbf{y};\boldsymbol{\theta}\right)$ in the following way

$$\mathbf{g}\left(\mathbf{y};\boldsymbol{\theta}\right)=\left[\begin{array}{c}\partial\ell\left(\mathbf{y}^1;\boldsymbol{\theta}^1\right)/\partial\boldsymbol{\theta}^1 \\ \partial\ell\left(\mathbf{y}^2;\boldsymbol{\theta}^2\right)/\partial\boldsymbol{\theta}^1 \\ \partial\ell\left(\mathbf{y};\boldsymbol{\theta}^1,\boldsymbol{\theta}^2,\boldsymbol{\theta}^{12}\right)/\partial\boldsymbol{\theta}^{12}\end{array}\right] \tag{191}$$

The Godambe information matrix takes the form (Joe [1997]):

$$\mathfrak{G}\left(\boldsymbol{\theta}_0\right)=\mathbf{D}^{-1}\mathbf{M}\left(\mathbf{D}^{-1}\right)^\top \tag{192}$$

where $\mathbf{D}=\mathbb{E}\left[\partial\mathbf{g}\left(\mathbf{y};\boldsymbol{\theta}\right)^\top/\partial\boldsymbol{\theta}\right]$ and $\mathbf{M}=\mathbb{E}\left[\mathbf{g}\left(\mathbf{y};\boldsymbol{\theta}\right)^\top\mathbf{g}\left(\mathbf{y};\boldsymbol{\theta}\right)\right]$.

4.1.2 Semi-parametric estimation

In this paragraph, we assume that there exist non-parametric estimators of the survival functions denoted by $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$. In this case, the log-likelihood becomes

$$\ell\left(\mathbf{y};\boldsymbol{\theta}^{12}\right)=\sum_{i=1}^{n}\ell\left(\mathbf{y}_i;\boldsymbol{\theta}^{12}\right) \tag{193}$$

with

$$\begin{aligned}
\ell\left(\mathbf{y}_i;\boldsymbol{\theta}^{12}\right) \quad\propto\quad & \left(1-\delta_{1,i}^{-}\right)\left(1-\delta_{2,i}^{-}\right)\left(1-\delta_{1,i}^{+}\right)\left(1-\delta_{2,i}^{+}\right)\cdot\ln\breve{c}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)+\\
& \delta_{1,i}^{-}\left(1-\delta_{2,i}^{-}\right)\left(1-\delta_{1,i}^{+}\right)\left(1-\delta_{2,i}^{+}\right)\cdot\ln\left(1-\partial_2\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)\right)+\\
& \left(1-\delta_{1,i}^{-}\right)\delta_{2,i}^{-}\left(1-\delta_{1,i}^{+}\right)\left(1-\delta_{2,i}^{+}\right)\cdot\ln\left(1-\partial_1\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)\right)+\\
& \left(1-\delta_{1,i}^{-}\right)\left(1-\delta_{2,i}^{-}\right)\delta_{1,i}^{+}\left(1-\delta_{2,i}^{+}\right)\cdot\ln\partial_2\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)+\\
& \left(1-\delta_{1,i}^{-}\right)\left(1-\delta_{2,i}^{-}\right)\left(1-\delta_{1,i}^{+}\right)\delta_{2,i}^{+}\cdot\ln\partial_1\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)+\\
& \delta_{1,i}^{-}\delta_{2,i}^{-}\left(1-\delta_{1,i}^{+}\right)\left(1-\delta_{2,i}^{+}\right)\cdot\ln\left(1-\hat{\mathbf{S}}_1\left(d_{1,i}\right)-\hat{\mathbf{S}}_2\left(d_{2,i}\right)+\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)\right)+\\
& \delta_{1,i}^{-}\left(1-\delta_{2,i}^{-}\right)\left(1-\delta_{1,i}^{+}\right)\delta_{2,i}^{+}\cdot\ln\left(\hat{\mathbf{S}}_2\left(d_{2,i}\right)-\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)\right)+\\
& \left(1-\delta_{1,i}^{-}\right)\delta_{2,i}^{-}\delta_{1,i}^{+}\left(1-\delta_{2,i}^{+}\right)\cdot\ln\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right)-\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)\right)+\\
& \left(1-\delta_{1,i}^{-}\right)\left(1-\delta_{2,i}^{-}\right)\delta_{1,i}^{+}\delta_{2,i}^{+}\cdot\ln\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(d_{1,i}\right),\hat{\mathbf{S}}_2\left(d_{2,i}\right);\boldsymbol{\theta}^{12}\right)-\\
& \delta_{1,i}^{\star}\delta_{2,i}^{\star}\cdot\ln\breve{\mathbf{C}}\left(\hat{\mathbf{S}}_1\left(z_{1,i}\right),\hat{\mathbf{S}}_2\left(z_{2,i}\right);\boldsymbol{\theta}^{12}\right)
\end{aligned} \tag{194}$$

$\boldsymbol{\theta}^{12}$ is then estimated by maximizing the log-likelihood $\ell\left(\mathbf{y};\boldsymbol{\theta}^{12}\right)$. It is called the *omnibus* (om) estimator by Genest and Werker [2002]. This estimation method has been first suggested[20] by Genest, Ghoudi and Rivest [1995] and Shih and Louis [1995], who both show that this semi-parametric estimator is consistent, asymptotically normal and fully efficient at independence. Klassen and Wellner [1997] extend the efficiency property in the case of the Normal copula.

The om estimation requires to use the non-parametric estimators $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$. Even if they are both rank-based and consistent[21], Genest and Werker [2002] suggest that "*it is quite improbable that the omnibus* estimator $\hat{\boldsymbol{\theta}}_{\mathsf{om}}^{12}$ is semi-parametrically efficient for any of the most common parametric copula models". However, this method gives better results that parametric methods in the case of mispecifications of the margins. Durrleman, Nikeghbali and Roncalli [2000a] suggest then to use the om estimator $\hat{\boldsymbol{\theta}}_{\mathsf{om}}^{12}$ to check that the MLE or IFM estimator is not biased. They give an example where both MLE and IFM estimators are biased. We consider now a Monte Carlo study where the survival times are two Weibull random variables — $T_1\sim\mathsf{Weibull}\left(0.75,0.10\right)$ and $T_2\sim\mathsf{Weibull}\left(0.75,0.10\right)$ — and the survival copula is a Normal copula with parameter ρ equal to 0.50. The size n of the sample is set to 200 and the number of replications is 500. To estimate the parameters, we assume that the survival times are exponential and the survival copula is Normal. In Figure 13, we have reported the pdf of the different estimators for the parameter ρ. For this particular example, IFM and om give better results than ML.

[20] But we find the same idea in the works of David Oakes (see for example Oakes [1994]) and Philip Hougaard (see for example Hougaard, Harvald and Holm [1992]).

[21] For example, we can use the Kaplan–Meier or Nelson–Aalen estimators.

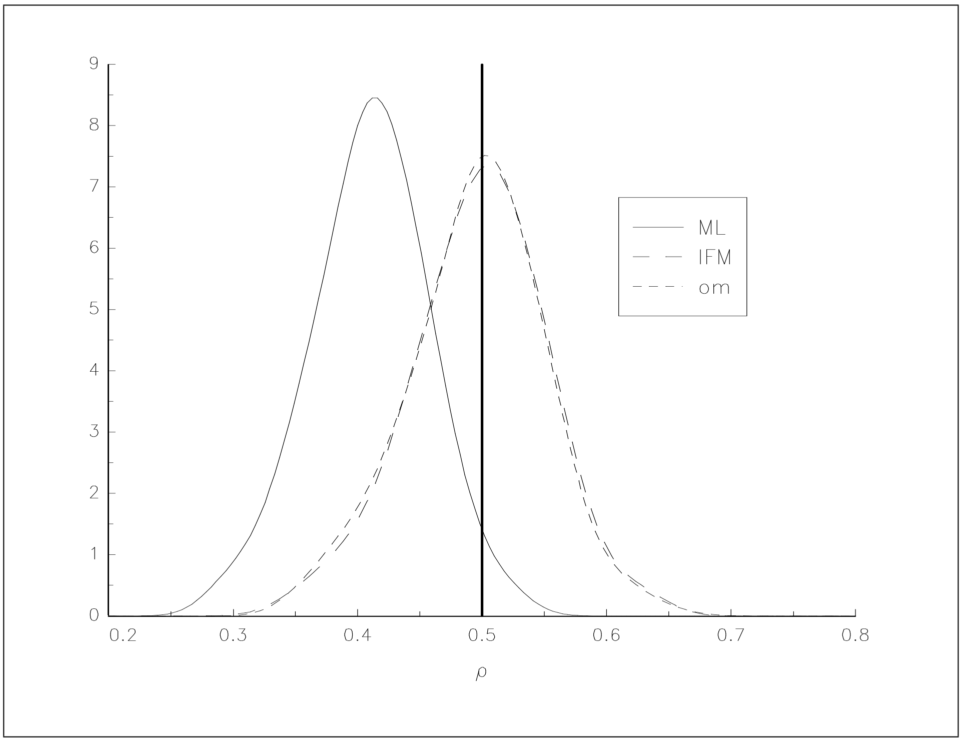

Figure 13: Probability density function of $\hat{\rho}_{\mathsf{MLE}}$, $\hat{\rho}_{\mathsf{IFM}}$ and $\hat{\rho}_{\mathsf{om}}$

4.2 EM algorithm

The EM algorithm is a general method introduced by to obtain the MLE in the case of incomplete data[22] (Dempster, Laird and Rubin [1977]). Suppose that the data consists in observed (but incomplete) data Y and missing (or latent) data Z. In terms of density, we can write

$$f(\boldsymbol{\theta} \mid Y) = \frac{f(\boldsymbol{\theta} \mid Y, Z) \cdot f(Z \mid Y)}{f(Z \mid \boldsymbol{\theta}, Y)} \tag{195}$$

The EM algorithm consists then in two steps:

1. E-step
   Find the expected value of the posterior density function:
   $$Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)}\right) = \int \ln f(\boldsymbol{\theta} \mid Y, Z) \, f\left(Z \mid Y, \boldsymbol{\theta}^{(i)}\right) \mathrm{d}Z \tag{196}$$
2. M-step
   Maximize the conditional posterior density function:
   $$\boldsymbol{\theta}^{(i+1)} = \arg\max_{\boldsymbol{\theta} \in \Theta} Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)}\right) \tag{197}$$

These two steps are repeated until the algorithm converges — we note $\hat{\boldsymbol{\theta}}_{\mathsf{EM}}$ the estimate.

This algorithm is well adapted for frailty models. We recall that the two survival times are independent given the frailty W. It comes that the posterior density function takes a very tractable expression (Guo and Rodríguez [1992]).

[22] see Tanner [1996] for a good exposition.

4.3 Kendall's tau estimator

The estimation problem of the Clayton model has been investigated by many people (for example Oakes [1982], Clayton and Cuzik [1985], Genest and Rivest [1993], Maguluri [1993], Glidden and Self [1999] and Fine and Jiang [2000]). Among the different approaches, the concordance estimator proposed by Oakes [1982] is the most simple. The idea is to estimate the parameter of the copula thanks to the method of moments and a moment condition based on Kendall's tau. We have

$$\hat{\theta} = \arg \{\tau(\theta) = \hat{\tau}\} \tag{198}$$

with $\tau(\theta) = \tau \left\langle \breve{\mathbf{C}}_{(\theta)} \right\rangle$ and $\hat{\tau}$ the value computed for the sample[23]. Phelps and Weissfeld [1997] show that this estimator "*performs well when the data are weakly dependent but does not perform well when the data are highly dependent or independent*". Nevertheless, the Oakes method is certainly the only method that could be implemented in finance for high dimensional problems. Let $\breve{\mathbf{C}}(\mathbf{u};\boldsymbol{\theta})$ a survival copula of dimension N. Suppose that all bivariate marginal copulas are in a given family with only one parameter: $\breve{\mathbf{C}}(\mathbf{u};\boldsymbol{\theta}) = \breve{\mathbf{C}}_{i,j}(u_i, u_j; \theta_{i,j})$ if $u_n = 1$ for all n except i and j. Suppose moreover that $\boldsymbol{\theta}$ is exactly the set of the parameters $\theta_{i,j}$ of the bivariate copulas. In this case, we have a one-to-one correspondence between $\theta_{i,j}$ and $\tau(T_i, T_j)$. We can then estimate the full vector of the parameters $\boldsymbol{\theta}$ in a similar way as the Oakes method:

$$\hat{\theta}_{i,j} = \arg \{\tau(T_i, T_j) = \hat{\tau}_{i,j}\} \tag{199}$$

This method is particular adapted for the Normal copula[24].

5 Financial applications

In this section, we consider different applications in finance[25]. For all of them, the survival copula is assumed to be the Normal copula. We do not discuss the appropriateness of this assumption, because as remarked by Bouyé, Durrleman, Nikeghbali, Riboulet and Roncalli [2000b], the Normal copula is almost the only 'industrial' copula function for finance.

Initially, we did the project to include a fourth financial application, which concerns the Life Time Value (LTV). LTV is an economic measure of the (potential) worth of a customer. One of the main point of LTV is to determine the length of time T a customer will remain active. Following Georges, Jacomy and Lazare [2000], LTV of the individual n corresponds to

$$\mathsf{LTV}_n(t) = \sum_k P_n(t, t_k) R_n(t_k) \mathbf{S}_n(t_k \mid t) \tag{200}$$

where P is the discount factor, R is the (potential) rentability function and $\mathbf{S}$ is the survival function of T. Using a continuous-time framework, we have

$$\mathsf{LTV}_n(t) = \int P_n(t, t+\delta) R_n(t+\delta) \mathbf{S}_n(t+\delta \mid t) \, \mathrm{d}\delta \tag{201}$$

LTV is a powerful tool for the bank in terms of segmentation and customer relationship management (CRM). For example, the bank could use LTV in order to decide how much to spend on keeping customers. Another application concerns mortgage: the bank could link the contract rate on the loan with the LTV of the customer. In general, given economic factors, LTV are assumed to be independent. However, this is not always the case. Thus, we could think that LTV of parents and children or husband and wife are 'correlated'. In this case, we could view the relationship between the bank and customers as stopping times. If we assume that the stopping times of the different members of a same family are independent, CRM will not be optimal.

[23] see Shih [1998] for the determination of $\hat{\tau}$ in presence of censoring.

[24] For the Normal copula, Durrleman, Nikeghbali and Roncalli [2000a] suggest to use the IFM or om estimator, which has an analytical expression.

[25] Contrary to other fields, for example actuarial sciences (see for instance Carriere [1994,2000], Frees, Carriere and Valdez [1996], Frees and Valdez [1998] and Valdez [2000]), the use of survival copulas is relatively new in finance.

5.1 Default, prepayment and credit lifetime

Modelling mortgage terminations is essential for the bank to value correctly the market price of the loan. For example, the price of the loan must incorporate the risk of default. Another source of termination is prepayment (Schwartz and Torus [1989], Frachot and Gourieroux [1995]). These two risks could be viewed as mortgage options or hidden options (Demey, Frachot and Riboulet [2000]). In general, these risks are assumed to be independent. However, this is not a satisfactory assumption (Schwartz and Torus [1992]).

We consider that the credit lifetime is linked to these events:

- either the counterparty defaults (thus it is the time-until-default $T_D^\star$),
- or a prepayment occurs (it is then the time-until-prepayment $T_P^\star$),
- or the credit goes to the maturity (we note it $T_M^\star$).

We assume that the time is measured in months. In order to make comparison, we standardize the survival times. Let T_D and T_P be defined as $T_D^\star / T_M^\star$ and $T_P^\star / T_M^\star$. The credit lifetime corresponds then to

$$T = \min\left(T_D, T_P, 1\right) \tag{202}$$

Let $\mathrm{d}t$ denote the standardized time unit, which is equal to $1/T_M^\star$. We can then consider the three following cases:

1. Default termination
   We have
   $$\Pr\left\{T_D = t, T_P > t \mid T > t - \mathrm{d}t\right\} \cdot \Pr\left\{T_D > t - \mathrm{d}t, T_P > t - \mathrm{d}t\right\} = \breve{\mathbf{C}}\left(\mathbf{S}_D\left(t - \mathrm{d}t\right), \mathbf{S}_P\left(t\right)\right) - \breve{\mathbf{C}}\left(\mathbf{S}_D\left(t\right), \mathbf{S}_P\left(t\right)\right) \tag{203}$$
2. Prepayment termination
   The probability of this event is
   $$\Pr\left\{T_D > t, T_P = t \mid T > t - \mathrm{d}t\right\} \cdot \Pr\left\{T_D > t - \mathrm{d}t, T_P > t - \mathrm{d}t\right\} = \breve{\mathbf{C}}\left(\mathbf{S}_D\left(t\right), \mathbf{S}_P\left(t - \mathrm{d}t\right)\right) - \breve{\mathbf{C}}\left(\mathbf{S}_D\left(t\right), \mathbf{S}_P\left(t\right)\right) \tag{204}$$
3. No default or prepayment termination
   In this case, the credit goes to maturity, or there has been no default nor prepayment and the credit isn't over yet ("real" censorship). It comes that
   $$\Pr\left\{T_D > t, T_P > t \mid T > t - \mathrm{d}t\right\} \cdot \Pr\left\{T_D > t - \mathrm{d}t, T_P > t - \mathrm{d}t\right\} = \breve{\mathbf{C}}\left(\mathbf{S}_D\left(t\right), \mathbf{S}_P\left(t\right)\right) \tag{205}$$

We note that we implicitely assume that when the credit lifetime and the maturity are the same, both T_D and T_P are "censored" and when prepayment (respectively default) occurs, T_D (respectively T_P) is censored.

We consider a database of personal loans starting from 1996 with 1050947 credits. For each credit, we have the maturity (or the theoretical credit lifetime), the credit lifetime and the cause of termination (default, prepayment or other). We would like to estimate the dependence between default and prepayment. More precisely, we would like to show that the two survival times are not independent. We do the following hypothesis: the margins $\mathbf{S}_D$ and $\mathbf{S}_P$ are exponential with parameter λ_D and λ_P, and the survival copula $\breve{\mathbf{C}}$ is a Normal copula with parameter ρ. Let t_i be the lifetime of credit i. The individual log-likelihood is then

$$\begin{aligned}
\ell\left(t_i, \delta_{D,i}, \delta_{P,i}; \boldsymbol{\theta}\right) &= \delta_{D,i} \ln\left(\breve{\mathbf{C}}\left(e^{-\lambda_D (t_i - \mathrm{d}t_i)}, e^{-\lambda_P t_i}; \rho\right) - \breve{\mathbf{C}}\left(e^{-\lambda_D t_i}, e^{-\lambda_P t_i}; \rho\right)\right) + \\
&\quad \delta_{P,i} \ln\left(\breve{\mathbf{C}}\left(e^{-\lambda_D t_i}, e^{-\lambda_P (t_i - \mathrm{d}t_i)}; \rho\right) - \breve{\mathbf{C}}\left(e^{-\lambda_D t_i}, e^{-\lambda_P t_i}; \rho\right)\right) + \\
&\quad \left(1 - \delta_{D,i} - \delta_{P,i}\right) \ln \breve{\mathbf{C}}\left(e^{-\lambda_D t_i}, e^{-\lambda_P t_i}; \rho\right)
\end{aligned} \tag{206}$$

with $\delta_{D,i}$ (respectively $\delta_{P,i}$) the indicator function that a default (respectively a prepayment) has occured. The vector $\boldsymbol{\theta}$ of parameters is of course $(\lambda_D, \lambda_P, \rho)$. We could estimate $\boldsymbol{\theta}$ by the ML method or by the IFM method. In this last case, we estimate the hazard rates with the following estimators — $\hat{\lambda}_D = \left(\sum_{i=1}^{n} t_i\right)^{-1} \sum_{i=1}^{n} \delta_{D,i}$ and $\hat{\lambda}_P = \left(\sum_{i=1}^{n} t_i\right)^{-1} \sum_{i=1}^{n} \delta_{P,i}$. We do the estimation for different years from 1996 to 1999 and three categories of credit (less than 3 years, between 3 and 5 years and from 5 to 10 years). We do not report the full results here, but the principal conclusions are the following:

- For all years and the three categories, ρ is negative (from -96% to -43%).
- There is no significant difference between the three categories.
- Censorship affects results in a decreasing way: thus, when credits are longer, or have started very recently, it is very unlikely to observe any default or prepayment. This is why the measure of dependence ρ is strongly negative when there is little censorship, less negative as censorship occurs more frequently.

Of course, these results are biaised because the model is not realistic. In particular, we do not take account interest rates movements, "Burn out" effect, etc. (Aïssaoui and Frachot [1999]). Moreover, it is difficult to assume that the hazard rate is constant (Baud and Trang [1999]). However, we believe that using a more realistic model does not change the main conclusion: **default and prepayment are negatively dependent**.

5.2 Measuring the risk of credit portfolios

One of the main issue concerning credit risk is without doubt the modelling of joint default distribution. Li [2000a] and Maccarinelli and Maggiolini [2000] suggest that copulas could be a suitable tool for such a problem. Indeed, a default is generally described by a *survival* function $\mathbf{S}(t) = \Pr\{T > t\}$, which indicates the probability that a security will attain age t. The survival time T is called the *time-until-default* in Li [2000a].

We consider here the problem of the risk measure of a credit portfolio and the economic capital allocation:

> *The estimated economic capital needed to support a bank's credit risk exposure is generally referred to as its required economic capital for credit risk. The process for determining this amount is analogous to value at risk methods used in allocating economic capital against market risks. Specifically, the economic capital for credit risk is determined so that the estimated probability of unexpected credit loss exhausting economic capital is less than some target insolvency rate.*
>
> *Capital allocation systems generally assume that it is the role of reserving policies to cover expected credit losses, while it is that of economic capital to cover unexpected credit losses. Thus, required economic capital is the additional amount of capital necessary to achieve the target insolvency rate, over and above that needed for coverage of expected loss (document [1], page 14).*

Let $\mathfrak{L}$ denote the credit loss random variable with distribution $\mathbf{F}_{\mathfrak{L}}$. The expected loss is then equal to

$$\mathrm{EL} = \mathbb{E}\left[\mathfrak{L}(t_0)\right] \tag{207}$$

whereas the unexpected loss is given by

$$\mathrm{UL} = \mathrm{TL} - \mathrm{EL} \tag{208}$$

TL is the target loss. Following business pratices, TL is most of the times a credit loss quantile at a specified confidence level α:

$$\mathrm{TL}(\alpha) = \mathbf{F}_{\mathfrak{L}}^{-1}(\alpha) \tag{209}$$

In this case, $1-\alpha$ could be interpreted as the "target insolvency rate" and depends of the rating target of the bank. We give here the values of α which are generally used:

Rating target	BBB	A	AA	AAA
α	99.75%	99.9%	99.95%	99.97%

We finally deduce that the *Capital-at-Risk* CaR (or economic capital EC) of a credit portfolio is

$$\mathrm{CaR}(\alpha) = \mathrm{UL}(\alpha) = \mathbf{F}_{\mathfrak{L}}^{-1}(\alpha) - \mathbb{E}\left[\mathfrak{L}(t_0)\right] \tag{210}$$

To compute this Capital-at-Risk, the bank may use an internal model or a benchmak model (CreditMetrics, CreditRisk+, KMV, etc.). These benchmark model have been extensively studied (Koyluoglu and Hickman [1998], Gordy [2000], Crouhy, Galai and Mark [2000]) and could produce different risk measures for the same portfolio (Crouhy [1999]). We could explain this result by the fact that these models use different specifications about the distribution of the survival times and dependence functions. For example, Li [2000a] shows that CreditMetrics uses implicitely a Normal copula. In CreditRisk+, the dependence function is related to a frailty model (Coutant, Martineu, Messines, Riboulet and Roncalli [2001]). Another main difference comes from the definition of a credit event. The Basel Commitee on Banking Supervision provides two definitions (document [1]):

- In the default mode (DM) paradigm, "*a credit loss arises only if a borrower defaults within the planning horizon*".
- In the mark-to-market (MtM) paradigm, "*a credit loss can arise in response to detoriation in an asset's credit quality short of default*".

In our point of view, CreditMetrics is a MtM model whereas CreditRisk+ is a DM model.

In this work, we focus on the DM paradigm (see Coutant, Martineu, Messines, Riboulet and Roncalli [2001] for the MtM paradigm) and show how it is related to survival modelling. We consider a very simple model where the actualized **potential** loss $\mathfrak{L}$ of the credit portfolio is defined by

$$\mathfrak{L}(t_0) = \sum_{n=1}^{N} P_n(t_0, t_n) \cdot B_n(t_n) \cdot (1 - R_n(t_n)) \cdot \mathbf{1}_{[T_n \leq t_n]} \tag{211}$$

where $P_n(t_0, t_n)$ is the discount factor for the credit line n, $B_n(t_n)$ the recovery basis, $R_n(t_n)$ the recovery rate and t_n the maturity of the credit. Without any loss of generality, we assume that $P_n(t_0, t_n)$ is equal to one. We note that the previous expression of $\mathfrak{L}$ could be written as

$$\mathfrak{L}(t_0) = \sum_{n=1}^{N} \mathrm{EAD}_n(t_n) \cdot \mathrm{LGD}_n(t_n) \cdot \mathrm{PD}_n(t_n) \tag{212}$$

where EAD is the 'exposure at default', LGD is the 'loss given default' and PD is the 'probability of default'. We then obtain a model similar to the IRB approach defined by the Basel Commitee on Banking Supervision (document [2]). However, we remark that we do not use a time frame of one year. We suppose that the time horizon for monitoring the risk of the credit line n is its maturity. This is sometimes the method used in banks, because credits could not be traded (see [1] page 17).

Another difference comes from the dependence between the probabilities of default. In the IRB approach, they are assumed independent (before taking into account the *granularity* adjustment). However, the dependence function (or "correlated defaults") may be a key point to compute the economic capital as suggested by numerical results of Duffie and Singleton [1999] and Lindskog and McNeil [2001]. This is particular true in the point of view of credit portfolio managers. Let us illustrate this point. We assume that $B_n(t_n)$ is known,

$R_n(t_n)$ is a uniform distribution $\mathcal{U}_{[a_n,b_n]}$ and T_n is an absolutely continuous survival time. In this case, the distribution $\mathbf{F}_{\mathfrak{L}}$ of the loss has a singular component in zero

$$\begin{aligned} \mathbf{F}_{\mathfrak{L}}(0) &= \Pr\{\mathfrak{L}(t_0)=0\} \\ &= \mathbf{S}(t_1,\ldots,t_N) \end{aligned} \tag{213}$$

and an absolutely continuous component in $]0,\infty]$. The portfolio is composed of two credit lines, which have the same characteristics[26]. We assume that the survival times are exponential with $\lambda_n = 0.10$ and the survival copula is Normal with a matrix $\boldsymbol{\rho}$ of parameters of the following form:

$$\boldsymbol{\rho} = \begin{bmatrix} 1 & \rho \\ & 1 \end{bmatrix}$$

In Figure 14, we have represented the density of the loss in the case of the product copula $\mathbf{C}^{\perp}$. We note that the dependence function influences the value of economic capital, because of the target quantile computation. Suppose that the bank imposes a risk limitation $\mathbf{L}$ to the credit portfolio manager. In Figure 15, we have represented the probability $\Pr\{\mathfrak{L}(t_0) \geq \mathbf{L}\}$ for different values of the coefficient ρ. If we set $\mathbf{L}$ to 100, we obtain a probability of 26.9%, 20.3% and 14.5% when ρ takes respectively the values -0.75, 0 and 0.75. Figure 16 gives the risk measure of this credit portfolio for three target solvency rates. In this case, we remark that differences induced by ρ depends on the confidence level α.

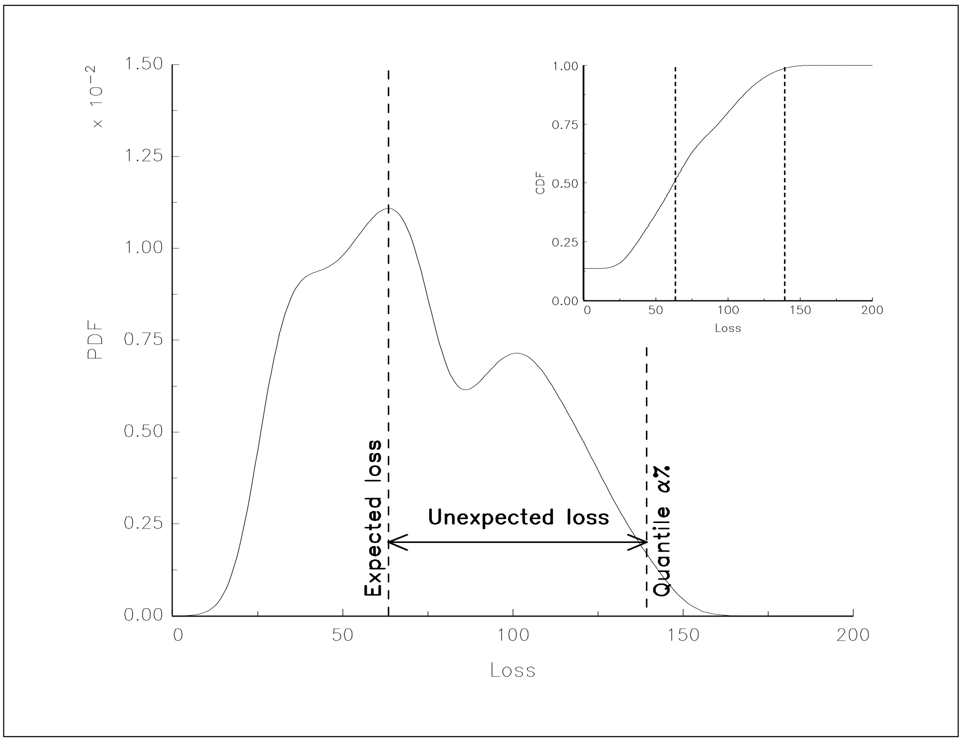

Figure 14: Density of the loss and allocated economic capital

We suppose now that the credit manager would add a third credit line in the portfolio. The 'correlation' between the two previous default times is set to 50%. We write also the matrix $\boldsymbol{\rho}$ of the Normal copula parameters

[26] $B = 100$, $t_n = 10$ years and $[a_n, b_n] = [25\%, 75\%]$.

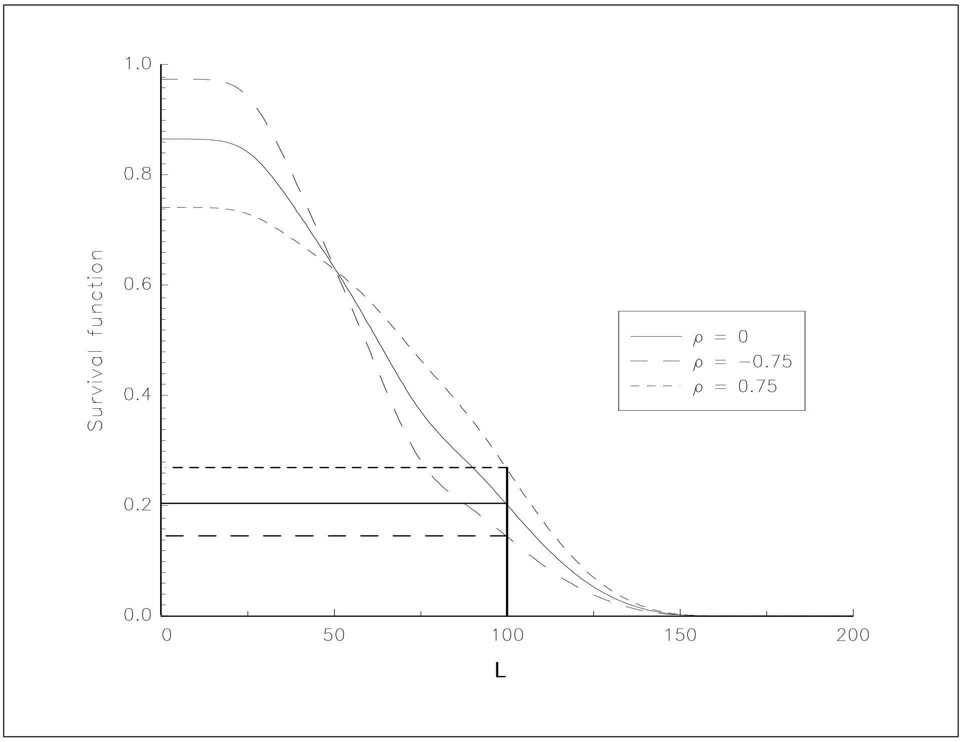

Figure 15: Probability $\Pr\{\mathfrak{L}(t_0) \geq \mathbf{L}\}$ for different values of ρ

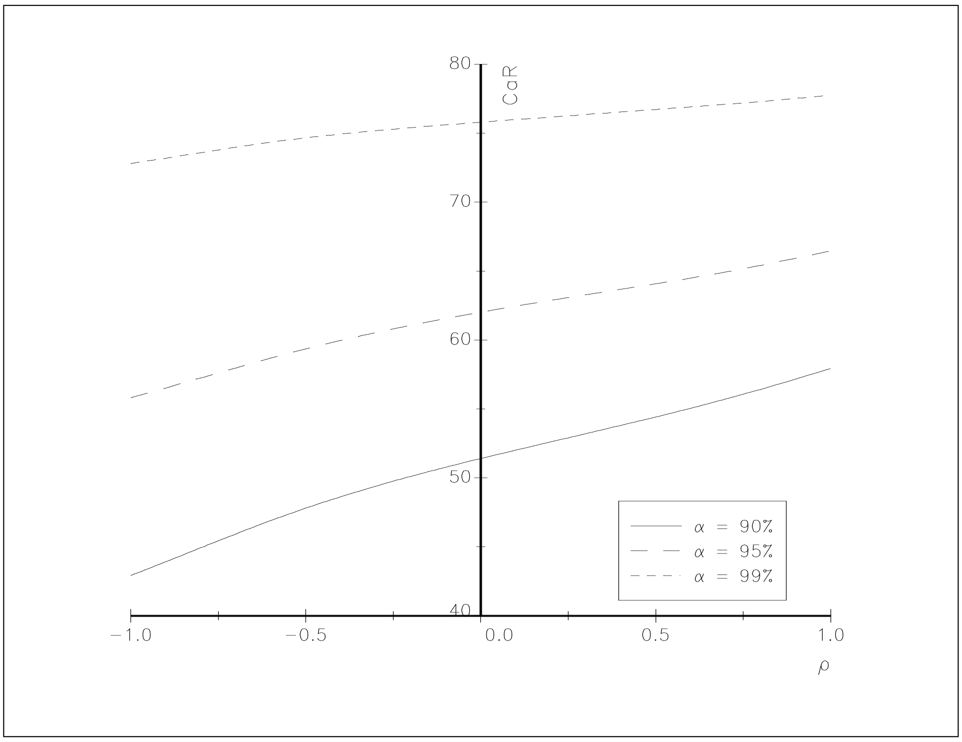

Figure 16: Capital-at-risk for different values of α

in the following way:

$$\boldsymbol{\rho} = \begin{bmatrix} 1 & 0.5 & \rho_{1,3} \\ & 1 & \rho_{2,3} \\ & & 1 \end{bmatrix}$$

We consider the credit lines with the following characteristics:

Credit line	B	t_n	λ_n	a_n	b_n	$\rho_{1,3}$	$\rho_{2,3}$
A	100	10	0.10	25%	75%	0	0
B	100	10	0.10	25%	75%	-0.50	-0.50
C	100	10	0.10	25%	75%	0.50	0.50
D	100	10	0.10	25%	75%	-0.25	0.25
E	100	10	0.05	75%	75%	0.50	0.50
F	100	10	0.10	50%	75%	0.50	0.50
G	150	10	0.10	75%	75%	0.50	0.50

We could then compute the **marginal** Capital-at-Risk induced by the third credit line. We obtain the following results:

α	A	B	C	D	E	F	G
90%	12.2	2.6	17.9	12.4	21.9	10.9	10.1
95%	16.3	10.3	20.3	16.4	27.1	12.3	11.4
99%	24.8	20.9	26.9	24.8	35.9	15.3	12.8

If we compare the first four credit lines, the only difference between them is the values taken by the coefficients $\rho_{1,3}$ and $\rho_{2,3}$. We obtain significant differences between the marginal economic capital. For the last three credit lines, the coefficients $\rho_{1,3}$ and $\rho_{2,3}$ are the same as for the credit line C, but the other characteristics are different. This example suggests that the economic capital of a credit portfolio is very sensitive to recovery rates, hazard rates and default correlations. As a result, the copula will play a major role in the evaluation of risk-adjusted performance using risk-adjusted return on capital (Punjabi [1998]). For example, if we define the performance as the internal rate of return between profitability and economic capital, we obtain the following results[27] ($\alpha = 99\%$):

Previous portfolio	A	B	C	D	E	F	G
14.58%	14.68%	15.12%	14.44%	14.68%	13.50%	15.80%	16.13%

It comes that the copula between the default times is important if the credit risk manager has a RAROC or ROE objective.

A more interesting application of copulas concerns the problem of economic capital allocation, not for a credit portfolio of one desk, but for the credit portfolio for the bank as a whole. In this case, it is interesting to use the copula framework to define risk-bucket capital charge (Gordy [2001]) or to propose closed-form expression of the economic capital (Ieda, Marumo and Yoshiba [2000]). These different points are studied in Coutant, Martineu, Messines, Riboulet and Roncalli [2001].

5.3 Pricing credit derivatives

We consider now the pricing of credit derivatives. Schönbucher [1998] defines credit derivatives as "*derivative securities whose payoff depends on the credit quality of a certain issuer. This credit quality can be measured by the credit rating of the issuer or by the yield spread of his bonds over the yield of a comparable default-free bond*". We remind that the default-free bond price with maturity t is given at time t_0 by

$$P(t_0, t) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_{t_0}^{t} r(s)\ \mathrm{d}s\right) \mid \mathcal{F}_{t_0}\right] \tag{214}$$

[27] The profitability is 150 for the first two credit lines and 100 for the third credit lines.

with $\mathbb{Q}$ the *martingale* probability measure and $r(t)$ the instantaneous interest rate. In the case where the interest rate and the survival time of the defaultable bond are independent, the price of the defaultable bond $\breve{P}(t_0, t)$ is given by (proposition 1 of MADAN and UNAL [2000]):

$$\breve{P}(t_0, t) = R \cdot P(t_0, t) + (1 - R) \cdot P(t_0, t) \cdot S(t) \tag{215}$$

where R is the constant recovery. In the case where $S(t)$ is one, we have of course $\breve{P}(t_0, t) = P(t_0, t)$. If we set R to zero, the credit spread $CS(t_0, t)$ is

$$\begin{aligned} CS(t_0, t) &= -\frac{1}{(t - t_0)} \ln\left(\frac{\breve{P}(t_0, t)}{P(t_0, t)}\right) \\ &= \frac{1}{(t - t_0)} \Lambda(t) \end{aligned} \tag{216}$$

We remark that the survival function (or the hazard function) of the defaultable bond plays an important role for its pricing. In the case of the valuation of basket credit derivatives, the important thing is the joint survival function, not only the univariate characterization of the survival times (LI [2000a]). Let us consider the case of a default digital put option (DDP). In SCHÖNBUCHER [1998], the European DDP "*pays off 1 at t iff there has been a default at some time before (or including) t*". If we assume that the interest rate and the first default $\tau = \bigwedge_{n=1}^{N} T_n$ are independent, the price is then

$$\begin{aligned} \mathsf{DDP}(t_0, t) &= \mathbb{E}\left[e^{-\int_{t_0}^{t} r(s)\,\mathrm{d}s} \mathbf{1}_{[\tau < t]}\right] \\ &= (1 - S_\tau(t)) P(t_0, t) \\ &= \left(1 - \breve{\mathbf{C}}(\mathbf{S}_1(t), \ldots, \mathbf{S}_N(t))\right) P(t_0, t) \end{aligned} \tag{217}$$

The price of this option is then influenced by the survival functions $\mathbf{S}_n(t)$ and the dependence $\breve{\mathbf{C}}$ between the survival times. We consider N defaultable securities with exponential survival times ($\lambda_n = 5\%$). We assume that the matrix $\boldsymbol{\rho}$ of parameters is of the following form:

$$\boldsymbol{\rho} = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ & 1 & \ddots & \vdots \\ & & 1 & \rho \\ & & & 1 \end{bmatrix} \tag{218}$$

The interest rate r is constant and is equal to 5%. In Figure 17, we have reported the DDP premium according to the maturity t and the parameter ρ in the case $N = 2$. We verify that $\rho_1 \geq \rho_2$ implies $\mathsf{DDP}_1(t_0, t) \leq \mathsf{DDP}_2(t_0, t)$. The relation between $\mathsf{DDP}(t_0, t)$ and the maturity t is more complex because there are two effects: $1 - S_\tau(t)$ is an increasing function of t, but $P(t_0, t)$ is a decreasing function of t. In Figure 18, we show the impact of the number of defaultable securities on the premium.

Let us consider now the simple case of the credit default swap (CDS) given by LI [2000b]. The product is defined by the following characteristics:

- there are two counterparties A (the default protection seller) and B (the buyer) and a bond issuer (or a reference security) C;
- the counterparty B pays to A a fix leg (or a swap rate) ℓ at time t_m ($m = 1, \ldots, M$) **until default or maturity** of the CDS;
- B receives **at default** a payment[28] $C(t_m)$ from A if the bond issuer C defaults; however, he only receives a fraction $R(t_m)$ of $C(t_m)$ if A defaults too.

[28] which is generally the difference between the recovery value and the notional value of the bond.

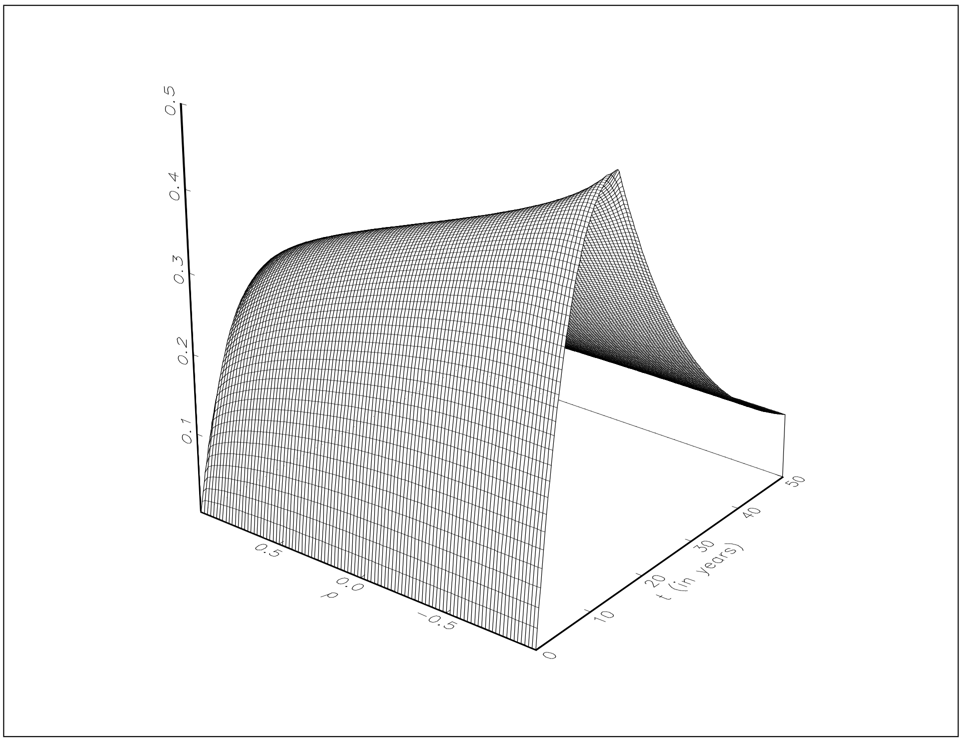

Figure 17: Influence of the parameter ρ on the DDP premium

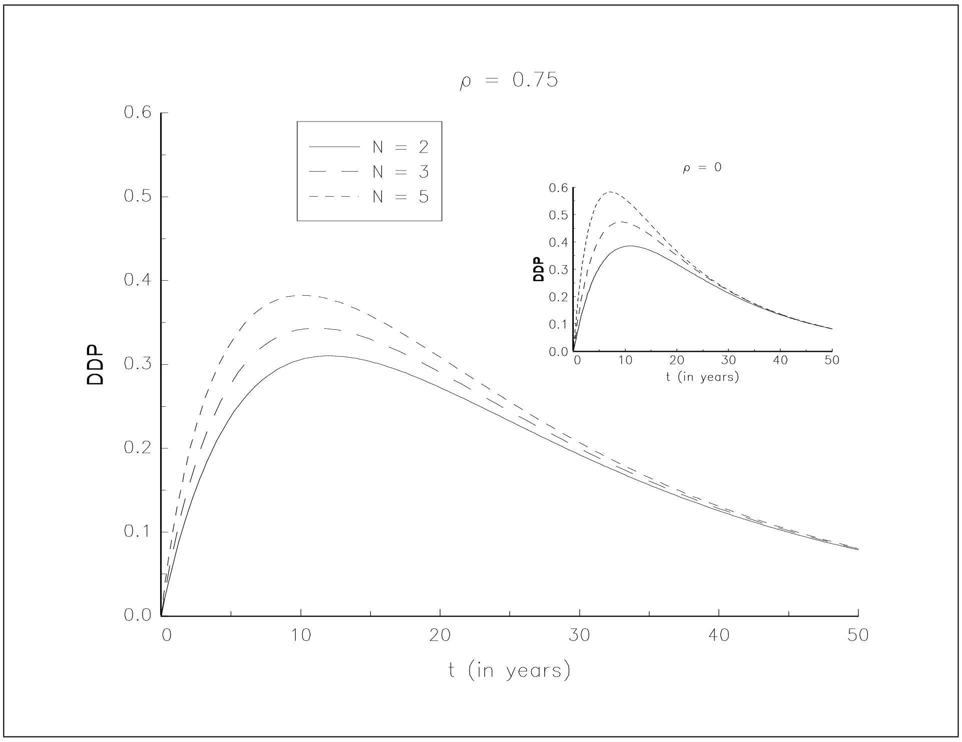

Figure 18: Influence of the number of defaultable securities on the DDP premium

It comes that the present value of the payment is

$$PV(t_0)=\sum_{m=1}^{M}P(t_0,t_m)\left[C(t_m)\Pr\{T_C\in[t_{m-1},t_m[\,,T_A>t_m\}+C(t_m)R(t_m)\Pr\{T_C\in[t_{m-1},t_m[\,,T_A\le t_m\}\right] \tag{219}$$

The present value of the periodic legs is

$$\begin{aligned} V(t_0) &= \ell\sum_{m=0}^{\min(T_C\wedge T_A,t_{M-1})}P(t_0,t_m) \\ &= \ell\sum_{m=1}^{M}P(t_0,t_{m-1})\Pr\{T_C>t_{m-1},T_A>t_{m-1}\} \end{aligned} \tag{220}$$

If we assume that the absence of arbitrage implies $PV(t_0)=V(t_0)$, we then obtain

$$\ell=\frac{\sum_{m=1}^{M}P(t_0,t_m)\left[C(t_m)\Pr\{T_C\in[t_{m-1},t_m[\,,T_A>t_m\}+C(t_m)R(t_m)\Pr\{T_C\in[t_{m-1},t_m[\,,T_A\le t_m\}\right]}{\sum_{m=1}^{M}P(t_0,t_{m-1})\Pr\{T_C>t_{m-1},T_A>t_{m-1}\}} \tag{221}$$

ℓ depends then on the bivariate survival function $\mathbf{S}$ of (T_C,T_A):

$$\ell=\frac{\sum_{m=1}^{M}P(t_0,t_m)C(t_m)\left[(1-R(t_m))(S(t_{m-1},t_m)-S(t_m,t_m))+R(t_m)(S_C(t_{m-1})-S_C(t_m))\right]}{\sum_{m=1}^{M}P(t_0,t_{m-1})S(t_{m-1},t_{m-1})} \tag{222}$$

If $R(t_m)=0$, this formula becomes

$$\ell=\frac{\sum_{m=1}^{M}P(t_0,t_m)C(t_m)\left[S(t_{m-1},t_m)-S(t_m,t_m)\right]}{\sum_{m=1}^{M}P(t_0,t_{m-1})S(t_{m-1},t_{m-1})} \tag{223}$$

We consider a CDS with annual payment dates. The parameters take the following values: $C(t_m)=1$, T_A and T_C are two exponential survival times with hazard rate $\lambda_A=\lambda_C=5\%$. Moreover, we assume that the interest rate r is constant. We have reported the value of the leg ℓ in Figures 19 and 20. In Figure 19, we remark the influence of the parameter ρ of the Normal copula. Moreover, we see how moves ℓ with parameters M or r. In this example, we have

$$\rho_1\ge\rho_2\ (\text{or } \mathbf{C}_{\rho_1}\succ\mathbf{C}_{\rho_2})\Longrightarrow\ell_1\le\ell_2 \tag{224}$$

In Figure 20, the interest rate r is equal to 5% and the value of M is five years. We verify that $R_1(t_m)\ge R_2(t_m)$ implies $\ell_1\ge\ell_2$.

Other credit derivatives could be considered. In particular, the copula approach seems to be adequate to price the *first-to-default*. A first-to-default (FtoD) is a contingent claim that pays at the first of N credit events an amount $\varpi(\tau)$. We assume that the price of the FtoD is given by the expected value of the amount paid

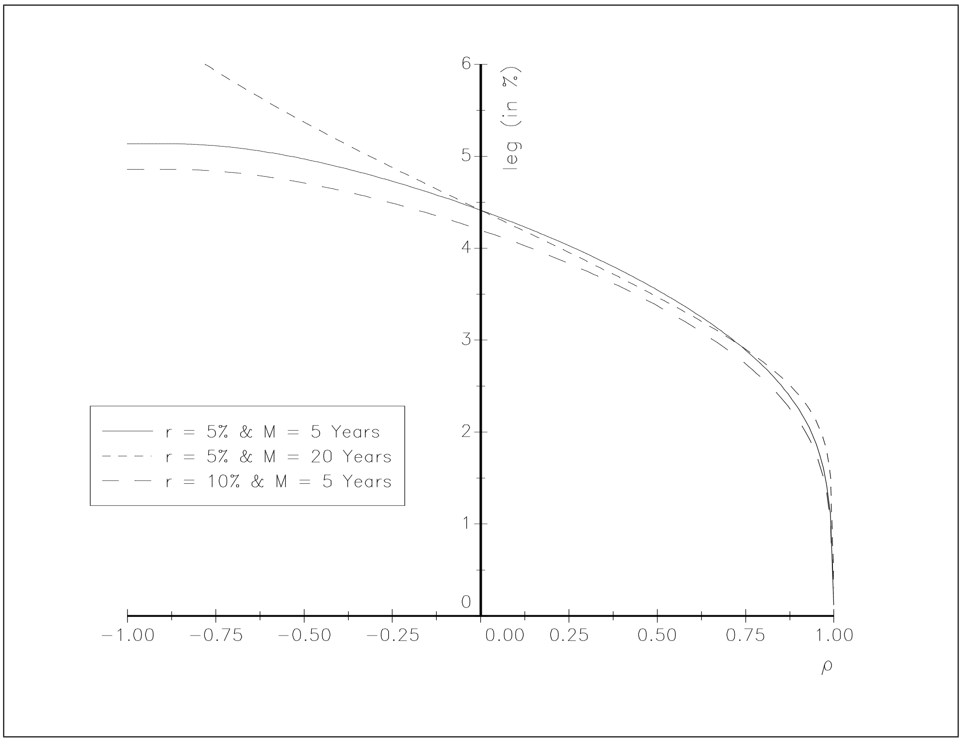

Figure 19: Value of the CDS leg ℓ with $R(t_m) = 0$

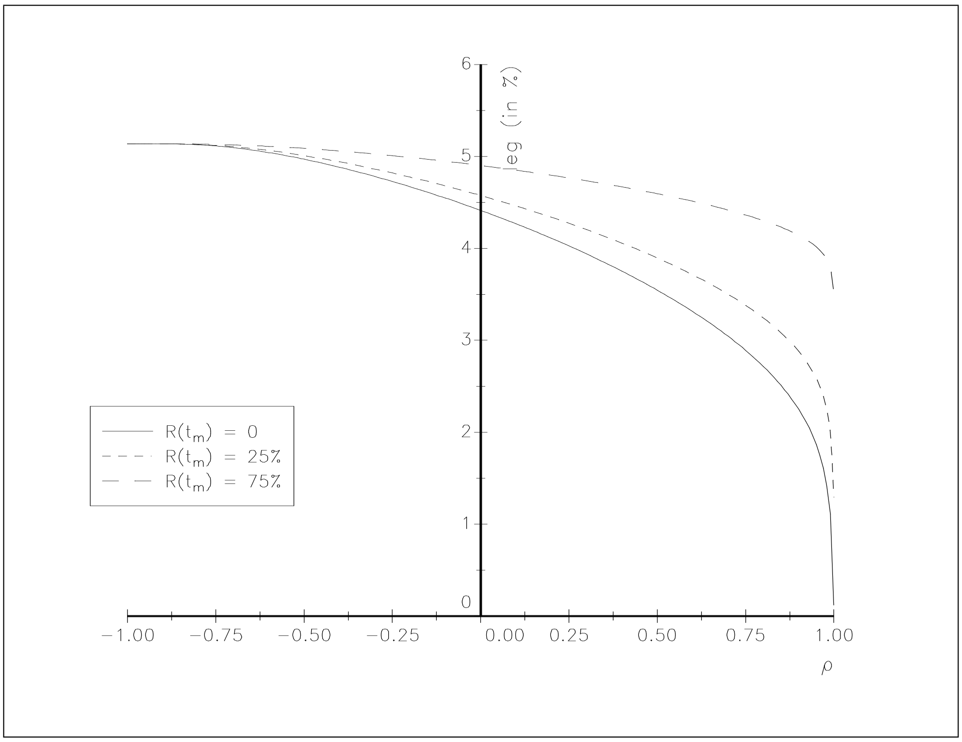

Figure 20: Value of the CDS leg ℓ with $R(t_m) \neq 0$

$\varpi(\tau)$:

$$\begin{aligned}\mathsf{FtoD}(t_0,t) &= \mathbb{E}\left[\varpi(\tau)\, e^{-r(\tau)}\mathbf{1}_{[\tau<t]}\right] \\ &= \int_{t_0}^{t}\varpi(s)\, e^{-r(s)} f_\tau(s)\;\mathrm{d}s \\ &= \sum_{n=1}^{N}\int_{t_0}^{t}\varpi(s)\, e^{-r(s)} f_n(s)\,\partial_n\breve{\mathbf{C}}\left(\mathbf{S}_1(s),\ldots,\mathbf{S}_N(s)\right)\,\mathrm{d}s \end{aligned} \tag{225}$$

More generally, we could consider credit derivatives based on the n^{th}-to-default and we note $P_{n:N}(t_0,t)$ the corresponding price. We have

$$\begin{aligned}P_{n:N}(t_0,t) &= \mathbb{E}\left[\varpi(T_{n:N})\, e^{-r(T_{n:N})}\mathbf{1}_{[T_{n:N}<t]}\right] \\ &= \int_{t_0}^{t}\varpi(s)\, e^{-r(s)} f_{n:N}(s)\;\mathrm{d}s \end{aligned} \tag{226}$$

We have of course $P_{1:N}(t_0,t) = \mathsf{FtoD}(t_0,t)$ and $P_{N:N}(t_0,t) = \mathsf{LtoD}(t_0,t)$ (*last-to-default*). In what follows, we assume that $\varpi(t) = 1$. We consider the previous example of the DDP option. In Figures 21 and 22, we have represented the prices $\mathsf{FtoD}(t_0,t)$ for different values of ρ and t. The only difference between the two graphs is the value of the hazard rate of the second defaultable security. We remark that the dependence plays a more important role when the survival times of the defaultable securities are "close" than when the survival times of the defaultable securities are different. In Figures 23, 24 and 25, we suppose that the hazard rates are the same and are equal to 5%. We see the impact of the maturity t and the parameter ρ. In Figure 25, we have reported the premium of several n^{th}-to-default. The dashed horizontal line indicates the value of the first-to-default in the case $N = 1$:

$$P_{1:1}(t_0,t) = \int_{t_0}^{t} e^{-r(s)} f(s)\;\mathrm{d}s \tag{227}$$

For this example, we remark that $P_{1:1}(t_0,t) \le P_{1:N}(t_0,t)$ and $P_{1:1}(t_0,t) \ge P_{N:N}(t_0,t)$ and we verify that $P_{n:N}(t_0,t) = P_{1:1}(t_0,t)$ if the dependence function is $\mathbf{C}^{+}$.

These examples show the importance of the dependence function in pricing credit derivatives. However, we have only built illustrations with a monotone dependence. When the copula is neither PQD nor NQD, its impact on prices could be very complex (Coutant, Martineu, Messines, Riboulet and Roncalli [2001]). Moreover, we have studied simple credit derivatives. In the case of more complicated products, the price is critically dependent on the survival copula. This is for example the case of CBO (Davis and Lo [2000]). We refer interested readers to Coutant, Martineu, Messines, Riboulet and Roncalli [2001] for a more complete treatement of these topics.

References

[1] Basel Committee on Banking Supervision, Credit Risk Modelling: Current Practices and Applications, April 1999

[2] Basel Committee on Banking Supervision, The Internal Ratings-Based Approach — Consultative Document, Supporting document to the New Basel Capital Accord, January 2001

[3] Abbring, J.H. and G.J. Van den Berg [2000], The non-parametric identification of the mixed proportional hazards competing risks model, Tinbergen Institute, *Discussion Paper*, **2000-066/3**

[4] Aïssaoui, F. and A. Frachot [1999], Modèles d'évaluation des remboursements anticipés des prêts immobiliers, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Note interne*

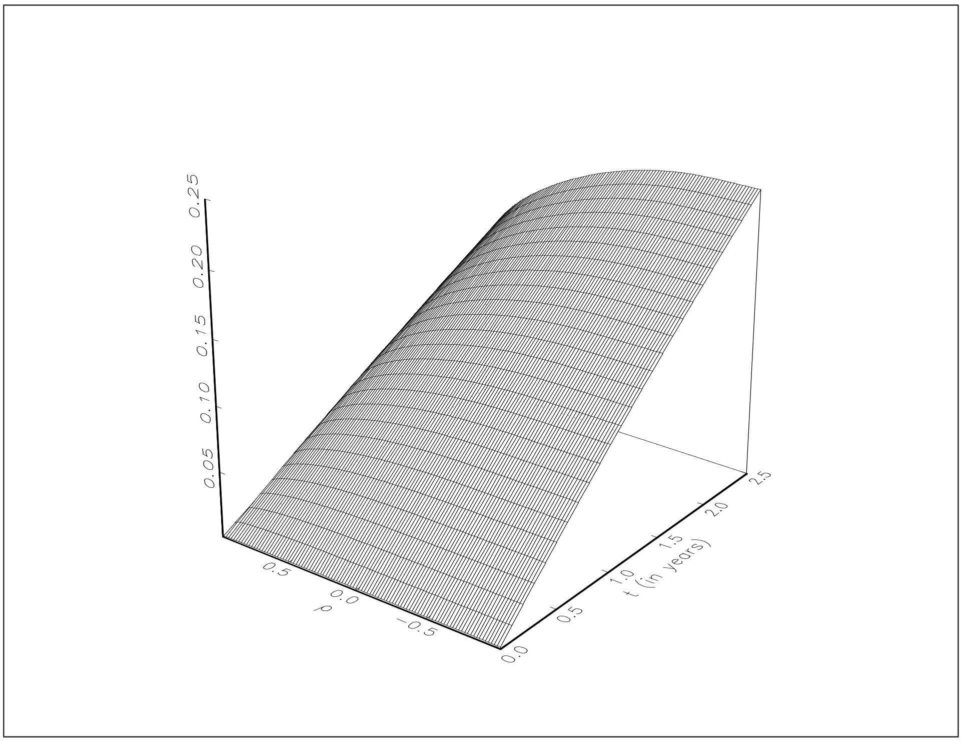

Figure 21: Value of the FtoD premium ($N = 2$, $\lambda_1 = 5\%$ and $\lambda_2 = 5\%$)

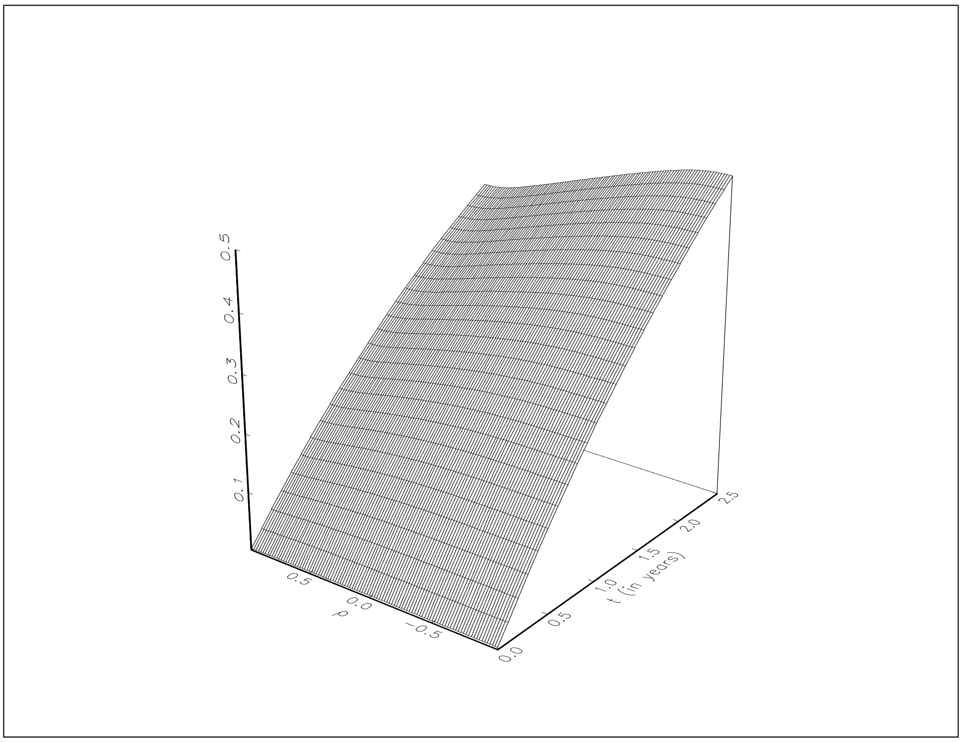

Figure 22: Value of the FtoD premium ($N = 2$, $\lambda_1 = 5\%$ and $\lambda_2 = 25\%$)

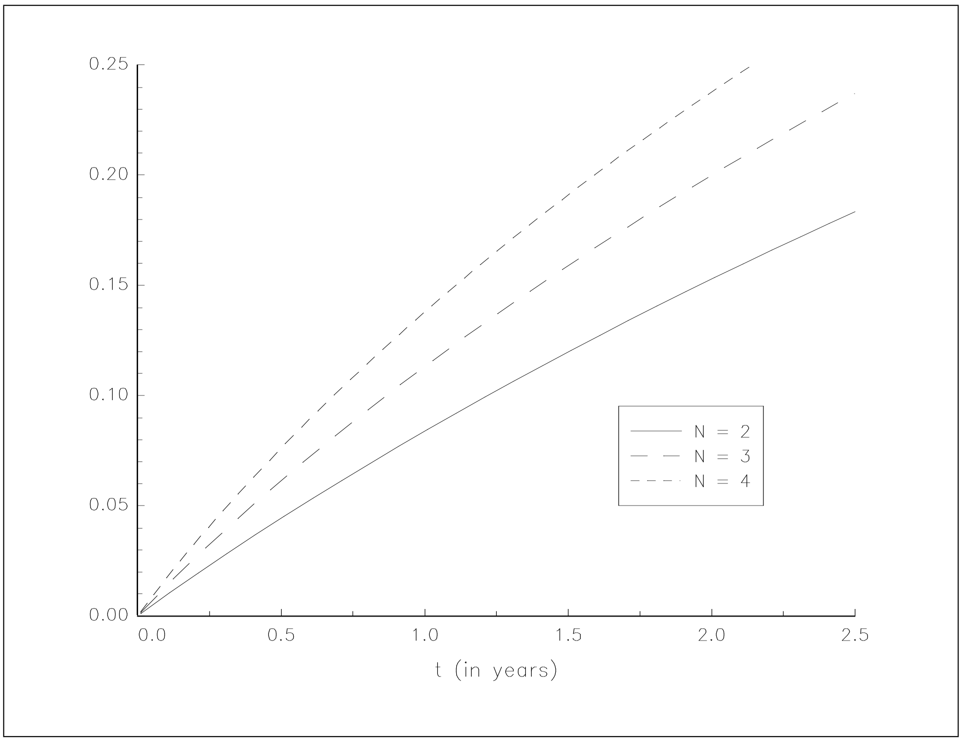

Figure 23: Value of the FtoD premium for $\rho = 0.5$

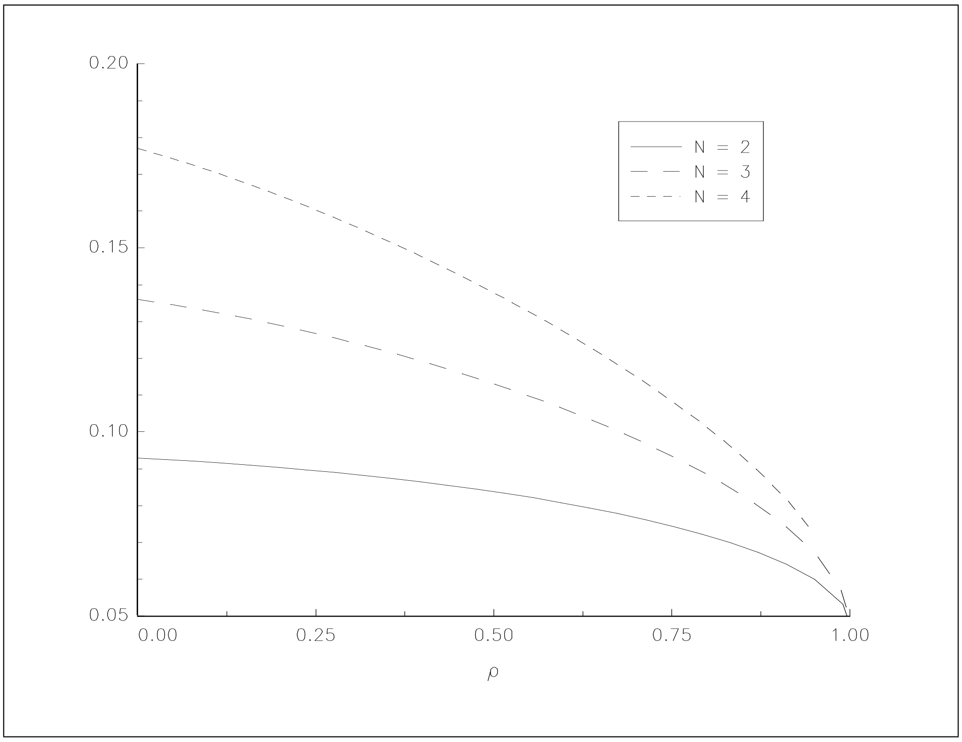

Figure 24: Value of the FtoD premium for $t = 1$

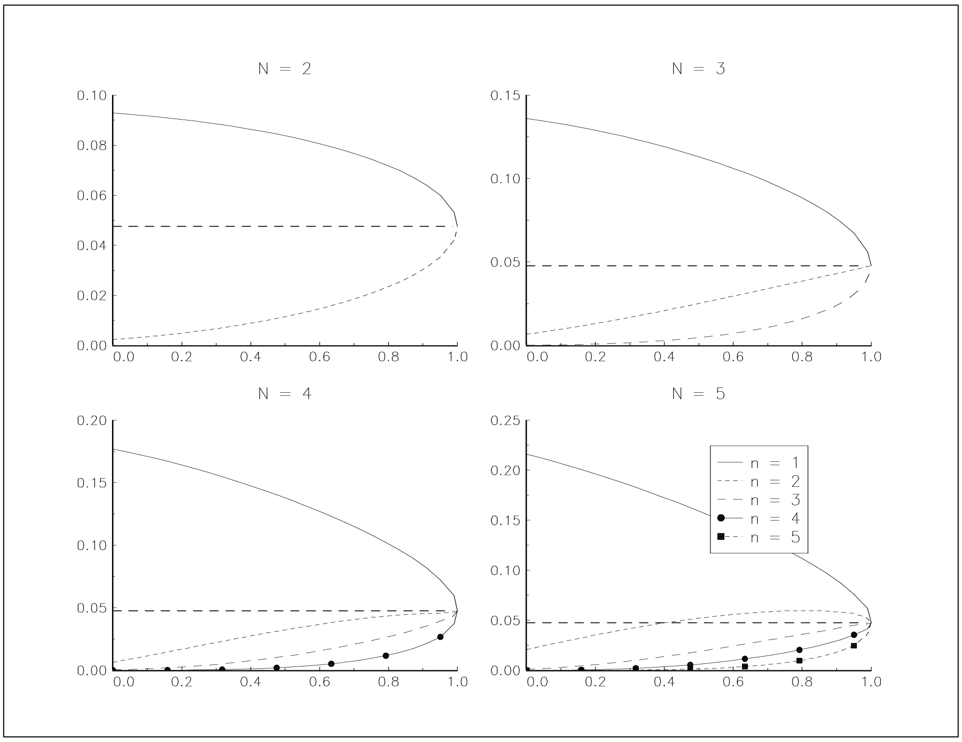

Figure 25: Value of the premium $P_{n:N}\left(t_0, t\right)$ for $t = 1$

[5] Andersen, P.K., O. Borgan, R.D. Gill and N. Keiding [1993], Statistical Models Based on Counting Processes, *Springer Series in Statistics*, Springer-Verlag, New York

[6] Anderson, J.E. and T.A. Louis [1995], Survival analysis using a scale change random effects model, *Journal of the American Statistical Association*, **90**, 669-679

[7] Anderson, J.E., T.A. Louis, N.V. Holm and B. Harvald [1992], Time-dependent association measures for bivariate survival distributions, *Journal of the American Statistical Association*, **87**, 641-650

[8] Bagdonavicius, V., S. Malov and M. Nikulin [1998], On semiparametric regression estimation in copula models, Université de Bordeaux II, *Working Paper*, **98019**

[9] Bagdonavicius, V., S. Malov and M. Nikulin [1999], Characterizations and semiparametric regression estimation in Archimedean copula, *Journal of Applied Statistical Science*, **8(2-3)**, 137-153

[10] Baud, N. and T-T. Trang [1999], Analyse des remboursements anticipés des prêts immobiliers à taux fixe — une approche individuelle par les modèles de durée, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Note interne*

[11] Balakrishnan, N. and A. Clifford Cohen [1991], Order Statistics and Inference, Academic Press, San Diego

[12] Bandeen-Roche, K.J. and K-Y. Liang [1996], Modelling failure-time associations in data with multiple levels of clustering, *Biometrika*, **83(1)**, 29-39

[13] Barnett, V. [1980], Some bivariate uniform distributions, *Communications in Statistics – Theory and Methods*, **9(4)**, 453-461

[14] Basu, A.P. [1971], Bivariate failure rate, *Journal of the American Statistical Association*, **66**, 103-104

[15] Basu, A.P. [1988], Multivariate exponential distributions and their applications in reliability, in P.R. Krishnaiah and C.R. Rao (Eds.), *Handbook of Statistics, Vol. 7 — Quality Control and Reliability*, North-Holland, Amsterdam

[16] Billingsley, P. [1999], Convergence of Probability Measures, *Wiley Series in Probability and Statistics, second edition*, John Wiley & Sons, New York (first edition 1968)

[17] Block, H.W. and A.P. Basu [1974], A continuous bivariate exponential extension, *Journal of the American Statistical Association*, **69**, 1031-1037

[18] Block, H.W. and T.H. Savits [1980], Multivariate increasing failure rate average distributions, *Annals of Probability*, **8**, 793-801

[19] Bouyé, E., V. Durrleman, A. Nikeghbali, G. Riboulet and T. Roncalli [2000a], Copulas for finance — A reading guide and some applications, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper* (available from `http://gro.creditlyonnais.fr`)

[20] Bouyé, E., V. Durrleman, A. Nikeghbali, G. Riboulet and T. Roncalli [2000b], Copulas: an open field for risk management, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper* (available from `http://gro.creditlyonnais.fr`)

[21] Cambanis, S., G. Simmons and W. Stout [1976], Inequalities for $\mathcal{E}k\left(X,Y\right)$ when the marginals are fixed, *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, **36**, 285-294

[22] Carriere, J.F. [1994], Dependent decrement theory, *Transactions: Society of Actuaries*, **XLVI**, 45-74

[23] Carriere, J.F. [2000], Bivariate survival models for coupled lives, *Scandinavian Actuarial Journal*, **100(1)**, 17-32

[24] 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(1)**, 141-151

[25] Clayton, D.G. and J. Cuzick [1985], Multivariate generalization of the proportional hazards model (with discussion), *Journal of the Royal Statistical Society – Series A*, **148(2)**, 82-117

[26] Coles, S., J. Currie and J. Tawn [1999], Dependence measures for extreme value analyses, Department of Mathematics and Statistics, Lancaster University, *Working Paper*

[27] Coutant, S., P. Martineu, J. Messines, G. Riboulet and T. Roncalli [2001], Credit risk modelling with copulas, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper* (available from `http://gro.creditlyonnais.fr`)

[28] Cox, D.R. [1972], Regression models and life-tables (with discussion), *Journal of the Royal Statistical Society – Series B*, **34**, 187-220

[29] Crouhy, M. [1999], Measuring credit risk: comparative analysis of CreditVaR models, Presentation at the conference "Credit Risk Models and the Regulation of Banks Capital Requirements" (TABE, 03/30/1999)

[30] Crouhy, M., D. Galai and R. Mark [2000], A comparative analysis of current credit risk models, *Journal of Banking and Finance*, **24**, 59-117

[31] Csörgő, M. [1983], Quantile Processes with Statistical Applications, *CBMS-NSF Regional Conference Series in Applied Mathematics*, **42**, Society for Industrial and Applied Mathematics, Philadelphia

[32] Dabrowska, D.M. [1988], Kaplan-Meier estimate on the plane, *Annals of Statistics*, **16(4)**, 1475-1489

[33] Dabrowska, D.M. [1989a], Kaplan-Meier estimate on the plane: weak convergence, LIL and the bootstrap, *Journal of Multivariate Analysis*, **29(2)**, 308-325

[34] Dabrowska, D.M. [1989b], Uniform consistency of the kernel conditional Kaplan-Meier estimate, *Annals of Statistics*, **17(3)**, 1157-1167

[35] Dabrowska, D.M. [1996], Weak convergence of a product integral dependence measure, *Scandinavian Journal of Statistics*, **23(4)**, 551-580

[36] David, H.A. [1970], Order Statistics, *Wiley Series in Probability and Mathematical Statistics*, John Wiley & Sons, New York

[37] Davis, M. and V. Lo [2000], Modelling default correlation in bond portfolios, Tokyo-Mitsubishi International plc, *Working Paper*

[38] Deheuvels, P. [1978], Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes, *Publications de l'Institut de Statistique de l'Université de Paris*, **23(3-4)**, 1-36

[39] Deheuvels, P. [1979a], Propriétes d'existence et propriétes topologiques des fonctions de dependance avec applications à la convergence des types pour des lois multivariées, *Comptes-Rendus de l'Académie des Sciences de Paris – Série 1*, **288**, 145-148

[40] Deheuvels, P. [1979b], La fonction de dépendance empirique et ses propriétés — Un test non paramétrique d'indépendance, *Académie Royale de Belgique – Bulletin de la Classe des Sciences – 5e Série*, **65**, 274-292

[41] Deheuvels, P. [1980], Non parametric tests of independence, in J.P. Raoult (Ed.), *Statistique non Parametrique Asymptotique — Proceedings, Actes des Journees Statistiques, Rouen, France, Juin 1979*, *Lecture Notes in Mathematics*, **821**, Springer-Verlag, New York

[42] Deheuvels, P. [1981], An asymptotic decomposition for multivariate distribution-free tests of independence, *Journal of Multivariate Analysis*, **11**, 102-113

[43] Demey, P., A. Frachot and G. Riboulet [2000], Note sur l'évaluation de l'option de remboursement anticipé, *Banque & Marchés*, **49**, 6-14 (available from `http://gro.creditlyonnais.fr`)

[44] Dempster, A.P., N.M. Laird and D.B. Rubin [1977], Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society – Series B*, **39**, 1-38

[45] Duffie, D. and K. Singleton [1999], Simulating correlated defaults, Graduate School of Business, Standford University, *Working Paper*

[46] Durrleman, V., A. Nikeghbali and T. Roncalli [2000a], Which copula is the right one?, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper* (available from `http://gro.creditlyonnais.fr`)

[47] Durrleman, V., A. Nikeghbali and T. Roncalli [2000b], A note about the conjecture on Spearman's rho and Kendall's tau, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper* (available from `http://gro.creditlyonnais.fr`)

[48] Embrechts, P., A.J. McNeil and D. Straumann [1999], Correlation: pitfalls and alternatives, *Risk magazine*, **12**, May, 69-71

[49] Embrechts, P., A.J. McNeil and D. Straumann [2000], Correlation and dependency in risk management: properties and pitfalls, in M. Dempster and H.K. Moffatt (Eds.), *Risk Management: Value at Risk and Beyond*, Cambridge University Press

[50] FAN, J., R.L. PRENTICE and L. HSU [2000], A class of weighted dependence measures for bivariate failure time data, *Journal of the Royal Statistical Society – Series B*, **62**, 181-190

[51] FELLER, W. [1968], An introduction to Probability Theory and Its Applications — Volume I, *Wiley Series in Probability and Mathematical Statistics*, *third edition*, John Wiley & Sons, New York (first edition 1950)

[52] FELLER, W. [1971], An introduction to Probability Theory and Its Applications — Volume II, *Wiley Series in Probability and Mathematical Statistics*, *second edition*, John Wiley & Sons, New York (first edition 1966)

[53] FERMANIAN, J.-D. [1997], Multivariate hazard rates under random censorship, *Journal of Multivariate Analysis*, **62(2)**, 273-309

[54] FINE, J.P. and H. JIANG [2000], On association in a copula with time transformation, *Biometrika*, **87(3)**, 559-571

[55] FRACHOT, A. and C. GOURIEROUX [1995], Titrisation et Remboursements anticipés, Economica, Paris

[56] FRANK, M.J. [1979], On the simultaneous associativity of $\mathbf{F}(x,y)$ and $x+y-\mathbf{F}(x,y)$, *Aequationes Mathematicae*, **19**, 194-226

[57] FRANK, M.J. [1991], Convolutions for dependent random variables, in G. Dall'Aglio, S. Kotz and G. Salinetti (Eds.), *Advances in Probability Distributions with Given Marginals (Beyond the Copulas)*, Kluwer Academic Publishers, Dordrecht

[58] FREES, E.W., J.F. CARRIERE and E.A. VALDEZ [1996], Annuity valuation with dependent mortality, *Journal of Risk and Insurance*, **63(2)**, 229-261

[59] FREES, E.W. and E.A. VALDEZ, [1998], Understanding relationships using copulas, *North American Actuarial Journal*, **2**, 1-25

[60] FREUND, J.E. [1961], A bivariate extension of the exponential distribution, *Journal of the American Statistical Association*, **56**, 971-977

[61] GAENSSLER, P. and W. STUTE [1987], Seminar on Empirical Processes, *DMV Seminar*, **9**, Birkhäuser Verlag Basel, Boston

[62] GALAMBOS, J. [1982], The role of functional equations in stochastic model building, *Aequationes Mathematicae*, **25**, 21-41

[63] GALAMBOS, J. and S. KOTZ [1978], Characterizations of probability distributions, *Lecture Notes in Mathematics*, **675**, Springer-Verlag, Heidelberg

[64] GENEST, C. [1987], Frank's family of bivariate distributions, *Biometrika*, **74(3)**, 549-555

[65] GENEST, C., K. GHOUDI and L-P RIVEST [1995], A semiparametric estimation procedure for dependence parameters in multivariate families of distributions, *Biometrika*, **82(3)**, 543-552

[66] GENEST, C. and J. MACKAY [1986a], Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données, *Canadian Journal of Statistics*, **14(2)**, 145-159

[67] GENEST, C. and J. MACKAY [1986b], The joy of copulas: Bivariate distributions with uniform marginals, *American Statistician*, **40**, 280-283

[68] GENEST, C. and L. RIVEST [1993], Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association*, **88**, 1034-1043

[69] Genest, C. and B.J.M. Werker [2002], Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence paramters in copula models, in C.M. Cuadras and J.A. Rodríguez Lallena (Eds.), *Proceedings of the Conference on Distributions with Given Marginals and Statistical Modelling (Barcelona, Espagne)*, Kluwer Academic Publishers, Dordrecht

[70] Georges, P., D. Jacomy and S-H. Lazare [2000], Mesure de la valeur Client (LTV), Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Note interne*

[71] Gill, R.D. [1992a], Multivariate survival analysis, *Theory of Probability and Its Applications*, **37(1)**, 18-31

[72] Gill, R.D. [1992b], Multivariate survival analysis — Part 2: Technique, *Theory of Probability and Its Applications*, **37(2)**, 284-301

[73] Gill, R.D. [1994], Lectures on survival analysis, in D. Bakry, R.D. Gill, and S. Molchanov (Eds), *Lectures on Probability Theory — Ecole d'Eté de Probabilités de Saint-Flour XXII-1992*, *Lectures Notes on Mathematics*, **1581**, Springer-Verlag, Berlin

[74] Gill, R.D. and S. Johansen [1990], A survey of product-integration with a view toward application in survival analysis, *Annals of Statistics*, **18(4)**, 1501-1555

[75] Gill, R.D. M.J. van der Laan, J.A. Wellner [1995], Inefficient estimators of the bivariate survival function for three models, *Annales de l'Institut Henri Poincaré – Calcul des Probabilités et Statistique*, **31(3)**, 545-597

[76] Glidden, D.V. and S.G. Self [1999], Semiparametric likelihood estimation in the Clayton-Oakes failure time model, *Scandinavian Journal of Statistics*, **26(3)**, 363-372

[77] Gordy, M.B. [2000], A comparative anatomy of credit risk models, *Journal of Banking and Finance*, **24**, 119-149

[78] Gordy, M.B. [2001], A risk-factor model foundation for ratings-based bank capital rules, Federal Reserve Board, *Working Paper*

[79] Guo, G. and G. Rodríguez [1992], Estimating a multivariate proportional hazards model for clustered data using the EM algorithm, with an application to child survival in guatemala, *Journal of the American Statistical Association*, **87**, 969-976

[80] Gumbel, E. J. [1960], Bivariate exponential distributions, *Journal of the American Statistical Association*, **55**, 698-707

[81] Gumbel, E. J. [1961], Bivariate logistic distributions, *Journal of the American Statistical Association*, **56**, 335-349

[82] Hanley, J.A. and M.N. Parnes [1983], Nonparametric estimation of a multivariate distribution in the presence of censoring, *Biometrics*, **39**, 129-139

[83] Heckman, J.J. and B.E. Honoré [1989], The identifiability of the competing risks model, *Biometrika*, **76(2)**, 325-330

[84] Honoré, B.E. [1993], Identification results for duration models with multiple spells, *Review of Economic Studies*, **60(1)**, 241-246

[85] Hougaard, P. [1986a], Survival models for heterogeneous populations derived from stable distributions, *Biometrika*, **73(2)**, 387-396

[86] Hougaard, P. [1986b], A class of multivariate failure time distributions, *Biometrika*, **73(3)**, 671-678

[87] Hougaard, P. [1987], Modelling multivariate survival, *Scandinavian Journal of Statistics*, **14(4)**, 291-30

[88] HOUGAARD, P. [1999], Fundamentals of survival data, *Biometrics*, **55**, 13-22

[89] HOUGAARD, P. [2000], Analysis for Multivariate Survival Data, *Statistics for Biology and Health*, Springer-Verlag, New York

[90] HOUGAARD, P., B. HARVALD and N.V. HOLM [1992], Measuring the similarities between the lifetimes of adult Danish twins born between 1881-1930, *Journal of the American Statistical Association*, **87**, 17-24

[91] HUTCHINSON, T.P. and C.D. LAI [1990], Continous Bivariate Distributions — Emphasising Applications, Rumsby Scientific Publishing, Adelaide

[92] HUTCHINSON, T.P. and C.D. LAI [1991], The Engineering Statistician's Guide to Continous Bivariate Distributions, Rumsby Scientific Publishing, Adelaide

[93] IEDA, A., K. MARUMO and T. YOSHIBA [2000], A simplified method for calculating the credit risk of lending portfolios, Institute for Monetary and Economic Studies, Bank of Japan, *Discussion Paper*, **2000-E-10**

[94] JOE, H. [1997], Multivariate Models and Dependence Concepts, *Monographs on Statistics and Applied Probability*, **73**, Chapman & Hall, London

[95] JOE, H. and T. HU [1996], Multivariate distributions from mixtures of max-infinitely divisible distributions, *Journal of Multivariate Analysis*, **57(2)**, 240-265

[96] JOE, H. and C. MA [2000], Multivariate survival functions with a min-stable property, *Journal of Multivariate Analysis*, **75(1)**, 13-35

[97] JOE, H. and J.J. XU [1996], The estimation method of inference functions for margins for multivariate models, Department of Statistics, University of British Columbia, *Technical Report*, **166**

[98] JOHNSON, N. and S. KOTZ [1975], On some generalized Farlie-Gumbel-Morgenstern distributions, *Communications in Statistics*, **4**, 415-427

[99] KLAASSEN, C.A.J. and J.A. WELLNER [1997], Efficient estimation in the bivariate normal copula model: normal margins are least favorable, *Bernoulli*, **3(1)**, 55-77

[100] KOTZ, S., N. BALAKRISHNAN and N.L. JOHNSON [2000], Continuous Multivariate Distributions — Volume 1: Models and Applications, John Wiley & Sons, New York

[101] KOYLUOGLU, H.U. and A. HICKMAN [1998], A generalized framework for credit risk portfolio models, *Working Paper*

[102] LANCASTER, T. [1979], Econometric methods for the duration of unemployment, *Econometrica*, **47**, 939-956

[103] LANCASTER, T. [1990], The Econometric Analysis of Transition Data, *Econometric Society Monographs*, **17**, Cambridge University Press, Cambridge

[104] LEE, L. [1979], Multivariate distributions having Weibull properties, *Journal of Multivariate Analysis*, **9**, 267-277

[105] LEUNG, K-M., R.M. ELASHOFF and A.A. AFIFI [1997], Censoring issues in survival analysis, *Annual Review of Public Health*, **18**, 83-104

[106] LI, D.X. [2000a], On default correlation: a copula function approach, *Journal of Fixed Income*, **9(4)**, 43-54

[107] LI, D.X. [2000b], The valuation of basket credit derivatives: a copula function approach, *Workshop on Options in Financial Products: Approaches to Valuation*, 8 December, Toronto

[108] Lindskog, F. and A.J. McNeil [2001], Common poisson shock models, RiskLab, *Research Paper*

[109] Ling, C-H. [1965], Representation of associative functions, *Publicationes Mathematicae (Publ. Math. Debrecen)*, **12**, 189-212

[110] Maccarinelli, M. and V. Maggiolini [2000], The envolving practice of credit risk management in global financial institutions, *Risk Conference*, 26/27 September, Paris

[111] Maguluri, G. [1993], Semiparametric estimation of association in a bivariate survival function, *Annals of Statistics*, **21(3)**, 1648-1662

[112] Madan, D. and H. Unal [2000], A two-factor hazard rate model for pricing risky debt and the term structure of credit spreads, *Journal of Financial and Quantitative Analysis*, **35(1)**, 43-65

[113] Manatunga, A.K. and D. Oakes [1996], A measure of association for bivariate frailty distributions, *Journal of Multivariate Analysis*, **56(1)**, 60-74

[114] Marshall, A.W. and I. Olkin [1967], A multivariate exponential distribution, *Journal of the American Statistical Association*, **62**, 30-44

[115] Marshall, A.W. and I. Olkin [1988], Families of multivariate distributions, *Journal of the American Statistical Association*, **83**, 834-841

[116] Nelsen, R.B. [1995], Copulas, characterization, correlation, and counterexamples, *Mathematics Magazine*, **68**, 193-198

[117] Nelsen, R.B. [1999], An Introduction to Copulas, *Lectures Notes in Statistics*, **139**, Springer-Verlag, New York

[118] Nielsen, G.G., R.D. Gill, P.K. Andersen and T.I.A. Sorensen [1992], A counting process approach to maximum likelihood estimation in frailty models, *Scandinavian Journal of Statistics*, **19(1)**, 25-43

[119] Oakes, D. [1982], A model for association in bivariate survival data, *Journal of the Royal Statistical Society – Series B*, **44(3)**, 414-422

[120] Oakes, D. [1986], Semiparametric inference in a model for association in bivariate survival data, *Biometrika*, **73(2)**, 353-361

[121] Oakes, D. [1989], Bivariate survival models induced by frailties, *Journal of the American Statistical Association*, **84**, 487-493

[122] Oakes, D. [1994], Multivariate survival distributions, *Journal of Nonparametric Statistics*, **3**, 343-354

[123] Phelps, A.L. and L.A. Weissfeld [1997], A comparison of dependence estimators in bivariate copula models, *Communications in Statistics – Simulation and Computation*, **26(4)**, 1583-1597

[124] Pollard, D. [1990], Empirical Processes: Theory and Applications, *NSF-CBMS Regional Conference Series in Probability and Statistics*, **2**, Institute of Mathematical Statistics, Hayward, CA

[125] Prentice, R.L., J.D. Kalbfleisch, A.V. Peterson, N. Flournoy, V.T. Farewell and N.E. Breslow [1978], The analysis of failure times in the presence of competing risks, *Biometrics*, **34**, 541-554

[126] Punjabi, S. [1998], Many happy returns, *Risk*, **June**, 71-76

[127] Resnick, S.I. [1987], Extreme Values, Point Processes and Regular Variation, Springer-Verlag, New York

[128] Scarsini, M. [1984], On measures of concordance, *Stochastica*, **8**, 201-218

[129] Schönbucher, P.J. [1998], Pricing credit risk derivatives, London School of Economics, Financial Markets Group, *Working Paper*

[130] Schwartz, E. and W. Torous [1989], Prepayment and the valuation of mortgage-backed securities, *Journal of Finance*, **44(2)**, 375-392

[131] Schwartz, E. and W. Torous [1992], Prepayment, default and the valuation of mortgage pass-through securities, *Journal of Business*, **65(2)**, 221-239

[132] Schweizer, B. and A. Sklar [1983], Probabilistic Metric Spaces, *Probability and Applied Mathematics*, North-Holland, New York

[133] Schweizer, B. and E. Wolff [1981], On nonparametric measures of dependence for random variables, *Annals of Statistics*, **9**, 879-885

[134] Shaked, M. and J.G. Shanthikumar [1987], The multivariate hazard construction, *Stochastic Processes and their Applications*, **24**, 241-258

[135] Shih, J.H. [1998], A goodness-of-fit test for association in a bivariate survival model, *Biometrika*, **85(1)**, 189-200

[136] Shih, J.H. and T.A. Louis [1995], Inferences on the asociation parameter in copula models for bivariate survival data, *Biometrics*, **51**, 1384-1399

[137] Shorack, G.R. and J.A. Wellner [1986], Empirical Processes with Applications to Statistics, John Wiley & Sons, New York

[138] Sklar, A. [1959], Fonctions de repartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229-231

[139] Sklar, A. [1996], Random variables, distribution functions and copulas — A personal look backward and forward, in L. Rüschendorf, B. Schweizer and M.D. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics*, Institute of Mathematical Statistics, Hayward, CA

[140] Song, P. X-K. [2000], Multivariate dispersion models generated from Gaussian copula, *Scandinavian Journal of Statistics,* **27(2)**, 305-320

[141] Tanner, M.A. [1996], Tools for Statistical Inference — Methods for the Exploration of Posterior Distributions and Likelihood Functions, *Springer Series in Statistics*, *third edition*, Springer-Verlag, New York (first edition 1991)

[142] Tchen, A.H. [1980], Inequalities for distributions with given marginals, *Annals of Statistics*, **8(4)**, 814-827

[143] Tsiatis, A. [1975], A nonidentifiability aspect of the problem of competing risks, *Proceedings of the National Academy of Sciences of the USA*, **72**, 20-22

[144] Valdez, E.A. [2000], Bivariate analysis of survivorship and persistency, Nanyang Technological University, *Working Paper*

[145] Van den Berg, G.J. [1997], Association measures for durations in bivariate hazard rate models, *Journal of Econometrics*, **79(2)**, 221-245

[146] van der Vaart, A.W. and J.A. Wellner [1996], Weak convergence and empirical processes, *Springer Series in Statistics*, Springer-Verlag, New York

[147] Vaupel, J.W., K.G. Manton and E. Stallard [1979], The impact of heterogeneity in individual frailty and the dynamics of mortality, *Demography*, **16**, 439-454

[148] Wang, S. [1998], Aggregation of correlated risk portfolios: models and algorithms, *Proceedings of the Casualty Actuarial Society*, **LXXXV**, 848-939

[149] Whitt, W. [1976], Bivariate distributions with given marginals, *Annals of Statistics*, **4(6)**, 1280-1289

[150] Zaik, E., J. Walter, G. Kelling and C. James [1996], RAROC at Bank of America: from theory to practice, *Journal of Applied Corporate Finance*, **9**, 83-93

[151] Zheng, M. and J.P. Klein [1994], A self-consistent estimator of marginal survival functions based on dependent competing risk data and an assumed copula, *Communications in Statistics – Theory and Methods*, **23(8)**, 2299-2311

[152] Zheng, M. and J.P. Klein [1995], Estimates of marginal survival for dependent competing risks based on an assumed copula, *Biometrika*, **82(1)**, 127-138

A Technical proofs

A.1 Proof of Theorem 2

In order to prove Theorem 2, we first need a lemma:

Lemma 24 *Let us consider N events $A_1, \ldots, A_N$. We note $E = \bigcap_{n=1}^{N} A_n$ the joint event. We have*

$$\Pr\left\{\overline{E}\right\} = \sum_{n=1}^{N} \left\{ (-1)^{n+1} \sum_{\{i_1,\ldots,i_n\} \in \{1,\ldots,N\}} \Pr\left\{\overline{A_{i_1}}, \ldots, \overline{A_{i_n}}\right\} \right\} \tag{228}$$

Proof. The following proof is adapted from Feller [1968]. In order to verify this formula, we will check if all events appear once and only once in the following formula:

$$\begin{aligned}\Pr\left\{\overline{E}\right\} &= \sum \Pr\left\{\overline{A_{i_1}}\right\} - \sum \Pr\left\{\overline{A_{i_1}}, \overline{A_{i_2}}\right\} + \sum \Pr\left\{\overline{A_{i_1}}, \overline{A_{i_2}}, \overline{A_{i_3}}\right\} - \ldots + \\ &\quad (-1)^{N+1} \sum \Pr\left\{\overline{A_{i_1}}, \ldots, \overline{A_{i_N}}\right\}\end{aligned} \tag{229}$$

How many times do we number the probabilities so that the event $\bigcap_{i=i_1}^{i_n} \overline{A_i}$ occurs but not the other events $\overline{A_{i_{n+1}}}, \ldots, \overline{A_{i_N}}$. The answer is the following: $\binom{n}{1}$ in the probabilities $\Pr\left\{\overline{A_{i_1}}\right\}$, $-\binom{n}{2}$ in the probabilities $\Pr\left\{\overline{A_{i_1}}, \overline{A_{i_2}}\right\}$, etc. Moreover, we remark that

$$\begin{aligned}\sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} &= -\left[\sum_{n=0}^{n} (-1)^{i} \binom{n}{i} - 1\right] \\ &= -\left[(1-1)^n - 1\right] \\ &= 1\end{aligned} \tag{230}$$

This completes the proof. ■

Let $\bar{\mathbf{C}}$ be the joint survival function (Nelsen [1999], p. 28):

$$\bar{\mathbf{C}}(u_1, \ldots, u_N) = \Pr\{U_1 > u_1, \ldots, U_N > u_N\} \tag{231}$$

where $(U_1, \ldots, U_N)$ is a vector of uniform random variables with copula $\mathbf{C}$. We remark that

$$\breve{\mathbf{C}}(\mathbf{S}_1(t_1), \ldots, \mathbf{S}_N(t_N)) = \bar{\mathbf{C}}(1 - \mathbf{S}_1(t_1), \ldots, \mathbf{S}_N(t_N)) \tag{232}$$

To find an expression of $\bar{\mathbf{C}}$, we denote A_n the event $U_n > u_n$. Using the previous lemma, we have

$$\begin{aligned}
\bar{\mathbf{C}}(u_1,\ldots,u_N) &= \Pr\{A_1,\ldots,A_N\} \\
&= 1-\Pr\left\{\overline{A_1,\ldots,A_N}\right\} \\
&= 1-\sum_{n=1}^{N}\left\{(-1)^{n+1}\sum_{\{i_1,\ldots,i_n\}\in\{1,\ldots,N\}}\Pr\left\{\overline{A_{i_1}},\ldots,\overline{A_{i_n}}\right\}\right\} \\
&= 1+\sum_{n=1}^{N}\left\{(-1)^{n}\sum_{\{i_1,\ldots,i_n\}\in\{1,\ldots,N\}}\Pr\left\{\overline{A_{i_1}},\ldots,\overline{A_{i_n}}\right\}\right\} \qquad (233)
\end{aligned}$$

We remark that

$$\begin{aligned}
1 &= \mathbf{C}(1,\ldots,1) \\
\Pr\left\{\overline{A_1}\right\} &= \mathbf{C}(u_1,1,\ldots,1) = u_1 \qquad (234) \\
\Pr\left\{\overline{A_1},\overline{A_2}\right\} &= \mathbf{C}(u_1,u_2,1,\ldots,1) \\
\Pr\left\{\overline{A_1},\ldots,\overline{A_n}\right\} &= \mathbf{C}(u_1,\ldots,u_n,1,\ldots,1) \qquad (235)
\end{aligned}$$

We then retreive the expression (39) page 7.

Let us consider some special cases. If N is equal to one, we have $\bar{\mathbf{C}}(u_1) = \sum_{n=0}^{1}\left[(-1)^n\sum_{\mathbf{v}\in\mathcal{Z}(1-n,1,1)}\mathbf{C}(\mathbf{v})\right] = \mathbf{C}(1)-\mathbf{C}(u_1) = 1-u_1$ and $\breve{\mathbf{C}}(u_1) = u_1$. It comes that $\mathbf{S}(t_1) = \breve{\mathbf{C}}(\mathbf{S}_1(t_1)) = \mathbf{S}_1(t_1)$. Univariate survival distributions can then be derived using the copula framework. If $N=2$, we have $\bar{\mathbf{C}}(u_1,u_2) = \sum_{n=0}^{2}\left[(-1)^n\sum_{\mathbf{v}\in\mathcal{Z}(2-n,2,1)}\mathbf{C}(\mathbf{v})\right] = \mathbf{C}(1,1)-\mathbf{C}(u_1,1)-\mathbf{C}(1,u_2)+\mathbf{C}(u_1,u_2) = 1-u_1-u_2+\mathbf{C}(u_1,u_2)$. It comes that $\breve{\mathbf{C}}(u_1,u_2) = u_1+u_2-1+\mathbf{C}(1-u_1,1-u_2)$. We retreive the well known result. In the case $N=3$, we have $\bar{\mathbf{C}}(u_1,u_2,u_3) = \sum_{n=0}^{3}\left[(-1)^n\sum_{\mathbf{v}\in\mathcal{Z}(3-n,3)}\mathbf{C}(\mathbf{v})\right] = \mathbf{C}(1,1,1)-\mathbf{C}(u_1,1,1)-\mathbf{C}(1,u_2,1)-\mathbf{C}(1,1,u_3)+\mathbf{C}(u_1,u_2,1)+\mathbf{C}(u_1,1,u_3)+\mathbf{C}(1,u_2,u_3)-\mathbf{C}(u_1,u_2,u_3) = 1-u_1-u_2-u_3+\mathbf{C}(u_1,u_2)+\mathbf{C}(u_1,u_3)+\mathbf{C}(u_2,u_3)-\mathbf{C}(u_1,u_2,u_3)$. It comes that

$$\begin{aligned}
\breve{\mathbf{C}}(u_1,u_2,u_3) = {} & u_1+u_2+u_3-2+\mathbf{C}(1-u_1,1-u_2)+\mathbf{C}(1-u_1,1-u_3)+ \\
& \mathbf{C}(1-u_2,1-u_3)-\mathbf{C}(1-u_1,1-u_2,1-u_3) \qquad (236)
\end{aligned}$$

In the case $N=4$, we obtain

$$\begin{aligned}
\bar{\mathbf{C}}(u_1,u_2,u_3,u_4) = {} & \mathbf{1}-u_1-u_2-u_3-u_4+\mathbf{C}(u_1,u_2)+\mathbf{C}(u_1,u_3)+\mathbf{C}(u_1,u_4)+\mathbf{C}(u_2,u_3)+\mathbf{C}(u_2,u_4)+ \\
& \mathbf{C}(u_3,u_4)-\mathbf{C}(u_1,u_2,u_3)-\mathbf{C}(u_1,u_2,u_4)-\mathbf{C}(u_1,u_3,u_4)-\mathbf{C}(u_2,u_3,u_4)+\mathbf{C}(u_1,u_2,u_3,u_4) \qquad (237)
\end{aligned}$$

and

$$\begin{aligned}
\breve{\mathbf{C}}(u_1,u_2,u_3,u_4) = {} & u_1+u_2+u_3+u_4-3+\mathbf{C}(1-u_1,1-u_2)+\mathbf{C}(1-u_1,1-u_3)+\mathbf{C}(1-u_1,1-u_4)+ \\
& \mathbf{C}(1-u_2,1-u_3)+\mathbf{C}(1-u_2,1-u_4)+\mathbf{C}(1-u_3,1-u_4)-\mathbf{C}(1-u_1,1-u_2,1-u_3) \\
& -\mathbf{C}(1-u_1,1-u_2,1-u_4)-\mathbf{C}(1-u_1,1-u_3,1-u_4)-\mathbf{C}(1-u_2,1-u_3,1-u_4) \\
& +\mathbf{C}(1-u_1,1-u_2,1-u_3,1-u_4) \qquad (238)
\end{aligned}$$

In order to prove completely Theorem 2, we have also to show that $\breve{\mathbf{C}}$ is a copula function. We can proceed as the bivariate case, but it is very difficult to verify that $\breve{\mathbf{C}}$ is N-increasing. We first note that the margins of

$\breve{\mathbf{C}}$ are uniform and $\breve{\mathbf{C}}$ is a grounded function. We assume that $\mathbf{C}$ is absolutely continuous. It comes that

$$\partial_{1,2,\ldots,N}\breve{\mathbf{C}}\left(u_1,,\ldots,u_N\right) = \left(-1\right)^N \partial_{1,2,\ldots,N}\mathbf{C}\left(1-u_1,,\ldots,1-u_N\right) \times \underbrace{\left(-1\right)\times\ldots\times\left(-1\right)}_{N\text{ times}} \tag{239}$$

We deduce that $\breve{\mathbf{C}}$ is a positive measure, and so $\breve{\mathbf{C}}$ is a copula function. In the case where $\mathbf{C}$ is not absolutely continuous, we assume that the N-increasing property holds.

A.2 Proof of Theorem 3

Using Lemma 24 and denoting A_n the event $U_n \leq u_n$, we obtain

$$\begin{aligned}
\mathbf{C}\left(u_1,\ldots,u_N\right) &= \Pr\left\{A_1,\ldots,A_N\right\} \\
&= 1+\sum_{n=1}^{N}\left\{\left(-1\right)^n \sum_{\{i_1,\ldots,i_n\}\in\{1,\ldots,N\}} \Pr\left\{\overline{A_{i_1}},\ldots,\overline{A_{i_n}}\right\}\right\} \\
&= \sum_{n=0}^{N}\left[\left(-1\right)^n \sum_{\mathbf{v}(u_1,\ldots,u_n,\ldots,u_N)\in\mathcal{Z}(N-n,N,0)} \bar{\mathbf{C}}\left(v_1,\ldots,v_n,\ldots,v_N\right)\right] \\
&= \sum_{n=0}^{N}\left[\left(-1\right)^n \sum_{\mathbf{v}(u_1,\ldots,u_n,\ldots,u_N)\in\mathcal{Z}(N-n,N,0)} \breve{\mathbf{C}}\left(1-v_1,\ldots,1-v_n,\ldots,1-v_N\right)\right] \qquad (240) \\
&= (*)
\end{aligned}$$

If $N=1$, we have

$$\begin{aligned}
(*) &= \breve{\mathbf{C}}\left(1\right) - \breve{\mathbf{C}}\left(1-u_1\right) \\
&= \bar{\mathbf{C}}\left(0\right) - \bar{\mathbf{C}}\left(u_1\right) \\
&= 1-\left(1-u_1\right) \\
&= u_1
\end{aligned} \tag{241}$$

If $N=2$, we verify that

$$\begin{aligned}
(*) &= \breve{\mathbf{C}}\left(1,1\right) - \breve{\mathbf{C}}\left(1-u_1\right) - \breve{\mathbf{C}}\left(1-u_2\right) + \breve{\mathbf{C}}\left(1-u_1,1-u_2\right) \\
&= 1-\left(1-u_1\right)-\left(1-u_2\right)+\left[1-u_1-u_2+\mathbf{C}\left(u_1,u_2\right)\right] \\
&= \mathbf{C}\left(u_1,u_2\right)
\end{aligned} \tag{242}$$

In the case $N=3$, it comes that

$$\begin{aligned}
(*) &= 1-\left(1-u_1\right)-\left(1-u_2\right)-\left(1-u_3\right)+1-u_1-u_2+\mathbf{C}\left(u_1,u_2\right)+ \\
&\quad 1-u_1-u_3+\mathbf{C}\left(u_1,u_3\right)+1-u_2-u_3+\mathbf{C}\left(u_2,u_3\right)- \\
&\quad \left[1-u_1-u_2-u_3+\mathbf{C}\left(u_1,u_2\right)+\mathbf{C}\left(u_1,u_3\right)+\mathbf{C}\left(u_2,u_3\right)-\mathbf{C}\left(u_1,u_2,u_3\right)\right] \\
&= \mathbf{C}\left(u_1,u_2,u_3\right)
\end{aligned} \tag{243}$$

In the case $N=4$, the relationship is verified too.

A.3 Proof of Theorem 23

We have

$$\begin{aligned}
\mathbf{F}_{n:N}(t) &= \Pr\{T_{n:N} \leq t\} \\
&= \Pr\{\text{at least } n \text{ of the survival times } T_1, \ldots, T_N \text{ are smaller than } t\} \\
&= \mathbf{F}_{n+1:N}(t) + \Pr\{\text{exactly } n \text{ of the survival times } T_1, \ldots, T_N \text{ are smaller than } t\} \qquad (244)
\end{aligned}$$

In the same way, we obtain

$$\mathbf{S}_{n+1:N}(t) = \mathbf{S}_{n:N}(t) + \Pr\{\text{exactly } n \text{ of the survival times } T_1, \ldots, T_N \text{ are smaller than } t\} \qquad (245)$$

We can use the previous relationships to determine a recurrence formula

$$[\mathbf{S}_{n+1:N}(t) - \mathbf{S}_{n:N}(t)] = a(n; N) - b(n; N)[\mathbf{S}_{n:N}(t) - \mathbf{S}_{n-1:N}(t)] \qquad (246)$$

However, using this equation to prove Theorem 23 is long and tedious. Another way to proceed is to remark that we can write $\mathbf{F}_{n:N}(t)$ in the following manner

$$\mathbf{F}_{n:N}(t) = \sum_{k=n}^{N} \left[\alpha(k, n; N) \sum_{\mathbf{v}(\mathbf{F}_1(t), \ldots, \mathbf{F}_N(t)) \in \mathcal{Z}(N-k, N, 1)} \mathbf{C}(v_1, \ldots, v_N) \right] \qquad (247)$$

Moreover, the coefficients $\alpha(k, n; N)$ do not depend on the copula function and the margins. Let us then consider the case where the copula function is the product copula $\mathbf{C}^{\perp}$ and the survival times are *i.i.d.* with distribution $\mathbf{F}$. In this case, we know that $\mathbf{F}_{n:N}$ is the tail probability of a binomial distribution (David [1970])

$$\mathbf{F}_{n:N}(t) = \sum_{k=n}^{N} \binom{N}{k} \mathbf{F}^k(t) [1 - \mathbf{F}(t)]^{N-k} \qquad (248)$$

It comes that

$$\begin{aligned}
\mathbf{F}_{n:N}(t) &= \mathbf{F}_{n+1:N}(t) + \binom{N}{n} \mathbf{F}^n(t) [1 - \mathbf{F}(t)]^{N-n} \\
&= \mathbf{F}_{n+1:N}(t) + \binom{N}{n} \mathbf{F}^n(t) \sum_{k=0}^{N-n} (-1)^k \binom{N-n}{k} \mathbf{F}^k(t) \\
&= \mathbf{F}_{n+1:N}(t) + \sum_{k=0}^{N-n} (-1)^k \binom{N}{n} \binom{N-n}{k} \mathbf{F}^{k+n}(t) \\
&= \mathbf{F}_{n+1:N}(t) + \sum_{k=n}^{N} (-1)^{k-n} \binom{N}{n} \binom{N-n}{k-n} \mathbf{F}^k(t) \\
&= \sum_{l=n}^{N} \sum_{k=l}^{N} (-1)^{k-l} \binom{N}{l} \binom{N-l}{k-l} \mathbf{F}^k(t) \\
&= \sum_{k=n}^{N} \sum_{l=n}^{k} (-1)^{k-l} \binom{N}{l} \binom{N-l}{k-l} \mathbf{F}^k(t) \qquad (249)
\end{aligned}$$

Using the equation (247), we have also

$$\mathbf{F}_{n:N}(t) = \sum_{k=n}^{N} \alpha(k, n; N) \binom{N}{k} \mathbf{F}^k(t) \qquad (250)$$

We deduce that

$$\begin{aligned}
\alpha(k,n;N) &= \frac{\sum_{l=n}^{k}(-1)^{k-l}\binom{N}{l}\binom{N-l}{k-l}}{\binom{N}{k}} \\
&= \sum_{l=n}^{k}(-1)^{k-l}\frac{k!}{(k-l)!l!} \\
&= \sum_{l=n}^{k}(-1)^{k-l}\binom{k}{l}
\end{aligned} \tag{251}$$

We consider now special cases:

- $N = 1$

$$\mathbf{F}_{1:1}(t) = \mathbf{F}_1(t)$$

- $N = 2$

$$\begin{aligned}
\mathbf{F}_{2:2}(t) &= \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t)) \\
\mathbf{F}_{1:2}(t) &= \mathbf{F}_1(t) + \mathbf{F}_2(t) - \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t))
\end{aligned}$$

- $N = 3$

$$\begin{aligned}
\mathbf{F}_{3:3}(t) &= \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t)) \\
\mathbf{F}_{2:3}(t) &= \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t)) + \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_3(t)) + \mathbf{C}(\mathbf{F}_2(t), \mathbf{F}_3(t)) - 2\cdot\mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t)) \\
\mathbf{F}_{1:3}(t) &= \mathbf{F}_1(t) + \mathbf{F}_2(t) + \mathbf{F}_3(t) - \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t)) - \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_3(t)) - \mathbf{C}(\mathbf{F}_2(t), \mathbf{F}_3(t)) + \\
&\quad \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t))
\end{aligned}$$

- $N = 4$

$$\begin{aligned}
\mathbf{F}_{4:4}(t) &= \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t)) \\
\mathbf{F}_{3:4}(t) &= \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t)) + \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_4(t)) + \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_3(t), \mathbf{F}_4(t)) + \\
&\quad \mathbf{C}(\mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t)) - 3\cdot\mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t)) \\
\mathbf{F}_{2:4}(t) &= \sum_{j>i}\mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t)) - 2\sum_{k>j>i}\mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t)) + 3\cdot\mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t)) \\
\mathbf{F}_{1:4}(t) &= \sum_{i}\mathbf{F}_i(t) - \sum_{j>i}\mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t)) + \sum_{k>j>i}\mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t)) - \\
&\quad \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t))
\end{aligned}$$

- $N = 5$

$$\begin{aligned}
\mathbf{F}_{5:5}(t) &= \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t), \mathbf{F}_5(t)) \\
\mathbf{F}_{4:5}(t) &= \sum_{l>k>j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t), \mathbf{F}_l(t)) - 4 \cdot \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t), \mathbf{F}_5(t)) \\
\mathbf{F}_{3:5}(t) &= \sum_{k>j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t)) - 3 \sum_{l>k>j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t), \mathbf{F}_l(t)) + \\
&\quad 6 \cdot \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t), \mathbf{F}_5(t)) \\
\mathbf{F}_{2:5}(t) &= \sum_{j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t)) - 2 \sum_{k>j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t)) + \\
&\quad 3 \sum_{l>k>j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t), \mathbf{F}_l(t)) + 4 \cdot \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t), \mathbf{F}_5(t)) \\
\mathbf{F}_{1:5}(t) &= \sum_{i} \mathbf{F}_i(t) - \sum_{j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t)) + \sum_{k>j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t)) - \\
&\quad \sum_{l>k>j>i} \mathbf{C}(\mathbf{F}_i(t), \mathbf{F}_j(t), \mathbf{F}_k(t), \mathbf{F}_l(t)) + \mathbf{C}(\mathbf{F}_1(t), \mathbf{F}_2(t), \mathbf{F}_3(t), \mathbf{F}_4(t), \mathbf{F}_5(t))
\end{aligned}$$

B Conditional distribution in the Normal copula

Let us consider the $N-$copula function $\mathbf{C}(\mathbf{u})$. We partition the vector $\mathbf{u}$ at the nth row to give

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ u_{n-1} \\ \hline u_n \\ \vdots \\ u_N \end{bmatrix} \tag{252}$$

The conditional distribution of U_1 given U_2 is equal to $\mathbf{u}_2$ is given by the Bayes Theorem:

$$\begin{aligned}
\mathbf{C}_{1|2}(\mathbf{u}) &= \Pr\{U_1 \leq \mathbf{u}_1 \mid U_2 = \mathbf{u}_2\} \\
&= \frac{\Pr\{U_1 \leq \mathbf{u}_1, U_2 = \mathbf{u}_2\}}{\Pr\{U_2 = \mathbf{u}_2\}}
\end{aligned} \tag{253}$$

Another expression of $\mathbf{C}_{1|2}$ is

$$\begin{aligned}
\mathbf{C}_{1|2}(\mathbf{u}) &= \lim_{\mathbf{h}_2 \to \mathbf{0}} \frac{\Pr\{U_1 \leq \mathbf{u}_1, \mathbf{u}_2 \leq U_2 \leq \mathbf{u}_2 + \mathbf{h}_2\}}{\Pr\{\mathbf{u}_2 \leq U_2 \leq \mathbf{u}_2 + \mathbf{h}_2\}} \\
&= \lim_{\mathbf{h}_2 \to \mathbf{0}} \frac{\mathbf{C}(\mathbf{u}_1, \mathbf{u}_2 + \mathbf{h}_2) - \mathbf{C}(\mathbf{u}_1, \mathbf{u}_2)}{\mathbf{C}(\mathbf{1}, \mathbf{u}_2 + \mathbf{h}_2) - \mathbf{C}(\mathbf{1}, \mathbf{u}_2)} \\
&= \frac{\partial_{n,\ldots,N} \mathbf{C}(\mathbf{u}_1, \mathbf{u}_2)}{\partial_{n,\ldots,N} \mathbf{C}(\mathbf{1}, \mathbf{u}_2)}
\end{aligned} \tag{254}$$

In the case where $n = N$, we retreive the result

$$\Pr\{U_1 \leq u_1, \ldots, U_{N-1} \leq u_{N-1} \mid U_N = u_N\} = \partial_N \mathbf{C}(u_1, \ldots, u_N) \tag{255}$$

The relationship between the conditional distributions $\mathbf{C}_{1|2}$ and $\mathbf{F}_{1|2}$ is given by the following equation

$$\mathbf{F}_{1|2}(x_1,\ldots,x_N)=\mathbf{C}_{1|2}(\mathbf{F}_1(x_1),\ldots,\mathbf{F}_N(x_N)) \tag{256}$$

To prove this result, we remark that

$$\partial_{n,\ldots,N}\mathbf{F}(\mathbf{x}_1,\mathbf{x}_2)=\mathbf{F}_{1|2}(\mathbf{x}_1,\mathbf{x}_2)\,\partial_{n,\ldots,N}\mathbf{F}(+\infty,\mathbf{x}_2) \tag{257}$$

Using the fact that $\mathbf{F}(x_1,\ldots,x_N)=\mathbf{C}(\mathbf{F}_1(x_1),\ldots,\mathbf{F}_N(x_N))$, it comes that

$$\begin{aligned}\mathbf{F}_{1|2}(x_1,\ldots,x_N) &= \frac{\partial_{n,\ldots,N}\mathbf{C}(\mathbf{F}_1(x_1),\ldots,\mathbf{F}_N(x_N))\prod_{i=n}^{N}f_i(x_i)}{\partial_{n,\ldots,N}\mathbf{C}(\mathbf{F}_1(+\infty),\ldots,\mathbf{F}_{n-1}(+\infty),\mathbf{F}_n(x_n),\ldots,\mathbf{F}_N(x_N))\prod_{i=n}^{N}f_i(x_i)}\\ &= \frac{\partial_{n,\ldots,N}\mathbf{C}(\mathbf{F}_1(x_1),\ldots,\mathbf{F}_N(x_N))}{\partial_{n,\ldots,N}\mathbf{C}(1,\ldots,1,\mathbf{F}_n(x_n),\ldots,\mathbf{F}_N(x_N))}\end{aligned} \tag{258}$$

In the case of the Normal distribution, we can show that if $X\sim\mathcal{N}(\mathbf{0},\boldsymbol{\rho})$, then the conditional distribution of X_1 given that X_2 is equal to $\mathbf{x}_2$ is normal with mean $\mu_{1|2}$ and covariance $\Sigma_{1|2}$ where[29]

$$\mu_{1|2}=\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\mathbf{x}_2 \tag{259}$$

and

$$\Sigma_{1|2}=\boldsymbol{\rho}_{11}-\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\boldsymbol{\rho}_{12}^{\top} \tag{260}$$

It comes that

$$\mathbf{C}_{1|2}(\mathbf{u}_1,\mathbf{u}_2;\boldsymbol{\rho})=\Phi(\bar{\mathbf{x}}_1;\bar{\boldsymbol{\rho}}) \tag{261}$$

with $\bar{\mathbf{x}}_1=\left[\Phi^{-1}(\mathbf{u}_1)-\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\Phi^{-1}(\mathbf{u}_2)\right]\setminus\boldsymbol{\sigma}$, $\boldsymbol{\rho}_{11}-\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\boldsymbol{\rho}_{12}^{\top}=\boldsymbol{\sigma}\odot\boldsymbol{\sigma}^{\top}\odot\bar{\boldsymbol{\rho}}$ and $\boldsymbol{\sigma}=\operatorname{diag}^{\frac{1}{2}}\left(\boldsymbol{\rho}_{11}-\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\boldsymbol{\rho}_{12}^{\top}\right)$ Note that we can express this conditional probability as a function of the Normal copula of dimension $n-1$ because we have

$$\mathbf{C}_{1|2}(\mathbf{u}_1,\mathbf{u}_2;\boldsymbol{\rho})=\mathbf{C}\left(\Phi\left(\left[\Phi^{-1}(\mathbf{u}_1)-\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\Phi^{-1}(\mathbf{u}_2)\right]\setminus\boldsymbol{\sigma}\right);\left[\boldsymbol{\rho}_{11}-\boldsymbol{\rho}_{12}\boldsymbol{\rho}_{22}^{-1}\boldsymbol{\rho}_{12}^{\top}\right]\setminus\boldsymbol{\sigma}\setminus\boldsymbol{\sigma}^{\top}\right) \tag{262}$$

C Contributions to the likelihood

C.1 The case of censoring

In the bivariate case, we observe $\left(D_1,D_2,\delta_1^-,\delta_2^-,\delta_1^+,\delta_2^+\right)$. Moreover, we assume that the distribution of $\mathbf{X}=\left(D_1,D_2,C_1^-,C_2^-,C_1^+,C_2^+\right)$ is of the form

$$\mathbf{C}^{\perp}\left(\mathbf{C}(\mathbf{F}_1(t_1),\mathbf{F}_2(t_2)),\mathbf{C}^{\perp}\left(\mathbf{G}_1^-\left(c_1^-\right),\mathbf{G}_2^-\left(c_2^-\right),\mathbf{G}_1^+\left(c_1^+\right),\mathbf{G}_2^+\left(c_2^+\right)\right)\right) \tag{263}$$

That's imply that the censoring times are independent, and that the survival times are independent of the censoring times. We could then distinguish the following cases:

- If neither T_1 nor T_2 is censored, we have

$$\begin{aligned}\Pr\{D_1\le d_1,D_2\le d_2\} &\propto \Pr\{T_1\le d_1,T_2\le d_2\}\\ &= 1-\mathbf{S}_1(d_1)-\mathbf{S}_2(d_2)+\breve{\mathbf{C}}(\mathbf{S}_1(d_1),\mathbf{S}_2(d_2))\end{aligned} \tag{264}$$

  and

$$L\propto\breve{c}(\mathbf{S}_1(d_1),\mathbf{S}_2(d_2))\,f_1(d_1)\,f_2(d_2) \tag{265}$$

[29] see Kotz, Balakrishnan and Johnson [2000] page 112.

- If T_1 is right-censored — $\left(\delta_1^-, \delta_2^-, \delta_1^+, \delta_2^+\right) = (0,0,1,0)$, we have

$$\begin{aligned}
\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\} &\propto \Pr\left\{C_1^+ \le d_1, T_2 \le d_2, T_1 > C_1^+\right\} \\
&= \iiint \mathbf{1}_{[c \le d_1, t_2 \le d_2, t_1 > c]} f\left(t_1, t_2\right) g_1^+\left(c\right)\, \mathrm{d}t_1\, \mathrm{d}t_2\, \mathrm{d}c \\
&= \int_0^{d_1} \left[\int_c^{\infty} \int_0^{d_2} f\left(t_1, t_2\right)\, \mathrm{d}t_1\, \mathrm{d}t_2\right] g_1^+\left(c\right)\, \mathrm{d}c \\
&= \int_0^{d_1} \left[\mathbf{S}_1\left(c\right) - \breve{\mathbf{C}}\left(\mathbf{S}_1\left(c\right), \mathbf{S}_2\left(d_2\right)\right)\right] g_1^+\left(c\right)\, \mathrm{d}c \qquad (266)
\end{aligned}$$

and

$$L \propto \partial_2 \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_1\right), \mathbf{S}_2\left(d_2\right)\right) f_2\left(d_2\right) g_1^+\left(d_1\right) \qquad (267)$$

Symetrically, if T_2 is right-censored — $\left(\delta_1^-, \delta_2^-, \delta_1^+, \delta_2^+\right) = (0,0,0,1)$, we have

$$L \propto \partial_1 \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_1\right), \mathbf{S}_2\left(d_2\right)\right) f_1\left(d_1\right) g_2^+\left(d_2\right) \qquad (268)$$

- If T_1 is left-censored — $\left(\delta_1^-, \delta_2^-, \delta_1^+, \delta_2^+\right) = (1,0,0,0)$, we have

$$\begin{aligned}
\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\} &\propto \Pr\left\{C_1^- \le d_1, T_2 \le d_2, T_1 \le C_1^-\right\} \\
&= \iiint \mathbf{1}_{[c \le d_1, t_2 \le d_2, t_1 \le c]} f\left(t_1, t_2\right) g_1^-\left(c\right)\, \mathrm{d}t_1\, \mathrm{d}t_2\, \mathrm{d}c \\
&= \int_0^{d_1} \left[\int_0^{c} \int_0^{d_2} f\left(t_1, t_2\right)\, \mathrm{d}t_1\, \mathrm{d}t_2\right] g_1^-\left(c\right)\, \mathrm{d}c \\
&= \int_0^{d_1} \left[1 - \mathbf{S}_1\left(c\right) - \mathbf{S}_2\left(d_2\right) + \breve{\mathbf{C}}\left(\mathbf{S}_1\left(c\right), \mathbf{S}_2\left(d_2\right)\right)\right] g_1^-\left(c\right)\, \mathrm{d}c \qquad (269)
\end{aligned}$$

and

$$L \propto \left(1 - \partial_2 \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_1\right), \mathbf{S}_2\left(d_2\right)\right)\right) f_2\left(d_2\right) g_1^-\left(d_1\right) \qquad (270)$$

Symetrically, if T_2 is left-censored — $\left(\delta_1^-, \delta_2^-, \delta_1^+, \delta_2^+\right) = (0,0,0,0)$, we have

$$L \propto \left(1 - \partial_1 \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_1\right), \mathbf{S}_2\left(d_2\right)\right)\right) f_1\left(d_1\right) g_2^-\left(d_2\right) \qquad (271)$$

- If T_1 and T_2 are right-censored — $\left(\delta_1^-, \delta_2^-, \delta_1^+, \delta_2^+\right) = (0,0,1,1)$, we have

$$\begin{aligned}
\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\} &\propto \Pr\left\{C_1^+ \le d_1, C_2^+ \le d_2, T_1 > C_1^+, T_2 > C_2^+\right\} \\
&= \iiiint \mathbf{1}_{[c_1 \le d_1, c_2 \le d_2, t_1 > c_1, t_2 > c_2]} f\left(t_1, t_2\right) g_1^+\left(c_1\right) g_2^+\left(c_2\right)\, \mathrm{d}t_1\, \mathrm{d}t_2\, \mathrm{d}c_1\, \mathrm{d}c_2 \\
&= \int_0^{d_1} \int_0^{d_2} \left[\int_{c_1}^{\infty} \int_{c_2}^{\infty} f\left(t_1, t_2\right)\, \mathrm{d}t_1\, \mathrm{d}t_2\right] g_1^+\left(c_1\right) g_2^+\left(c_2\right)\, \mathrm{d}c_1\, \mathrm{d}c_2 \\
&= \int_0^{d_1} \int_0^{d_2} \breve{\mathbf{C}}\left(\mathbf{S}_1\left(c_1\right), \mathbf{S}_2\left(c_2\right)\right) g_1^+\left(c_1\right) g_2^+\left(c_2\right)\, \mathrm{d}c_1\, \mathrm{d}c_2 \qquad (272)
\end{aligned}$$

and

$$L \propto \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_1\right), \mathbf{S}_2\left(d_2\right)\right) g_1^+\left(d_1\right) g_2^+\left(d_2\right) \qquad (273)$$

- If T_1 and T_2 are left-censored — $\left(\delta_1^-, \delta_2^-, \delta_1^+, \delta_2^+\right) = (1, 1, 0, 0)$, we have

$$\begin{aligned}
\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\} &\propto \Pr\left\{C_1^- \le d_1, C_2^- \le d_2, T_1 \le C_1^-, T_2 \le C_2^-\right\} \\
&= \iiiint \mathbf{1}_{[c_1 \le d_1, c_2 \le d_2, t_1 \le c_1, t_2 \le c_2]} f\left(t_1, t_2\right) g_1^-\left(c_1\right) g_2^-\left(c_2\right) \,\mathrm{d}t_1 \,\mathrm{d}t_2 \,\mathrm{d}c_1 \,\mathrm{d}c_2 \\
&= \int_0^{d_1} \int_0^{d_2} \left[\int_0^{c_1} \int_0^{c_2} f\left(t_1, t_2\right) \,\mathrm{d}t_1 \,\mathrm{d}t_2\right] g_1^-\left(c_1\right) g_2^-\left(c_2\right) \,\mathrm{d}c_1 \,\mathrm{d}c_2 \\
&= \int_0^{d_1} \int_0^{d_2} \left[1 - \mathbf{S}_1\left(c_1\right) - \mathbf{S}_2\left(c_2\right) + \breve{\mathbf{C}}\left(\mathbf{S}_1\left(c_1\right), \mathbf{S}_2\left(c_2\right)\right)\right] g_1^-\left(c_1\right) g_2^-\left(c_2\right) \,\mathrm{d}c_1 \,\mathrm{d}c_2
\end{aligned} \tag{274}$$

  and

$$L \propto \left(1 - \mathbf{S}_1\left(d_1\right) - \mathbf{S}_2\left(d_2\right) + \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_1\right), \mathbf{S}_2\left(d_2\right)\right)\right) g_1^-\left(d_1\right) g_2^-\left(d_2\right) \tag{275}$$

- If T_1 is right–censored and T_2 is left-censored — $\left(\delta_1^-, \delta_2^-, \delta_1^+, \delta_2^+\right) = (0, 1, 1, 0)$, we have

$$\begin{aligned}
\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\} &\propto \Pr\left\{C_1^+ \le d_1, C_2^- \le d_2, T_1 > C_1^+, T_2 \le C_2^-\right\} \\
&= \iiiint \mathbf{1}_{[c_1 \le d_1, c_2 \le d_2, t_1 > c_1, t_2 \le c_2]} f\left(t_1, t_2\right) g_1^+\left(c_1\right) g_2^-\left(c_2\right) \,\mathrm{d}t_1 \,\mathrm{d}t_2 \,\mathrm{d}c_1 \,\mathrm{d}c_2 \\
&= \int_0^{d_1} \int_0^{d_2} \left[\int_{c_1}^{\infty} \int_0^{c_2} f\left(t_1, t_2\right) \,\mathrm{d}t_1 \,\mathrm{d}t_2\right] g_1^+\left(c_1\right) g_2^-\left(c_2\right) \,\mathrm{d}c_1 \,\mathrm{d}c_2 \\
&= \int_0^{d_1} \int_0^{d_2} \left[\mathbf{S}_1\left(c_1\right) - \breve{\mathbf{C}}\left(\mathbf{S}_1\left(c_1\right), \mathbf{S}_2\left(c_2\right)\right)\right] g_1^+\left(c_1\right) g_2^-\left(c_2\right) \,\mathrm{d}c_1 \,\mathrm{d}c_2
\end{aligned} \tag{276}$$

  and

$$L \propto \left(\mathbf{S}_1\left(d_1\right) - \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_1\right), \mathbf{S}_2\left(d_2\right)\right)\right) g_1^+\left(d_1\right) g_2^-\left(d_2\right) \tag{277}$$

  Symetrically, if T_1 is left–censored and T_2 is right-censored — $\left(\delta_1^-, \delta_2^-, \delta_1^+, \delta_2^+\right) = (1, 0, 0, 1)$, we have

$$L \propto \left(\mathbf{S}_2\left(d_2\right) - \breve{\mathbf{C}}\left(\mathbf{S}_1\left(d_1\right), \mathbf{S}_2\left(d_2\right)\right)\right) g_1^-\left(d_1\right) g_2^+\left(d_2\right) \tag{278}$$

C.2 The case of left truncation

As for the case of censoring, we distinguish different cases:

- We observe (D_1, D_2) iff $T_1 > Z_1$ and $T_2 > Z_2$ — $(\delta_1^\star, \delta_2^\star) = (1, 1)$. We have

$$\begin{aligned}
\Pr\left\{D_1 \le d_1, D_2 \le d_2 \mid Z_1 = z_1, Z_2 = z_2\right\} &= \frac{\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\}}{\Pr\left\{T_1 > z_1, T_2 > z_2\right\}} \\
&= \frac{\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\}}{\breve{\mathbf{C}}\left(\mathbf{S}_1\left(z_1\right), \mathbf{S}_2\left(z_2\right)\right)}
\end{aligned} \tag{279}$$

- We observe D_1 iff $T_1 > Z_1$ — $(\delta_1^\star, \delta_2^\star) = (1, 0)$. We have

$$\Pr\left\{D_1 \le d_1, D_2 \le d_2 \mid Z_1 = z_1\right\} = \frac{\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\}}{\mathbf{S}_1\left(z_1\right)} \tag{280}$$

  Symetrically, we have

$$\Pr\left\{D_1 \le d_1, D_2 \le d_2 \mid Z_2 = z_2\right\} = \frac{\Pr\left\{D_1 \le d_1, D_2 \le d_2\right\}}{\mathbf{S}_2\left(z_2\right)} \tag{281}$$