Lecture Notes of
Seminario Interdisciplinare di Matematica
Vol. 12(2015), pp. 255 – 271.

Some remarks on multivariate conditional hazard rates and dependence modeling

Fabio Spizzichino[1]

Abstract. We review and discuss some special aspects of the joint distribution of n non-negative random variables $X_1, \cdots, X_n$ in the absolutely continuous case. Our attention will be in particular concentrated on *Multivariate Conditional Hazard Rate* functions and on their role in the description of stochastic dependence among $X_1, \cdots, X_n$. The system of the M.C.H.R. functions is one of the tools that can be used to characterize the related joint probability distribution. Such a characterization is alternative, but equivalent, to the one based on the joint density function (or on other tools such as joint survival function, marginal distributions and survival copula). However these two types of characterizations imply completely different methods for building multivariate dependence models and for describing stochastic dependence among the variables $X_1, \cdots, X_n$. The method of the M.C.H.R. functions is specially adapt to describe *dynamic models* of dependence. As a main purpose of this discussion paper, we recall some basic definitions and aim to provide some related comments. Basic material had mainly appeared in the 1980's in the literature suggested by Reliability problems. Later on, related issues have also been considered, in the analysis of Financial Risk, under a somehow different language and in the perspective of different probability models. Some reflections on the comparison between the two methods, respectively based on the joint density and on the M.C.H.R. functions, are still needed and can reveal to be useful to people working in different applied fields. In particular, we shall dwell on a comparison between two different classes of dependence models: those defined in terms of *conditional independence* and those of the type *Load-Sharing* (a property that can be directly defined in terms of the M.C.H.R. functions). We will also try to briefly explain the reasons of interest in establishing such a comparison.

1. Introduction

As it is well known, the study of probability distributions of vectors of non-negative random variables emerges in several fields of Statistics and of Applied Probability, in a completely natural way. Typically, such random variables have

---

[1]F. Spizzichino, Università degli Studi di Roma "La Sapienza", Dipartimento di Matematica, Piazzale Aldo Moro 5, 00185 Roma, Italy; `fabio.spizzichino@uniroma1.it`

Partially supported by University Sapienza, Rome, through 2014 Research Project "Aspetti di dipendenza stocastica e rischio in modelli probabilistici".

Keywords. Hazard rate, conditional independence, load-sharing model, positive dependence, reliability, portfolio credit risk, default contagion

AMS Subject Classification. 60E00, 62A00, 62H00, 62N00.

the meaning of *waiting times*, *times to events, failure times,* or *times of default* and so on. In fact they are met, under such forms, in Queuing theory, Reliability, Biomedicine, Financial Risk, Actuarial Mathematics. More generally, such a study is useful in all those applied fields where statistical data are observed dynamically, waiting-times for different objects are considered simultaneously, and the phenomenon of time-elapsing takes a central role.

Both the statistical analysis of observed data from non-negative random variables and the description of the joint probability distributions of such variables respectively led mathematicians and statisticians to building special chapters of Probability and Statistics. These chapters are characterized by specific concepts, language and methods of their own.

In the present paper, we review and discuss some special aspects of the joint distribution of n non-negative random variables $X_1, \cdots, X_n$ in the absolutely continuous case. Our attention will be in particular concentrated on *Multivariate Conditional Hazard Rate (M.C.H.R.)* functions and on their role in the description of stochastic dependence among $X_1, \cdots, X_n$.

In the characterization of an absolutely continuous probability distribution on $[0, +\infty)$, namely of the probability distribution of a single non-negative random variable X admitting a probability density, the notion of *hazard rate* function emerges as a natural, and very useful, tool. In the multivariate case, the definition of the M.C.H.R.functions can be seen as a direct generalization of the concept of hazard rate. Of basic importance is the circumstance that a joint probability distribution can be characterized in terms of the system of its M.C.H.R. functions. Such a characterization is alternative, but equivalent, to the characterization of a joint probability distribution, based on other tools such as joint survival function, joint density function, marginal distributions and survival copula.

However these two types of characterizations imply completely different methods for building multivariate dependence models and for describing stochastic dependence among the variables $X_1, \cdots, X_n$.

The method of the M.C.H.R. functions is actually appropriate to describe *dynamic models* of dependence. Or, better, we can say that a multivariate probability model for a vector of non-negative random variables can be called of *dynamic type* whenever its formulation can be given directly in terms of the behavior of the M.C.H.R. functions, rather than in terms of the joint density function or of the connecting copula. Most of this material had appeared in the 1980's in the literature suggested by Reliability problems.

The use of M.C.H.R. functions in particular led to defining the so-called Load-Sharing models. In a few words, the latter are multivariate models, describing the circumstance that the instantaneous risk of a given (still surviving) component only depends on the current time and on the set of all the surviving components. These models have a clear interpretation in the realm of Reliability; for basic papers in the relevant literature see, in particular, [11], [26], [34].

Later on, under a somehow different language, interest in related topics has also emerged in the field of credit risk, in connection with *interacting intensities* and with the theme of *Default Contagion* (see e.g. [20]). On this topic a wide literature has been developed. Several aspects have been considered, also from the conceptual view-point, though by using a terminology different from the one arising in the Reliability field and in the perspective of different probability models.

Interesting historical remarks and a useful overview about M.C.H.R. functions have been provided in the recent article [32]. In the present discussion paper, we aim to present some comments on related points. As one main purpose, we discuss about the comparison between the two methods, respectively based on the joint density and on the M.C.H.R. functions, and about different dependence models corresponding to them. We think in fact that some reflections on these topics is still valuable nowadays and can be useful to people working in different applied fields. Concerning M.C.H.R. functions, we will briefly recall just those aspects that are strictly needed in view of our purposes. Load-Sharing models, even though very special, deserve some room in our discussion and we will in particular dwell on a comparison between them and models of conditional independence.

More precisely, the structure of the paper is as follows. In Section 2, after introducing some needed notation and useful preliminaries, we recall the definition of the M.C.H.R. functions. In Section 3 this concept will be illustrated by considering the special cases of Conditional Independence, Load-Sharing, and Exchangeability. Some basic mathematical aspects of the M.C.H.R. functions will be briefly recalled in Section 4. Section 5 will be finally devoted to a discussion and, in particular, to specific remarks that can constitute a basis for further reflections about stochastic dependence.

2. Notation, preliminaries and basic definitions

For a non-negative random variable X, with an absolutely continuous probability distribution, let G and g respectively denote its distribution function and probability density function. We furthermore denote by $\overline{G}$ its *survival function* and by r its *hazard rate function*, namely

$$\overline{G}(x) := 1 - G(x) = \mathbb{P}\left(X > x\right) \quad , \quad r(x) := \frac{g(x)}{\overline{G}(x)} \ ,$$

for $x > 0$. In the case of the exponential distribution of parameter θ, one has

$$\overline{G}(x) = \exp\{-\theta x\}, \qquad r(x) = \theta \ .$$

In the general case one can write $\overline{G}(x) = \exp\{-\int_0^x r\left(\xi\right)\,\mathrm{d}\xi\}$, or

$$\overline{G}(x) = \exp\{-R(x)\} \quad , \quad g(x) = r\left(\xi\right)\exp\{-R(x)\} \ , \tag{1}$$

denoting by $R(x) = \int_0^x r\left(\xi\right)\,\mathrm{d}\xi$ the *function of integrated hazard.*

All along this paper we will consider n non-negative random variables $X_1, \cdots, X_n$ defined on a same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A standard tool to describe their joint probability distribution is provided by the joint survival function $\overline{F} : [0, +\infty)^n \to [0, 1]$, defined by the position

$$\overline{F}\left(x_1, \cdots, x_n\right) := \mathbb{P}\left(X_1 > x_1, \cdots, X_n > x_n\right) \ .$$

By $\overline{G}_j : [0, +\infty) \to [0, 1]$ $(j = 1, \cdots, n)$ we denote the marginal survival function of X_j: for $x > 0$,

$$\overline{G}_j(x) := \mathbb{P}\left(X_j > x\right) = \overline{F}\left(0, \cdots 0, x, 0, \cdots, 0\right) \ ,$$

where $(0, \cdots 0, x, 0, \cdots, 0)$ is the vector with all the coordinates, but the j-th one, equal to 0 .

For simplicity's sake, we assume the survival functions $\overline{G}_j$'s to be everywhere positive, continuous and strictly decreasing all over $[0,+\infty)$. As a substantial assumption, on the other hand, we also assume $\overline{F}$ to be absolutely continuous, namely we assume the existence of a joint density function $f_{\boldsymbol{X}}(x_1,\cdots,x_n)$, so that $\overline{F}$ can be written in the form

$$\overline{F}(x_1,\cdots,x_n)=\int_{x_1}^{+\infty}\cdots\int_{x_n}^{+\infty}f_{\boldsymbol{X}}(\xi_1,\cdots,\xi_n)\,\mathrm{d}\xi_1\cdots\mathrm{d}\xi_n\ .$$

The *survival copula* is the function $K:[0,1]^n\to[0,1]$ defined, for $\boldsymbol{u}\equiv(u_1,\cdots,u_n)\in[0,1]^n$ by the position

$$K(\boldsymbol{u})=\overline{F}\left(\overline{G}_1^{-1}(u_1),\cdots,\overline{G}_n^{-1}(u_n)\right)\ .$$

We can thus also write

$$\overline{F}(x_1,\cdots,x_n)=K\left(\overline{G}_1(x_1),\cdots,\overline{G}_n(x_n)\right)\ .$$

The function K is easily seen to be a copula and it provides a tool to express stochastic dependence among $X_1,\cdots,X_n$. For these aspects, for the relation existing between the survival copula and the *connecting copula*, and for general properties of the concept of copula, see, e.g. basic references such as [23] and [13]; see also the very recent volume [9] and the list of references cited therein.

From now on we will simply write f in place of $f_{\boldsymbol{X}}$. The index will rather be added to the symbol of a joint density function only when other random vectors are considered.

Let us now come to the definition of M.C.H.R. functions. In a sense, such a definition can be seen as a direct extension of the univariate concept of the hazard rate function for a single non-negative random variable X. We need however to consider the vector $X_{(1)},\cdots,X_{(n)}$ of the order statistics of $X_1,\cdots,X_n$ and the assumption of absolute continuity of the joint survival function $\overline{F}$ becomes essential. Such an assumption, in particular, guarantees the condition

$$\mathbb{P}(X_1\neq X_2\neq\cdots\neq X_n)=1\ , \tag{2}$$

so that the following definition of a random permutation $(J(1),\cdots,J(n))$ of $[n]:=\{1,\cdots,n\}$ is well-posed: for $h=1,\cdots,n$, we set

$$X_{(h)}=X_{J(h)}\ .$$

We also set

$$Y_{\emptyset}:=X_{(1)}=\min_{1\leq j\leq n}X_j$$

and, for non-empty and complementary subsets $A,\widetilde{A}\subset[n]$,

$$Y_A=Y_A(X_1,\cdots X_n):=\min_{j\in\widetilde{A}}X_j\ .$$

For a fixed index $j\in[n]$, an ordered set $I=\{i_1,\cdots,i_k\}\subset[n]$ with $j\notin I$, and an ordered vector $0\leq t_1\leq\cdots\leq t_k$ the M.C.H.R. function $\lambda_j(t|I;t_1,\cdots,t_k)$ is defined as follows:

$$\lambda_j(t|I;t_1,\cdots,t_k):=\lim_{\Delta t\to 0}\frac{1}{\Delta t}P\{X_j\leq t+\Delta t|X_{i_1}=t_1,\cdots,X_{i_k}=t_k,Y_I>t\}\ . \tag{3}$$

$\lambda_j(t|I;t_1,\cdots,t_k)$ is then a hazard rate function associated to the conditional distribution of the variable X_j, given the observation of the *dynamic history*

(4) $$\boldsymbol{h}_t =: \{X_{i_1}=t_1,\cdots,X_{i_k}=t_k,Y_I>t\} .$$

Furthermore, one sets

(5) $$\lambda_j(t|\emptyset) := \lim_{\Delta t\to 0}\frac{1}{\Delta t}P\{X_j\le t+\Delta t|Y_\emptyset>t\} .$$

The conditional probabilities appearing in the r.h.s. of formula (3) make of course sense in view of the above assumption of absolute continuity.

The M.C.H.R. functions, on the other hand, are related to the more general concepts of *stochastic intensity* and *compensator*, dealt with in the theory of point processes (see in particular [12], [5], [1]). See also e.g. [16], [2], [20] and other references cited therein.

The M.C.H.R. functions can be obtained in terms of the joint probability density $f(x_1,\cdots,x_n)$. Vice-versa a formula to obtain $f(x_1,\cdots,x_n)$ starting from the knowledge of all the functions $\lambda_j(t|I;t_1,\cdots,t_k),\lambda_j(t|\emptyset)$ can also be written. This topic will be considered more in details in Section 4. The next Section will be devoted to demonstrate the above definition by focusing attention on some relevant special classes of probability distributions.

3. Special classes of dependence models

In this section we consider three relevant classes of probability models for the n-tuple $X_1,\cdots,X_n$. These classes are respectively defined by Conditional Independence, Load-Sharing Dependence, Symmetric Dependence or Exchangeability. For such models, we in particular show the special forms of the joint density functions and of the system of the M.C.H.R. functions. Some further specific aspects will be analyzed in the three different cases.

3.1. Conditional independence.

Defined on the probability space $(\Omega,\mathcal{F},\mathbb{P})$ let us have, for some $d=1,2,\cdots$, a d-dimensional random vector $\Theta:\Omega\to\mathbb{R}^d$ with joint probability density $\pi_0(\theta)$ and such that the joint distribution of the vector $(\boldsymbol{X},\Theta)$ is also absolutely continuous. We now focus our attention on the specific case when $X_1,\cdots,X_n$ are *conditionally independent* given Θ. Denote by $g_j(\cdot;\theta),\overline{G}_j(\cdot;\theta)$, and $r_j(\cdot;\theta)$ the conditional density, the conditional survival function, and the conditional hazard rate of X_j given $(\Theta=\theta)$, respectively.

We can immediately write the joint density $f(x_1,\cdots,x_n)$ in the form

$$f(x_1,\cdots,x_n)=\int_{\mathbb{R}^d}\prod_{j=1}^n g_j(x_j;\theta)\pi_0(\theta)\,\mathrm{d}\theta .$$

For what concerns the computation of the corresponding system of the M.C.H.R. functions, one can argue as explained below. First, for clarity's sake, it is convenient to preliminary focus attention on the following univariate setting. Let X be a scalar random variable on $(\Omega,\mathcal{F},\mathbb{P})$ such that the joint distribution of (X,Θ) is absolutely continuous and denote by $g(\cdot;\theta),\overline{G}(\cdot;\theta)$, and $r(\cdot;\theta)$ the conditional density, the conditional survival function, and the conditional hazard rate of X given $(\Theta=\theta)$, respectively. As to the marginal density function and survival function of X we can respectively write

$$g(x) = \int_{\mathbb{R}^d} g(x;\theta)\,\pi_0(\theta)\;\mathrm{d}\theta \quad , \quad \overline{G}(x) = \int_{\mathbb{R}^d} \overline{G}(x;\theta)\pi_0(\theta)\;\mathrm{d}\theta\ .$$

As to the hazard rate of X, we then obtain

$$r(x) = \frac{\int_{\mathbb{R}^d} g(x|\theta)\,\pi_0(\theta)\;\mathrm{d}\theta}{\int_{\mathbb{R}^d} \overline{G}(x|\theta)\pi_0(\theta)\;\mathrm{d}\theta}\ ,$$

that we rewrite

$$r(x) = \frac{\int_{\mathbb{R}^d} \left(\frac{g(x;\theta)}{\overline{G}(x;\theta)}\right) \overline{G}(x;\theta)\pi_0(\theta)\;\mathrm{d}\theta}{\int_{\mathbb{R}^d} \overline{G}(x;\theta)\pi_0(\theta)\;\mathrm{d}\theta}\ .$$

By the Bayes formula, the function

$$\frac{\overline{G}(x;\theta)\pi_0(\theta)}{\int_{\mathbb{R}^d} \overline{G}(x;\theta)\pi_0(\theta)\;\mathrm{d}\theta}$$

can be seen as the conditional density of Θ given the event $(X > x)$ and will be denoted by $\pi_t(\theta|X > x)$. In conclusion, we can interpret the hazard rate $r(x)$ in the form

$$r(x) = \int_{\mathbb{R}^d} r(x;\theta)\,\pi_t(\theta|X > x)\;\mathrm{d}\theta\ . \tag{6}$$

Let us now pass to the computation of the functions $\lambda_j(t|\emptyset)$. Fix $j \in [n]$ and consider the conditional density $\pi_t\left(\theta\middle|\bigcap_{i\neq j}(X_i > t)\right)$ of Θ given the event $\left(\bigcap_{i\neq j}(X_i > t)\right)$. Notice that $\lambda_j(t|\emptyset)$ can be seen as the value at t, of the hazard rate function of the conditional distribution of the variable X_j, given the event $\left(\bigcap_{i\neq j}(X_i > t)\right)$ and that we can write

$$\lambda_j(t|\emptyset) = \frac{g_j\left(t\middle|\bigcap_{i\neq j}(X_i > t)\right)}{\overline{G}_j\left(t\middle|\bigcap_{i\neq j}(X_i > t)\right)}\ .$$

Then, in view of the assumption of conditional independence,

$$\lambda_j(t|\emptyset) = \frac{\int_{\mathbb{R}^d} g_j(t;\theta)\pi_t\left(\theta\middle|\bigcap_{i\neq j}(X_i > t)\right)\;\mathrm{d}\theta}{\int_{\mathbb{R}^d} \overline{G}_j(t;\theta)\,\pi_t\left(\theta\middle|\bigcap_{i\neq j}(X_i > t)\right)\;\mathrm{d}\theta}\ .$$

Similarly to the conclusion in (6), we obtain

$$\lambda_j(t|\emptyset) = \int_{\mathbb{R}^d} r_j(t;\theta)\pi_t\left(\theta\middle|\bigcap_{i=1}^{n}(X_i > t)\right)\;\mathrm{d}\theta\ ,$$

or, in other words,

$$\lambda_j(t|\emptyset) = \mathbb{E}\left(r_j\left(t;\Theta\right)|Y_\emptyset > t\right) . \tag{7}$$

By using similar arguments, one can give $\lambda_j(t|I;t_1,\cdots,t_k)$ the expression of a conditional expectation as follows:

$$\lambda_j(t|I;t_1,\cdots,t_k) = \mathbb{E}\left(r_j\left(t;\Theta\right)|\boldsymbol{h}_t\right) , \tag{8}$$

where $\boldsymbol{h}_t$ is the *dynamic history* in (4).

In the present subsection, we have assumed the existence of a density function π_0 for the variable Θ just for notational convenience. Actually, this assumption is not at all necessary for the validity of the formulae (7) and (8). What is really essential is the assumption of conditional independence of $X_1,\cdots,X_n$ given Θ, and the existence of the conditional densities $g_j\left(\cdot;\theta\right)$'s.

3.2. **Load-sharing models.** We recall here a class of models which are defined directly in terms of the M.C.H.R. functions and that had originally emerged, in a completely natural way, in the fields of Reliability and Life-testing. These models find direct applications in several other applied fields and the related literature is by now enormously developed.

Definition 1. The random vector $(X_1,\cdots X_n)$ is distributed according to a *Load-Sharing Model* if, for any non-empty set $I\subset[n]$, there exist functions $\mu_j(t|I)$ such that, for all $0\le t_1\le\cdots\le t_{|I|}\le t$,

$$\lambda_j(t|I;t_1,\cdots,t_{|I|}) = \mu_j(t|I) .$$

The Load-Sharing Model is *time-homogeneous* when there exist non-negative numbers $\mu_j(I)$ and $\mu_j(\emptyset)$ such that, for any $t\ge 0$,

$$\mu_j(t|I) = \mu_j(I) \quad ; \quad \lambda_j(t|\emptyset) = \mu_j(\emptyset) . \tag{9}$$

Time-homogeneous Load-Sharing Models manifest a number of relevant aspects that generalize corresponding properties of the distributions of independent (not necessarily identically distributed) exponential lifetimes. Actually they can be seen as multivariate models with the *no-ageing* property. For relations between ageing and dependence, see e.g. [17] and references therein; see also the review paper [37].

We now aim to obtain the functional form of the joint density f of a Load-Sharing Model, in terms of the set of coefficients $\{\mu_j(t|I)\}_{I\subset[n]}$. For simplicity's sake, we will consider the time-homogeneous case. The following Lemmas will be used. First we recall the readers' attention on elementary properties of the minimum among a set of independent, exponential, random variables.

Lemma 2. *Let $Z_1,\cdots,Z_m$ be m independent random variables, exponentially distributed with parameters $\theta_1,\cdots,\theta_m$, respectively, and set $Y=\min_{j\in[m]}Z_j$. Then*

a) *Y has an exponential distribution with parameter $\theta=\sum_{j=1}^m\theta_j$;*

b) $P\left(Y=Z_j\right)=\dfrac{\theta_j}{\sum_{j=1}^{m}\theta_j}$*;*

c) *Y and the partition $\{(Y=Z_1),\cdots,(Y=Z_m)\}$ are stochastically independent.*

Let us now come back to a random vector $\boldsymbol{X} \equiv (X_1, \cdots, X_n)$ whose joint distribution is a time-homogeneous Load-Sharing Model characterized by the parameters $\mu_j(I)$ $(j \notin I \subset [n])$ and $\mu_j(\emptyset)$ $(j \in [n])$.

Lemma 3. *The variable $Y_\emptyset := X_{(1)}$ is exponentially distributed with parameter $\theta_\emptyset := \sum_{j=1}^n \mu_j(\emptyset)$.*

Proof. By definition, one generally has

$$r_{Y_\emptyset}(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathbb{P}\left(\cup_{j=1}^n (X_j \leq t + \Delta t) | Y_\emptyset > t\right) .$$

In view of the condition of absolute continuity among the variables $X_1, \cdots, X_n$, we can on the other hand write

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathbb{P}\left(\cup_{j=1}^n (X_j \leq t + \Delta t) | Y_\emptyset > t\right) =$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \sum\nolimits_{j=1}^n \mathbb{P}\left(X_j \leq t + \Delta t | Y_\emptyset > t\right) ,$$

and thus obtain $r_{Y_\emptyset}(t) = \theta_\emptyset$.

□

Remark 4. For $t > 0$, let $\boldsymbol{t} \equiv (t, \cdots, t)$ denote a vector with $n - k$ coordinates, all equal to t, the number k being suitable for the actual context. Let furthermore $I \equiv \{i_1, \cdots, i_k\} \subset [n]$. We notice that the conditional distribution of $\boldsymbol{X}_{\widetilde{I}} - \boldsymbol{t}$, given the observation of the dynamic history $\boldsymbol{h}_t$ in (4) has again the form of a Load-Sharing Model (of dimension $n - k$) with a new set of parameters. This new set is easily obtained by an appropriate change of the old parameters.

By combining Remark 4 with arguments used in the proof of Lemma 3, one can obtain

Lemma 5. *For $1 \leq k < n$, $I = \{i_1, \cdots, i_k\}$, and $0 < t_1 < \cdots < t_k < t$, the conditional distribution of the variable Y_I given the observation $\{X_{i_1} = t_1, \cdots, X_{i_k} = t_k, Y_I > t\}$ is exponential with parameter*

$$\theta_I := \sum_{j \in \widetilde{I}}^n \mu_j(I) .$$

Proposition 6. *Let $\boldsymbol{x} \equiv (x_1, \cdots, x_n) \in \mathbb{R}_+^n$ be such that $x_{(h)} = x_{j(h)}$. We have*

$$f(\boldsymbol{x}) = \tag{10}$$

$$= \mu_{j(1)}(\emptyset) \exp\left\{-x_{(1)}\theta_\emptyset\right\} \cdot \mu_{j(2)}(\{j(1)\}) \exp\left\{-\left(x_{(2)} - x_{(1)}\right)\theta_{\{j(1)\}}\right\} \cdot \cdots$$

$$\cdots \cdot \mu_{j(n)}(\{j(1), \cdots, j(n-1)\}) \exp\left\{-\left(x_{(n)} - x_{(n-1)}\right)\theta(\{j(1), \cdots, j(n-1)\})\right\} .$$

Proof. By definition of joint density we write

$$f(\boldsymbol{x}) = \lim_{\Delta x_h \to 0} \frac{\mathbb{P}\left(x_{(1)} \leq X_{j(1)} \leq x_{(1)} + \Delta x_1, \cdots, x_{(n)} \leq X_{j(n)} \leq x_{(n)} + \Delta x_n\right)}{\Delta x_1 \cdot \cdots \cdot \Delta x_n}$$

and, by the product formula for probabilities,

$$\mathbb{P}\left(x_{(1)} \leq X_{j(1)} \leq x_{(1)} + \Delta x_1, \cdots, x_{(n)} \leq X_{j(n)} \leq x_{(n)} + \Delta x_n\right) =$$
$$= \mathbb{P}\left(x_{(1)} \leq X_{j(1)} \leq x_{(1)} + \Delta x_1\right) \times$$
$$\times \prod_{i=2}^{n} \mathbb{P}\left(x_{(i)} \leq X_{j(i)} \leq x_{(i)} + \Delta x_i \middle| \cap_{k=1}^{i-1} \left\{x_{(k)} \leq X_{j(k)} \leq x_{(k)} + \Delta x_k\right\}\right) .$$

The conclusion then follows from Lemmas 2, 3, 5 and Remark 4.

□

3.3. **Exchangeability.** We consider here the case when $X_1, \cdots, X_n$ are exchangeable random variables, namely when the vectors $(X_{\pi(1)}, \cdots, X_{\pi(n)})$ and $(X_1, \cdots, X_n)$ are identically distributed, for any permutation $\boldsymbol{\pi} \equiv (\pi(1), \cdots, \pi(n))$ of $\{1, \cdots, n\} \equiv [n]$. In other words, there is a complete symmetry among $X_1, \cdots, X_n$, as far as their joint distribution is concerned. In terms of the joint density function f, this condition writes

$$f\left(x_1, \cdots, x_n\right) = f\left(x_{\pi(1)}, \cdots, x_{\pi(n)}\right) .$$

Such a condition of symmetry yields that the information contained in the observation of a *dynamic history*

$$\boldsymbol{h}_t =: \{X_{i_1} = t_1, \cdots, X_{i_k} = t_k, Y_I > t\}$$

is the same as for the observation

$$\{X_{(1)} = t_1, \cdots, X_{(k)} = t_k, Y_I > t\} .$$

The M.C.H.R. functions take the form

$$\lambda_j(t|I; t_1, \cdots, t_k) = \lambda(t|k; t_1, \cdots, t_k) , \tag{11}$$

$$\lambda_j(t|\emptyset) = \lambda(t|\emptyset) . \tag{12}$$

Typically, the condition of exchangeability arises from conditional independence, when we consider variables $X_1, \cdots, X_n$ which, conditionally on Θ, are (not only independent, but also) identically distributed.

One can obtain the condition of exchangeability for Load-Sharing models, too. This is obtained by letting the functions $\mu_j(t|I)$ to depend on t and on k, the cardinality of the set I and by dropping the dependence on the specific choices of $j \in [n]$ and of $I \subset [n]$.

Example 7. Let $\widehat{L} > 0$ be given. A very special and remarkable case of time-homogeneous, exchangeable, Load Sharing Model is obtained by letting $\lambda_j(t|\emptyset) = \widehat{L}/n$, $\lambda_j(t|I; t_1, \cdots, t_k) = \widehat{L}/(n-k)$. In this model $\widehat{L}$ is interpreted as a constant load, which is equally divided among all the surviving components. By adapting formula (10) to this case, one obtains the following special form of the joint density function for $T_1, \cdots, T_n$

$$f\left(\boldsymbol{x}\right) = \frac{\widehat{L}}{n} \exp\{-x_{(1)}\widehat{L}\} \cdot \frac{\widehat{L}}{n-1} \exp\{-\left(x_{(2)} - x_{(1)}\right)\widehat{L}\} \cdot \cdots \cdot$$
$$\cdots \cdot \widehat{L} \exp\{-\left(x_{(n)} - x_{(n-1)}\right)\widehat{L}\} = \frac{\widehat{L}^n}{n!} \exp\{-x_{(n)}\widehat{L}\} .$$

For further details on the M.C.H.R. functions in the exchangeable case, see [36].

4. General aspects of M.C.H.R. functions

For non-negative random variables $X_1, \cdots, X_n$ with a joint density function $f(x_1, \cdots, x_n)$, in this section we review some general aspects of the corresponding M.C.H.R. functions. From an heuristic point of view, such functions take a significant meaning when $X_1, \cdots, X_n$ are though of as the lifetimes of (stochastically dependent) biologic individuals, that start living at a same time 0, or of industrial devices that start working simultaneously at time 0, and so on. However, from a mathematical viewpoint, the M.C.H.R. functions are in any case defined by means of the formulas (3) and (5).

4.1. **Joint densities in terms of M.C.H.R. functions.** The M.C.H.R. functions associated to $(X_1, \cdots, X_n)$ are determined by the joint density $f(x_1, \cdots, x_n)$. Viceversa $f(x_1, \cdots, x_n)$ is determined by the knowledge of the entire systems of the M.C.H.R. functions. As we have seen in the previous Section, the formula (10) holds in the special case of Time-Homogeneous Load-Sharing Models. In the general case one can write, for $x_1 \leq x_2 \leq \cdots \leq x_n$, the formula

$$f(x_1, \cdots, x_n) = \lambda_1(x_1|\emptyset) \exp\left\{ -\int_0^{x_1} \sum_{j=1}^{n} \lambda_j(t|\emptyset) \right\} \mathrm{d}t \cdot$$

$$\cdot \prod_{h=2}^{n} \lambda_h(x_h|\{1, \cdots, h-1\}; x_1, \cdots, x_{h-1}) \cdot$$

$$\cdot \exp\left\{ -\int_{x_{h-1}}^{x_h} \sum_{j=h}^{n} \lambda_j(t|\{1, \cdots, h-1\}; x_1, \cdots, x_{h-1}) \right\} \mathrm{d}t$$

holds. Similar expressions are valid in the other regions of $\mathbb{R}_+^n$ defined by the conditions $x_{\pi(1)} \leq x_{\pi(2)} \leq \cdots \leq x_{\pi(n)}$, with π any permutation of $[n] = \{1, \cdots, n\}$. It is immediately seen from these formulas that conditions (11) and (12) are not only necessary, but also sufficient for $X_1, \cdots, X_n$ to be exchangeable.

4.2. **The total hazard variables and the total hazard construction.** We remind that, for a non-negative random variable X with an hazard rate function $r(x)$, the survival function and the probability density function $\overline{G}(x)$ and $g(x)$ can respectively be expressed in terms of the integrated hazard $R(x)$ as shown in (1). In view of our assumption that $\overline{G}$ is strictly decreasing all over $[0, \infty)$, $R(x)$ is strictly increasing, and we also consider its inverse function R^{-1}. Let us now define the *integrated hazard* of X as the non-negative random variable $\mathcal{H}$ with

$$\mathcal{H} = R(X) .$$

By considering the $\{0, 1\}$-valued random processes $\{\mathbf{1}_{\{X \leq t\}}\}_{t \geq 0}$ which starts in the state 1 at time $t = 0$ and jumps to 0 at the random time X, one can also write

$$\mathcal{H} = \int_0^\infty \mathbf{1}_{\{X \leq t\}} r(t) \, \mathrm{d}t .$$

The following result is obvious; nevertheless it is important in the present setting.

Proposition 8. *The random variable $\mathcal{H}$ is distributed according to a standard exponential distribution. Let T be a non-negative random variable with a standard exponential distribution. Then $R^{-1}(T)$ has the same probability distribution as X.*

Let us now consider the vector of non-negative random variables $\boldsymbol{X} \equiv (X_1, \cdots, X_n)$. In terms of the M.C.H.R. functions, one can appropriately extend to $\boldsymbol{X}$ the circumstances envisaged in the above Proposition. On this purpose we consider, on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the following random processes:

- for $j = 1, \cdots, n$, the $\{0,1\}$-valued random processes $\{\mathbf{1}_{\{X_j \leq t\}}\}_{t \geq 0}$ which starts in the state 1 at time $t = 0$ and jumps to 0 at the random time X_j;
- the set-valued process $\{\boldsymbol{I}_t\}_{t \geq 0}$ where, for $t > 0$, $\boldsymbol{I}_t \subset [n]$ is the ordered set containing the indices i such that $\mathbf{1}_{\{X_i \leq t\}} = 0$;
- the $\mathbb{R}_+$-valued process $\rho_t^{(j)}$ defined by the position

$$\rho_t^{(j)} = \mathbf{1}_{\{X_j \leq t\}} \lambda_j(t|\boldsymbol{I}_t; X_{(1)}, \cdots, X_{(|I_t|)}) \ .$$

Finally we consider, for $j = 1, \cdots, n$, the random variable

$$\mathcal{H}_j = \int_0^\infty \rho_t^{(j)} \, \mathrm{d}t \ ,$$

the *integrated hazard* of X_j.

Theorem 9. *The random variables $\mathcal{H}_1, \cdots, \mathcal{H}_n$ are independent identically distributed, with a standard exponential distribution.*

For details about the proof, see e.g. [27], [3]. This result reveals to be a very important one in the probabilistic analysis of the vector $\boldsymbol{X} \equiv (X_1, \cdots, X_n)$ and of related properties of dependence. The same result can be viewed from different standpoints, as a special case of more general results presented in more general settings in the field of Probability and Theory of Stochastic Processes.

The transformation Ψ that maps the vector $\boldsymbol{X} \equiv (X_1, \cdots, X_n)$ into $\mathcal{H} \equiv (\mathcal{H}_1, \cdots, \mathcal{H}_n)$ admits an "almost inverse transformation" Ψ^* such that the relation

$$\boldsymbol{X} = \Psi^* \left(\Psi \left(\boldsymbol{X} \right) \right)$$

holds almost surely. See e.g. [28] and references cited therein. Thus $\boldsymbol{X}$ has the same probability distribution of a vector obtained by a suitable transformation of a vector of i.i.d. exponential random variables. Also this result can be viewed from different standpoints, and it is a special case of more general results given in the literature. These results are at the basis of the method of the *total hazard construction* (see in particular [27]). The latter method has been applied in several different fields; see for example [39].

5. M.C.H.R. functions and dependence modeling

In this section we collect our comments concerning the role of M.C.H.R. functions and related aspects in stochastic dependence. We remind that the joint probability density and the system of the M.C.H.R. functions are two equivalent tools, both apt to characterize an absolutely continuous probability measure over $\mathbb{R}_+^n$. We will in particular point out analogies and differences between them. In fact, despite their substantial equivalence, these two tools are dissimilar in nature and they are adapt to respectively pointing out different aspects of the condition of absolute continuity.

5.1. **Different forms of inference.** First of all we notice that the joint probability density and the system of the M.C.H.R. functions induce two different methods to describe stochastic dependence among different variables. From one point of view, the use of the density function seems to directly permit more general forms of inference. One can consider, for two disjoint subsets of indexes $A', A'' \subset [n]$, observation of data of the form

$$X_i = x_i, i \in A'; X_l > t_l, l \in A'' ,$$

(namely both *failure data* and *survival data* are *observed*). Conditional on these data, the information about the non-observed variables X_j $(j \notin A' \cup A'')$ is described by the conditional density

$$f_{\boldsymbol{X}_{\widetilde{A' \cup A''}}}(\boldsymbol{x}_{\widetilde{A' \cup A''}} | \boldsymbol{X}_{A'} = \boldsymbol{x}_{A'}, \boldsymbol{X}_{A''} > \boldsymbol{t}_{A''}) .$$

Or, more in particular, one can consider data of the form

$$X_i = x_i, i \in A ,$$

(only *failure data* are observed). Conditional on these data, the information about the non-observed variables X_j $(j \notin A)$ is described by the conditional density

$$f_{\boldsymbol{X}_{\widetilde{A}}}(\boldsymbol{x}_{\widetilde{A}} | \boldsymbol{X}_A = \boldsymbol{x}_A) = \frac{f(\boldsymbol{x}_A, \boldsymbol{x}_{\widetilde{A}})}{f_{\boldsymbol{X}_A}(\boldsymbol{x}_A)} .$$

The latter formula shows that the computation of marginal densities $f_{\boldsymbol{X}_A}$ is needed for inference purposes. Such a computation is, in any case, feasible in terms of the joint density $f(\cdot)$.

As far as the use of the M.C.H.R. functions is concerned, we have seen that observed data must have the form of a dynamic history $\boldsymbol{h}_t$ as in (4). Namely the observed survival data, being of the specific form $Y_I > t$, concern all of the non-completely observed variables; furthermore it must be $t_1 < \cdots < t_k \leq t$. Given these data, inference then concerns with the residual lifetimes $X_j - t$, with $j \notin I$. The joint distribution of $X_j - t$ (for $j \notin I$) is described in terms of its *conditional* M.C.H.R. functions, which can easily be obtained in terms of the original M.C.H.R. functions of the entire vector $(X_1, \cdots, X_n)$.

It is interesting to notice such an analogy: when the joint distribution is assessed in terms of the joint density function, the conditional distribution of non-observed variables, given observed data, is still expressed in terms of a (conditional) density function; when the joint distribution is assessed in terms of the M.C.H.R. functions, the conditional distribution of non-observed variables, given observed data, is still expressed in terms of (up-dated) M.C.H.R. functions.

5.2. **Dependence modeling.** Conditional independence is a very natural and direct source of stochastic dependence among different random variables. Such a property often emerges in the presence of some factor Θ (a scalar quantity or a vector) which is unobservable to the analyst and relevant in the analysis at hand, at a time.

As well know, an extremely large literature has been dedicated to this issue, both for theoretical and foundational insight and from a modeling and applied viewpoint, in many fields of Science, Operation Research and Management. The theoretical analysis is strictly related to the basic paradigm of Subjective Probability and of Bayesian Statistics, according to which any unobservable quantity should be treated as a random variable, whose probability distribution describes the analyst's state of information. Basic aspects in this direction had been in particular pointed out by Bruno de Finetti (see e.g. [7], [8]) who used the term “dipendenza per arricchimento di informazione” (dependence through an increase of information). In the last few decades, the interest toward this source of dependence in the analysis of real situations has increased enormously. In particular we cite the fields of Reliability and of Financial Risk. In the former field, where the life-times of different industrial components are to be predicted, the factor Θ is often defined in terms of unknown levels of components' quality; in other cases Θ is related to environmental factors, playing a simultaneous influence on all the different components that are working in a same time period. See e.g. [19], [6], [35]. Θ can have the meaning of an unobservable macro-economic, or financial, factor in the field of Financial Risk, where this phenomenon of dependence has been also termed *information-based default contagion*. Often, the setting in the field of portfolio credit risk requires a more general treatment than the one provided by the M.C.H.R. functions. In fact, default times are observed for different assets which start to be observed in different instants. However the basic ideas are not so different to those underling the use of the M.C.H.R. functions.

The same issue also arises in many other fields; for an example, see e.g. [25].

In both the fields of Reliability and of Financial Risk, in any case, the constant factor Θ should sometimes be replaced by a stochastic process $\{\Theta_t\}_{t\geq 0}$. In such cases the formulas for updating the hazard rates of surviving units are obtained on the basis of a logic which is similar to that valid in the case of a constant factor Θ. From a conceptual point of view, we simply have to replace the formula (8) with a more general formula of the type

$$\lambda_j(t|I; t_1, \cdots, t_k) = \mathbb{E}\left(r_j\left(t; \Theta_t\right)|\boldsymbol{h}_t\right) .$$

However the computation of the conditional density $\pi_t(\cdot|\boldsymbol{h}_t)$, of the instantaneous value of Θ_t at time t given the observed history $\boldsymbol{h}_t$, is much more complex. In fact, since the updated information coming from $\boldsymbol{h}_t$ must be combined with the information concerning the time-evolution of Θ_t, techniques of *stochastic filtering* and *stochastic calculus* are needed. For instance, concerning dynamic models of interest in financial risk, see Chapter 9 in [20] and references cited therein.

In such models of conditional independence, as more generally in the *copula models*, the dependence structure among default times is exogenously specified, in terms of the joint density function or of the connecting or survival copula. Correspondingly, the default intensities are endogenously derived.

On the other hand, in several applied studies, stochastic dependence is not due to an increase of information about an unobservable factor. Rather the source of dependence dwells in the direct interactions between different units under analysis. And when the variables under observation are just lifetimes (or times to failures) of different units, then it can be natural to model such interactions in terms of the changes that the failure of a unit triggers on the hazard rate of the other units that are still surviving. These are the models with *interacting intensities*, where the joint density is endogenously derived. In the field of Financial Risk, *counterparty risk* and *default dependence* are described directly in terms of interacting intensities. In those reliability problems where lifetimes are analyzed for components that start to work simultaneously, on the other hand, these are models where stochastic dependence is described directly in terms of the M.C.H.R. functions. In particular, the condition of Load Sharing can give rise to a paradigmatic case for these models.

This brief discussion points out that two types of dependence models are met in *failure* or in *default* analysis: those where the joint density or the survival copula are specified exogenously and those where the joint density or the survival copula are specified endogenously.

Several times, the latter models with interacting intensities are the most natural to describe dependence among defaults. In particular, the properties briefly reviewed in the previous Section show that dependence can be described taking into account how variables of interest can be transformed in order to get i.i.d. exponential variables. Furthermore, the hazard rate construction can be used in the simulation of those variables.

As a drawback, it has been noticed that, for such models, the derivation of the joint density or the survival copula from knowledge of the interacting intensities can be a complex operation. In particular some of these models, that manifest simple forms of interacting intensities, give rise to relatively more complex forms for the one-dimensional marginal distributions.

Example 10. For the lifetimes X_1, X_2 of $n = 2$ reliability components, we consider the time-homogeneous Load-Sharing Model characterized by parameters λ and λ': initially, both the components share the instantaneous hazard rate λ; after the first component's failure is observed, the instantaneous hazard rate for the surviving component becomes λ'. This gives rise to an exchangeable Load-Sharing Model (such models have also be called *Ross Models*) and the corresponding density function is

$$f(x_1, x_2) = \lambda \, \lambda' \exp\{-(2\lambda - \lambda')\, x_{(1)} - \lambda' x_{(2)}\} \, .$$

See also [Ross, 1984]. Being $f(x_1, x_2)$ exchangeable, X_1, X_2 share the same one-dimensional marginal distribution. By a few computations one realizes that the corresponding marginal density has the form of a mixture of two different exponential densities. Thus, the survival copula corresponding to such a joint distribution has a complex form.

In the field of Financial Risk, in particular, the circumstance that marginal distributions (and survival copulas) are not available in closed form makes more involved *the calibration of the model to defaultable term structure data* (see [20]).

Dealing with dynamic dependence models in the reliability field, a special notion of *dynamic* (conditional) marginal distribution had been introduced in [30]; interesting comments and examples have been presented therein. Such a concept of dynamic marginal can potentially reveal to be useful in the developments of topics discussed here.

It is also to be mentioned, on the other hand, that for some models specified in terms of conditional independence, the derivation of the M.C.H.R. functions can vice-versa become rather complex. Furthermore, more generally, copula models can manifest weak points and drawbacks (see [21]). The latter are in any case different from those related with interacting intensities.

It may appear, in conclusion, that a separation can be traced between the two classes of different dependence models respectively defined by information-based default contagion and interacting intensities. However, generally, the distinction between such two classes can be considered rather subtle. Only, we can say, models of *conditional independence* and models of *Load-Sharing* give rise, respectively, to paradigmatic and extreme cases of the two classes.

Many other interesting models, not belonging to such classes, can be obtained by assuming that the Load-Sharing property holds conditionally on unknown parameters. More extensive discussions about these issues may be the subject of further research.

5.3. **Comparisons between properties of positive dependence.** In the different fields of applied probability, several notions of *positive dependence* among observable random variables have been defined, on the purpose of obtaining useful comparison and inequalities. In particular different such notions have been compared and chains of implications among them have been proven. Also for this field, the specialized literature is by now very extended and providing an adequate list of relevant references is beyond the purposes and possibilities of this paper. Some basic references can be found in [23], [13]. In any case, positive dependence is strictly related with concepts of *stochastic orderings*, and *stochastic monotonicity.* See [29], [38], [22], [31], [18]; see also [10].

In the special case of stochastic dependence triggered by conditional independence, more specifically, positive dependence is implied by the assumption that the observable variables have a same kind of monotonicity behavior with respect to variation of values taken by the unobservable variables. This idea can be formalized and several precise results can be obtained in terms of different concepts of stochastic monotonicity. Even very strong conditions of positive dependence can be obtained in this way.

In the case of Load-Sharing models, on the other hand, conditions of positive dependence are obtained under monotonicity of the coefficients $\mu_j(t|I)$ with respect to I: for any $t > 0$ and $j \in [n]$,

$$I' \subset I'' \Rightarrow \mu_j(t|I') \leq \mu_j(t|I'') \ .$$

Such a condition is strictly related to the concepts of *Weakened by Failure* (WBF) introduced in [4], [24], [29]. The WBF property can, more generally, manifest also in other cases of default contagion in interacting intensity models. It turns out that, under simple conditions, WBF is implied by *Multivariate Totally Positive Dependence of Order 2* (MTP_2). We recall that a vector of non-negative random variables $\boldsymbol{X} = (X_1, \cdots, X_n)$, with a joint density function f, is MTP_2 if [15]

$$f(\boldsymbol{x}') \cdot f(\boldsymbol{x}'') \leq f\left(\boldsymbol{x}' \bigwedge \boldsymbol{x}''\right) \cdot f\left(\boldsymbol{x}' \bigvee \boldsymbol{x}''\right) \ .$$

MTP_2 is in fact a very strong property of positive dependence; see e.g. [29].

In some sense, positive dependence created by interacting intensities may be perceived, from an heuristic point of view, as something stronger than positive dependence created by conditional independence. Actually this intuition can reveal to be false. For instance, a vector $\boldsymbol{X} = (X_1, \cdots, X_n)$ of conditionally independent lifetimes given a non-observable factor Θ, can even have the MTP_2 property under suitable conditions of stochastic monotonicity of the variables $X_1, \cdots, X_n$ with respect to Θ; see [33]; see also [14], for extensions to more general settings. We can thus see that conditional independence is able to give rise to stronger conditions of positive dependence than interacting intensity can do.

Acknowledgements.
This paper is dedicated to the dear memory of Moshe Shaked, with whom I have had along years several discussions on the theme of stochastic dependence.

References

[1] E. Arjas, *The failure and hazard processes in multivariate reliability systems*, Math. Oper. Res., 6(1981), 551–562.

[2] E. Arjas, *Information and Reliability. A Bayesian Perspective*, in Reliability and decision making, R. Barlow, C.A. Clarotti and F. Spizzichino eds., Chapman & Hall, London, 1993, 115–135.

[3] E. Arjas & P. Haara, *A note on the exponentiality of total hazards before failure*, J. Multivariate Anal. , 26(1988), 207–218.

[4] E. Arjas & I. Norros, *Life lengths and association: a dynamic approach*, Math. Oper. Res., (1984), 151–158.

[5] P. Brémaud, *Point processes and queues: martingale dynamics*, Springer Series in Statistics, Springer-Verlag, New York, 1981.

[6] E. Çinlar, M. Shaked & J.G. Shanthikumar, *On lifetimes influenced by a common environment*, Stochastic Process. Appl., 33(1989), 347–359.

[7] B. deFinetti, *La prévision: ses lois logique, ses sources subjectives*, Annales de l'IHP, 7(1), 1–68.

[8] B. deFinetti, *Teoria della probabilità*, Einaudi, Torino, 1970, (English translation: Theory of Probability, Wiley, New York, 1974.

[9] F. Durante & C. Sempi, *Principles of copula theory*, CRC Press, Boca Raton, 2015.

[10] R. Foschi & F. Spizzichino, *Reversing conditional orderings*, in Stochastic orders in reliability and risk, H. Li & X. Li eds., Lecture Notes in Statistics, 208(2013), 59–80.

[11] J.E. Freund, *A bivariate extension of the exponential distribution*, J. Amer. Statist. Assoc., 56(1961), 971–977.

[12] J. Jacod, *Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales*, Probab. Theory Related Fields, 31(1975), 235–253.

[13] H. Joe, *Multivariate models and multivariate dependence concepts*, Monographs on Statistics & Applied Probability, 73, Chapman and Hall, London, 1997.

[14] B.E. Khaledi & S. Kochar, *Dependence properties of multivariate mixture distributions and their applications*, Ann. Inst. Statist. Math, 53(2001), 620–630.

[15] S. Karlin & Y. Rinott, *Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions*, J. Multivariate Anal., 10(1980), 467-498.

[16] G. Koch, *A dynamical approach to reliability*, in Teory of reliability theory, R. Barlow & A. Serra eds., North Holland, 1985; Proceedings of a summer course in Reliability theory, Varenna (Italy), 1984.

[17] C.D. Lai & M. Xie, *Stochastic ageing and dependence for reliability*, Springer-Verlag, New York, 2006.

[18] H. Li H. & X. Li X., *Stochastic orders in reliability and risk*, Lecture Notes in Statistics, 208, Springer-Verlag, 2013.

[19] D.V. Lindley & N.D. Singpurwalla, *Multivariate distributions for the life lengths of components of a system sharing a common environment*, *J. Appl. Probab.*, 23(1986), 418–431.

[20] A.J. McNeil, R. Frey & P. Embrechts, *Quantitative risk management: concepts, techniques and tools*, Princeton Series in Finance, Revised Edition, Princeton Series in Finance, 2015.

[21] T. Mikosch, *Copulas: tales and facts*, Extremes, 9(2006), 1, 3–20.

[22] A. Muller & D. Stoyan, *Comparison methods for stochastic models and risks*, Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., Chichester, 2002.

[23] R.B. Nelsen, *An introduction to copulas*, Springer Series in Statistics, II edition, Springer-Verlag, 2006.

[24] I. Norros, *Systems weakened by failures*, Stochastic Process. Appl., 20(1985), 81–196.

[25] O. Purcaru O & M. Denuit, *Dependence in dynamic claim frequency credibility models*, ASTIN Bulletin, 33(2003), 23–40.

[26] S.M. Ross, *A model in which component failure rates depend on the working set*, Naval Res. Logist., 31(1984), 297-300.

[27] M. Shaked & J.G. Shanthikumar, *The multivariate hazard construction*, Stochastic Process. Appl., 24(1987), 241–258.

[28] M. Shaked & J.G. Shanthikumar, *Multivariate conditional hazard rates and the MIFRA and MIFR properties*, J. Appl. Probab., 25(1988), 150–168.

[29] M. Shaked & J.G. Shanthikumar, *Multivariate stochastic orderings and positive dependence in reliability theory*, Math. Oper. Res., 15(1990), 545–552.

[30] M. Shaked & J.G. Shanthikumar, *Dynamic conditional marginal distributions in reliability theory*, J. Appl. Probab., 30(1993), 421–428

[31] M. Shaked M. & J.G. Shanthikumar, *Stochastic orders*, Springer Series in Statistics, Springer-Verlag, New York, 2007.

[32] M. Shaked & J.G. Shanthikumar, *Multivariate conditional hazard rate functions: an overview*, Appl. Stoch. Models Bus. Ind., 31(2015), no. 3, 285-296.

[33] M. Shaked & F. Spizzichino, *Positive dependence properties of conditionally independent random lifetimes*, Math. Oper. Res., 23(1998), 944–959.

[34] Shechner Z., *A load-sharing model: The linear breakdown rule*, Naval Res. Logist., 31(1984), 137–144.

[35] N.D. Singpurwalla & M.A. Youngren, *Multivariate distributions induced by dynamic environments*, Scand. J. Stat., 20(1993), 251–261.

[36] F. Spizzichino, *Subjective probability models for lifetimes*, Chapman and Hall, CRC Press, Boca Raton, 2001.

[37] F. Spizzichino, *Aging and positive dependence*, Encyclopedia of Statistics for Quality and Reliability. Wiley & Son Limited, Chichester, 2007, 82–95.

[38] R. Szekli, *Stochastic ordering and dependence in applied probability*, Lecture Notes in Statistics, 97, Springer-Verlag, New York, 1995.

[39] F. Yu, *Correlated defaults in intensity-based models*, Math. Finance, 17(2)(2007), 155--173.