

Construction of k -Variate Survival Functions with Emphasis on the Case $k = 3$

Jerzy K. Filus¹, Lidia Z. Filus²

¹Dept. of Mathematics and Computer Science, Oakton College, Des Plaines, USA

²Dept. of Mathematics, Northeastern Illinois University, Chicago, USA

Email: jkfilus98@gmail.com, L-Filus@neiu.edu

How to cite this paper: Filus, J.K. and Filus, L.Z. (2020) Construction of k -Variate Survival Functions with Emphasis on the Case $k = 3$. *Applied Mathematics*, 11, 678-697.
<https://doi.org/10.4236/am.2020.117046>

Received: April 28, 2020

Accepted: July 25, 2020

Published: July 28, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The purpose of this paper is to present a general universal formula for k -variate survival functions for arbitrary $k = 2, 3, \dots$, given all the univariate marginal survival functions. This universal form of k -variate probability distributions was obtained by means of “dependence functions” named “joiners” in the text. These joiners determine all the involved stochastic dependencies between the underlying random variables. However, in order that the presented formula (the form) represents a legitimate survival function, some necessary and sufficient conditions for the joiners had to be found. Basically, finding those conditions is the main task of this paper. This task was successfully performed for the case $k = 2$ and the main results for the case $k = 3$ were formulated as Theorem 1 and Theorem 2 in Section 4. Nevertheless, the hypothetical conditions valid for the general $k \geq 4$ case were also formulated in Section 3 as the (very convincing) Hypothesis. As for the sufficient conditions for both the $k = 3$ and $k \geq 4$ cases, the full generality was not achieved since two restrictions were imposed. Firstly, we limited ourselves to the, defined in the text, “continuous cases” (when the corresponding joint density exists and is continuous), and secondly we consider positive stochastic dependencies only. Nevertheless, the class of the k -variate distributions which can be constructed is very wide. The presented method of construction by means of joiners can be considered competitive to the **copula** methodology. As it is suggested in the paper the possibility of building a common theory of both copulae and joiners is quite possible, and the joiners may play the role of tools within the theory of copulae, and vice versa copulae may, for example, be used for finding proper joiners. Another independent feature of the joiners methodology is the possibility of constructing many new stochastic processes including stationary and Markovian.

Keywords

Construction of Multivariate Probability Distributions via Joiners, Joiner versus **Copula** Methodology, A Possible Fusion of the Two Construction Methods, k -Variate **Survival Function** Scheme, $k = 3$ Case

1. Introduction

This work includes a continuation of our previous papers [1] [2] [3] and [4] on analysis and construction methods of multivariate survival functions, given their univariate marginals.

This subject was widely developed in the literature since the late 1950s and 60s (for example, see [5] [6] [7] [8] [9]) until nowadays. Many methods of construction of bivariate and multivariate probability distributions were developed more recently, only to mention the conditionally determined bivariate and multivariate distributions in [10] [11] [12]. Some similar but different methods of constructing models, under the name “parameter dependence method”, can be found in [13] [14]. As it turned out many of the multivariate distributions obtained by that method can also be obtained by the “method of triangular transformations” (especially by pseudoaffine and pseudopower transformations), see for example [15].

Many other options for construction one can be found in the references of [7].

The approach that is developed here and in [1] [2] [3] has its origin [4] in the model which is the Aalen version [16] of the famous Cox model [17] for stochastic dependence.

However, as a result of further development we finally were able to formulate (here and in [1] [2] [3]) a version of an emerging theory independent of its Aalen origin.

Thus, in our approach, as presented here and in [1] [2] [3], both the construction and the universal characterization of multivariate models rely on incorporating the so called “**joiners**”. These joiners we propose to name “**Aalen factors**” (but we will use the name “joiner” for short). They are the functions which fully determine all underlying stochastic dependencies between the considered random variables.

As a result, the so obtained k -variate survival functions gain a nice factored form

$$P(X_1 > x_1, \dots, X_k > x_k) = J_{1, \dots, k}(x_1, \dots, x_{k-1}, x_k) S_1(x_1) S_2(x_2) \dots S_k(x_k),$$

where $S_1(x_1), S_2(x_2), \dots, S_k(x_k)$ represent all the, given in advance, univariate marginal survival functions and the “dependence factor” $J_{1, \dots, k}(x_1, \dots, x_{k-1}, x_k)$ is the joiner. Clearly, the case, where $J_{1, \dots, k}(x_1, \dots, x_{k-1}, x_k) = 1$ everywhere, is equivalent to stochastic independence.

Now, the task of the construction of any k -variate probability distribution can

simply be formulated as follows: given all the univariate marginal distributions, represented by the survival functions $S_1(x_1), S_2(x_2), \dots, S_k(x_k)$, find a proper joiner $J_{1,\dots,k}(x_1, \dots, x_{k-1}, x_k)$ so that the product

$J_{1,\dots,k}(x_1, \dots, x_{k-1}, x_k) S_1(x_1) S_2(x_2) \dots S_k(x_k)$ is a valid k -variate survival function.

This task formulation, in terms of the survival functions, resembles the task of finding a proper **copula** for given univariate cdfs, say, $F_1(x_1), F_2(x_2), \dots, F_k(x_k)$ in order to obtain a k -variate cdf. as a stochastic model for some investigated *reality* (*i.e.*, “outside of mathematics”) represented by a given set of data.

It becomes then clear that, given, “essentially the same” input data, $F_1(x_1), F_2(x_2), \dots, F_k(x_k)$, the “method of joiners” stands as an alternative (and competitive [18]) method to the copula methodology.

On the other hand, as we point out in Section 2, there is a possibility to formulate a **common general theory** of both copula and joiner representations of bivariate and, possibly, also k -variate ($k \geq 3$) probability distributions.

This possibility follows from the fact that both representations (by copulas and by joiners) are equivalent as describing the same probability distribution. Therefore, any copula uniquely determines a corresponding joiner, and any joiner uniquely determines a copula. This fact indicates the way to find more copulas through joiners, as well as more joiners that correspond to known copulas. Moreover, the method of joiners may, likely, be used to develop more copulas theory for higher dimensions.

These remarks only signalize the possibility of such a common theory. We intend to develop this in more detail in the nearest future.

At the moment, we rather concentrate on the joiner representation and the joiner based methods of models construction.

In the next section we provide a short introduction to the joiners’ theory by providing a slightly different (as compared with our previous works) formulation of bivariate case. This case is fundamental to our considerations in sections 3 and 4 since for the k -variate ($k \geq 3$) survival functions we restrict our attention to “**bi-dependence**” only, which means only bivariate joiners may be different than 1 (see, [2]). This restriction dramatically simplifies theory of k -variate distributions as compared with the general theory (for arbitrary k) developed in [2].

In Section 3 we provide only a general scheme of k -variate distributions’ analytical form (for arbitrary k) under the bi-dependence assumption.

As for the k -variate distribution for which the joiner (in this case the joiner is reduced to product of bivariate joiners that correspond to all, or to some only, bivariate marginals) must fit the given k univariate marginals, we only formulate the main result as Hypothesis.

Even though the Hypothesis is very convincing, we were not able to provide a formal proof. The arguments, which made us strongly believe it holds, mostly (but not exclusively) follow from the fact that the same pattern as we presume holds for an arbitrary k , holds for the cases $k = 2$ and $k = 3$.

As mentioned, the (hypothetical) conclusion, extending these two cases to all k , is not only based on that analogy, but also on the naturalness of the assertion. The case $k = 3$ is the subject of Section 4. The main thesis about the 3-dimensional model is formulated and proven there in two theorems.

The general case, *i.e.*, k -variate distribution for any $k = 1, 2, \dots$, actually defines an infinite sequence of random variables $X_1, X_2, \dots, X_k, \dots$ which, as it is pointed out at end of Section 3, satisfies the Kolmogorov consistency condition. So, we actually defined a class of discrete time (now, k represents “time”) **stochastic processes** with a variety of interesting special cases.

Clearly, such processes need not to be very complicated if we assume that most of the bivariate joiners reduce to 1. Based on such possibility we may construct **m-Markovian** processes for $m = 1, 2, \dots$ (they are Markovian if $m = 1$). Also by adopting natural assumptions we may gain **stationarity** of some constructed stochastic processes. It is an exciting possibility that having only one bivariate distribution of any neighboring random variables, say X_{k-1}, X_k , we may easily construct both stationary and Markovian stochastic processes.

This subject is only touched upon in Section 3, but is out of scope of this work.

At the end of this Introduction we must notice that, according to our common assumption that every joiner is not bigger than 1 (*i.e.*, its corresponding representation, by the defined throughout the text function $\Psi(\cdot, \cdot)$, is nonnegative) we limited ourselves, to positive stochastic dependencies. Extension to models also comprising negative dependencies is quite possible, but requires a more complex theory.

2. The Bivariate Case

Before constructing new classes of multivariate and, especially, tri-variate survival functions, in this section we repeat and slightly modify the bivariate cases which involve their bivariate universal forms, referring for more details to our previous papers [1] [3] [4].

Thus, according to those papers, any bivariate survival function $S(x_1, x_2) = P(X_1 > x_1, X_2 > x_2)$ of an arbitrary random vector (X_1, X_2) can be represented as:

$$S(x_1, x_2) = S_1(x_1)S_2(x_2)J(x_1, x_2), \quad (1)$$

where $S_1(x_1), S_2(x_2)$ are the marginal functions of $S(x_1, x_2)$ and the function $J(x_1, x_2)$, that was called the joiner, determines all the stochastic dependence between the random variables X_1, X_2 .

As it was argued in [3], form (1) is universal in the sense that every bivariate survival function has the unique representation (1).

Unlike papers [2] [3] in this work we restrict our attention to the so called “**continuous**” case [1] which means we adopt the assumption that both hazard rates of the marginals $S_1(x_1), S_2(x_2)$, say $\lambda_1(x_1), \lambda_2(x_2)$, do exist and are continuous.

Moreover, we assume that for any considered joiner $J(x_1, x_2)$ there is

unique representation by the continuous function $\Psi(x_1, x_2)$ such that:

$$J(x_1, x_2) = \exp \left[- \int_0^{x_1} \int_0^{x_2} \Psi(t, u) dt du \right] \tag{2}$$

The above assumptions allow us to represent the bivariate survival function (1) in the following exponential form:

$$S(x_1, x_2) = \exp \left[- \int_0^{x_1} \lambda_1(t) dt - \int_0^{x_1} \int_0^{x_2} \Psi(t, u) dt du - \int_0^{x_2} \lambda_2(u) du \right] \tag{3}$$

The problem that occurs at this stage is to find an answer to the following question [2]:

Given are a fixed, but arbitrary, pair of marginals $S_1(x_1), S_2(x_2)$ represented by the corresponding continuous hazard rates $\lambda_1(x_1), \lambda_2(x_2)$. What conditions must be satisfied by any continuous function of two nonnegative real variables, say $G(x_1, x_2)$, so that the product $S_1(x_1)S_2(x_2)G(x_1, x_2)$ is a legitimate survival function, *i.e.*, $G(x_1, x_2) = J(x_1, x_2)$ for a proper joiner $J(x_1, x_2)$ associated with $\lambda_1(x_1), \lambda_2(x_2)$.

Necessary conditions, for $G(x_1, x_2) = J(x_1, x_2)$, where $J(x_1, x_2)$ is a proper joiner, are that

$$G(x_1, 0) = 1, \text{ for each } x_1, \text{ and } G(0, x_2) = 1 \text{ for each } x_2, \tag{4}$$

and consequently $G(0, 0) = 1$, see [1] or [4].

Condition (4) is, however, not sufficient.

Assuming there is an exponential representation of $G(x_1, x_2)$

$$G(x_1, x_2) = \exp \left[- \int_0^{x_1} \int_0^{x_2} \Phi(t, u) dt du \right],$$

where $\Phi(t, u)$ is a continuous function satisfying (4), equality (3) may be re-written into the form

$$R(x_1, x_2) = \exp \left[- \int_0^{x_1} \lambda_1(t) dt - \int_0^{x_1} \int_0^{x_2} \Phi(t, u) dt du - \int_0^{x_2} \lambda_2(u) du \right] \tag{5}$$

The question whether, for the marginals given by the hazard rates $\lambda_1(x_1), \lambda_2(x_2)$, the function $R(x_1, x_2)$ is a legitimate bivariate survival function (in symbols, whether $R(x_1, x_2) = S(x_1, x_2)$) reduces to the question whether $\Phi(t, u)$ determines a proper joiner, *i.e.*, whether $\Phi(t, u) = \Psi(t, u)$, where the function $\Psi(t, u)$ determines a proper joiner $J(x_1, x_2)$ by (2).

Suppose the function $G(x_1, x_2)$ satisfies the necessary conditions (4).

Now, the sufficient condition for $\Phi(t, u) = \Psi(t, u)$ (for some proper function $\Psi(t, u)$, given by (2) is equivalent to the requirement that the second mixed derivatives $\partial^2/\partial x_1 \partial x_2$ and $\partial^2/\partial x_2 \partial x_1$ of $R(x_1, x_2)$, as given by (5), are equal to each other and are nonnegative for all x_1, x_2 .

Obviously they do exist and are continuous by the earlier assumed continuity of the functions $\lambda_1(t), \lambda_2(u)$ and $\Psi(t, u)$.

The nonnegativity requirement for $\partial^2/\partial x_1\partial x_2 R(x_1, x_2)$ is equivalent to the simple common fact that the joint density of any random vector (X_1, X_2) , if it exists, must be nonnegative.

As it follows from the form of (3), other properties, characterizing that density's antiderivative (the cdf.) are satisfied automatically.

Thus, to obtain the sufficient condition for $R(x_1, x_2) = S(x_1, x_2)$ it is enough to set

$$\partial^2/\partial x_1\partial x_2 R(x_1, x_2) \geq 0, \tag{6}$$

where $R(x_1, x_2)$ is given by (5).

After calculating the second mixed derivative from (5), then simplifying underlying expressions and setting expressions with negative signs to the right-hand side of the inequality, one obtains (6) in the form of the following integral inequality:

$$\left[\lambda_1(x_1) + \int_0^{x_2} \Phi(x_1, u) du \right] \cdot \lambda_2(x_2) + \int_0^{x_1} \Phi(t, x_2) dt \geq \Phi(x_1, x_2) \tag{7}$$

Now, the task of finding the bivariate distribution, given the marginals as represented by the hazard rates $\lambda_1(x_1)$, $\lambda_2(x_2)$, reduces to solving integral inequality (7) with respect to the only unknown function $\Phi(x_1, x_2)$. This means, any solution $\Phi(x_1, x_2) = \Psi(x_1, x_2)$ of (7) is functionally dependent on the marginal distributions $S_1(x_1)$, $S_2(x_2)$, here represented by $\lambda_1(x_1)$, $\lambda_2(x_2)$. So, in the continuous case, the set of all the solutions $\Phi(x_1, x_2)$ of (7) uniquely determines the set of all bivariate distributions having the same fixed marginals. All the functions $\Phi(x_1, x_2)$, satisfying conditions (4) and (7), will be denoted by the symbol $\Psi(x_1, x_2)$. So that whenever writing $\Psi(x_1, x_2)$ we will mean the function representing a proper joiner and the corresponding, by (3), function denoted by " $S(x_1, x_2)$ " will be treated as a legitimate bivariate survival function.

In the case $\Phi(x_1, x_2) \geq 0$, for all x_1, x_2 , (positive stochastic dependence case) if the inequality:

$$\lambda_1(x_1)\lambda_2(x_2) \geq \Phi(x_1, x_2) \tag{8}$$

holds then (7) holds too.

Thus, the conditions $\Phi(x_1, x_2) \geq 0$ and (8) are sufficient conditions for $\Phi(x_1, x_2) = \Psi(x_1, x_2)$, where $\Psi(x_1, x_2)$ determines a proper bivariate survival function $S(x_1, x_2)$ by (3).

Notice, however, that condition (8) is not a necessary condition.

Nevertheless, any solution of (8) is a solution of (7) and is very easy to find.

The simplest set of such solutions obviously is given by $\Psi(x_1, x_2) = a\lambda_1(x_1)\lambda_2(x_2)$, where for the constant parameter a (to be statistically estimated) we require $0 \leq a \leq 1$.

One then obtains as a special case of (3) the following model:

$$S(x_1, x_2) = \exp \left[-\int_0^{x_1} \lambda_1(t) dt - a \int_0^{x_1} \int_0^{x_2} \lambda_1(t) \lambda_2(u) dt du - \int_0^{x_2} \lambda_2(u) du \right] \tag{9}$$

Model (9), which is somehow related to the first Gumbel bivariate exponential [6], is the most natural (and, possibly, kind of “canonical”) solution to the problem of finding the joint distribution of random variables X_1, X_2 , given the margins represented by the hazard rates.

We *expect* many applications of this model according to the common opinion that the simplest (“but not yet simpler”) models mostly are the best reflections of modeled realities.

Model (9) can be generalized to the following one:

$$S(x_1, x_2) = \exp \left[-\int_0^{x_1} \lambda_1(t) dt - ac(x_1, x_2) \int_0^{x_1} \int_0^{x_2} \lambda_1(t) \lambda_2(u) dt du - \int_0^{x_2} \lambda_2(u) du \right] \quad (10)$$

where an additional factor of the middle term in the exponent of (10) is any continuous function $c(x_1, x_2)$ satisfying: $0 \leq c(x_1, x_2) \leq 1$ for all nonnegative x_1, x_2 . Moreover, again $0 \leq a \leq 1$.

In particular, we propose the following natural model, with

$$c(x_1, x_2) = \exp[-\gamma x_1 x_2]:$$

$$S(x_1, x_2) = \exp \left[-\int_0^{x_1} \lambda_1(t) dt - a \exp[-\gamma x_1 x_2] \int_0^{x_1} \int_0^{x_2} \lambda_1(t) \lambda_2(u) dt du - \int_0^{x_2} \lambda_2(u) du \right] \quad (11)$$

where the constant real parameter γ (to be estimated) satisfies $\gamma \geq 0$.

The theory of **joiner representations** (as developed in this work and in [1] [2] [3] [4]) is competitive [18] to the theory of **copulas** [9].

However, it can be shown that the two theories are equivalent in the sense that there is a one-to-one relationship between joiners and copulas, at least in the bivariate case. Thus, having a joiner one immediately obtains the corresponding unique copula and vice versa. As it is well-known, every copula is “good” to any pair of marginal distributions, but, as follows from (7), not every “joiner” fits the marginals given by hazard rates $\lambda_1(x_1), \lambda_2(x_2)$. On the other hand substituting into any copula the given marginals, one obtains back (through that copula) the joiner that fits the substituted marginals. Thus, in such a way, one can obtain all the proper joiners through all the copulas that are known.

It is important to notice that, in applications to practical problems, it is easier to find a joiner that reflects a modeled reality than a proper copula (That “easiness” in finding a proper joiner follows from the fact that the models involving joiners are closely related to the Aalen [16] version of the Cox [17] model for stochastic dependencies.).

On the other hand, such a proper joiner determines the corresponding copula. This facilitates the choice of proper copula. In that sense the joiner approach may be considered as an enrichment of the copula methodology. This subject of a possible **common theory** of copulas and joiners will be included in our next publication which is now in preparation.

3. A Class of k -Variate Survival Functions

Before starting a more detailed analysis of tri-variate survival functions, which is

our main goal, we first introduce a simplified version of k -variate survival functions for any $k \geq 3$. This version turns out to be a special case of the most general k -variate model presented by formula (1) in [2].

The model here presented is much simpler as it only describes the “pairwise stochastic dependence” which was defined in [2]. According to the terminology in [2] such distributions are “three-independent”.

As it will be seen, this simplified case still preserves a significant amount of generality.

Consider the following simplified formula for a k -variate survival function of the random vector (X_1, \dots, X_k) :

$$\begin{aligned} S(x_1, \dots, x_k) &= P(X_1 > x_1, \dots, X_k > x_k) \\ &= J_{1, \dots, k}^*(x_1, \dots, x_{k-1}, x_k) S_1(x_1) S_2(x_2) \dots S_k(x_k) \end{aligned} \tag{12}$$

As mentioned, this formula may be considered as a special case of formula (1) in [2].

Now, however, as we assume **pairwise dependence** only, the joiner in (12) reduces to the following product of bivariate joiners:

$$J_{1, \dots, k}^*(x_1, \dots, x_k) = \prod_{1 \leq i < j \leq k} J_{i,j}(x_i, x_j). \tag{13}$$

Comparing (13) to the joiner as present in formula (3) of [2], one sees that (13) is obtainable from the most general case considered in [2] by setting to 1 all the joiners which are not bivariate.

Resuming, formula (12) can be rewritten as:

$$S(x_1, \dots, x_k) = \left\{ \prod_{1 \leq j < k \leq k} J_{i,j}(x_i, x_j) \right\} S_1(x_1) S_2(x_2) \dots S_k(x_k), \tag{14}$$

where any single bivariate joiner $J_{i,j}(x_i, x_j)$ is assumed to fit well to the corresponding two marginal survival functions $S_i(x_i), S_j(x_j)$ so that the products $J_{i,j}(x_i, x_j) S_i(x_i) S_j(x_j)$ are legitimate bivariate survival functions of the (marginal) random vectors (X_i, X_j) .

Upon the conditions $J_{r,l}(0, x_l) = J_{r,l}(x_r, 0) = 1$ for all r, l ($1 \leq r < l \leq k$) realize that by substituting $x_r = x_l = 0$ for all x_r and x_l different from any member of the fixed pair $\{x_i, x_j\}$ one obtains from (14), as the bivariate marginal, the (well defined) survival function:

$$S(x_i, x_j) = J_{i,j}(x_i, x_j) S_i(x_i) S_j(x_j). \tag{15}$$

In the continuous case this means that all the k hazard rates $\lambda_1(x_1), \dots, \lambda_k(x_k)$ do exist and are continuous, and every bivariate joiner $J_{i,j}(x_i, x_j)$ has a unique representation by a continuous function, say $\Psi_{i,j}(x_i, x_j)$, so that

$$J_{i,j}(x_i, x_j) = \exp \left[- \int_0^{x_i} \int_0^{x_j} \Psi_{i,j}(u_i, u_j) du_i du_j \right]. \tag{16}$$

Naturally, for all pairs of hazard rates $\lambda_i(x_i), \lambda_j(x_j)$ the corresponding functions $\Psi_{i,j}(x_i, x_j)$ are chosen in such a way that, for $1 \leq i < j \leq k$, the following

inequalities:

$$\left[\lambda_i(x_i) + \int_0^{x_j} \Psi_{i,j}(x_i, u_j) du_j \right] \cdot \left[\lambda_j(x_j) + \int_0^{x_i} \Psi_{i,j}(t_i, x_j) dt_i \right] \geq \Psi_{i,j}(x_i, x_j), \quad (17)$$

hold for all x_i, x_j , given any fixed pair of indexes i, j .

Inequalities (17) have the same structure as inequality (7).

With the above continuity assumptions, formula (14) can be rewritten in the following exponential form:

$$S(x_1, \dots, x_k) = \exp \left[- \int_0^{x_1} \lambda_1(t) dt - \dots - \int_0^{x_k} \lambda_k(t) dt - \sum_{1 \leq i < j \leq k} \int_0^{x_i} \int_0^{x_j} \Psi_{i,j}(u_i, u_j) du_i du_j \right] \quad (18)$$

Notice that if we again substitute into (18) $x_r = x_l = 0$ for all x_r and x_l different from any member of the fixed pair $\{x_i, x_j\}$, we obtain (15) in exponential form:

$$S^{(2)}(x_i, x_j) = \exp \left[- \int_0^{x_i} \lambda_i(t) dt - \int_0^{x_j} \Psi_{i,j}(u_i, u_j) du_i du_j \right] - \int_0^{x_j} \lambda_j(u_j) du_j \quad (19)$$

which upon condition (17) is a well-defined bivariate survival function.

Remark

It's easy to find out that if we set in (18) any $k - r$ ($1 \leq r < k$) variables (among all the variables x_1, \dots, x_k) to 0, we obtain the r -dimensional marginal survival function of the remaining random variables, say X_{i_1}, \dots, X_{i_r} which, syntactically, has an identical form as the k -dimensional survival function (18). This observation implies the following:

- 1) Since the pattern given by (18) is valid for every $k = 2, 3, \dots$ we, in fact, defined (at least theoretically) an infinite sequence of probability distributions.
- 2) The above observation on r -dimensional marginals of every k -dimensional distribution, (for each $r = 1, 2, \dots, k - 1$) clearly indicates that for the underlying sequence of random variables X_1, \dots, X_k, \dots ($k = 1, 2, \dots$) the Kolmogorov and Daniels consistency requirement is satisfied in this case. Therefore a fairly wide class of **stochastic processes** $\{X_k\}_{k=1,2,\dots}$ with discrete "time" k is defined.
- 3) These stochastic processes will be significantly simplified to **Markovian** if we set all the joiners $J_{i,j}(x_i, x_j)$ in (14) (not necessarily in the continuous case) to 1 or (in the continuous case) all the functions $\Psi_{i,j}(u_i, u_j)$ in (18) to 0 for all pairs (i, j) such that $j - i > 1$. Recall, we only considered situations where $i < j$.

Thus, under the foregoing assumption, only the "adjacent" random variables, say X_{i-1}, X_i are dependent, while, for example, variables X_{i-2}, X_i for any i such that $i - 2 \geq 1$ are independent.

If, however, we additionally let the joiners $J_{i-2,i}(x_{i-2}, x_i)$ be $\neq 1$, but all the joiners $J_{i-q,i}(x_{-q}, x_i) = 1$ for $q \geq 3$, then we obtain the **"bi-Markovian"** stochastic process in the sense that the probability distribution of any X_i only depends on realizations of the random variables X_{i-1} and X_{i-2} while it is in-

dependent of any realization of random variables “earlier” than X_{i-2} .

Quite similarly to bi-Markovian, one can define m -Markovian stochastic processes for $m = 1, 2, \dots$, so that the probability distribution of any $X_i (i > m)$ depends only on realizations of the random variables $X_{i-m}, X_{i-m+1}, \dots, X_{i-1}$, but does not depend on realizations of any of the random variables $X_1, X_2, \dots, X_{i-m-1}$.

In the latter case it is enough to set all the joiners $J_{s,i}(x_s, x_i)$ to 1, for $s = 1, 2, \dots, i - m - 1$.

Of course, in applications, the choice of joiner (*i.e.*, which one was to be set to 1) depends on the character of a modeled reality. Anyway, there is no need always to rely on Markovian ($m = 1$) processes.

4) Suppose, for the defined above Markovian ($m = 1$) stochastic processes, we assume that all the underlying hazard rates have the same functional form, *i.e.*,

$$\lambda(u) = \lambda_1(u) = \lambda_2(u) = \dots = \lambda_k(u) = \dots$$

Also assume one functional form $J_{i,j}(u, t) = J(u, t)$ for all the bivariate joiners present in defining formula (14).

Then the so defined Markovian (as well as any m -Markovian) process will be **stationary**, and all we need to analyze and for any further computations regarding the whole so encountered process, it is enough to know one bivariate survival function, say

$$S(x_{i-1}, x_i) = \exp \left[- \int_0^{x_{i-1}} \lambda(t) dt - \int_0^{x_{i-1}} \int_0^{x_i} \Psi(t, u) dt du - \int_0^{x_i} \lambda(u) du \right], \quad (20)$$

for all $i = 2, 3, \dots$.

In more general cases (not necessarily “continuous”) one may consider, instead, the survival function in the form:

$$S^{(2)}(x_{i-1}, x_i) = S(x_{i-1})S(x_i)J(x_{i-1}, x_i) \text{ for all } i = 1, 2, \dots$$

5) As one can realize, an interesting new theory of both random vectors and stochastic processes emerges. This is, however, not in the scope of this work and we postpone that subject for future research. □

As for (19) we already know that if [for all pairs (x_i, x_j)] given the marginals $\lambda_i(x_i), \lambda_j(x_j)$ the corresponding functions $J_{i,j}(x_i, x_j)$ satisfy the conditions $J_{i,j}(0, x_j) = J_{i,j}(x_i, 0) = 1$ uniformly for all nonnegative x_i, x_j , and if they also satisfy inequality (17) with respect to their representations $\Psi_{i,j}(x_i, x_j)$, then (19) represents all the valid marginal bivariate survival functions.

Unfortunately, in the general k -variate case ($k \geq 4$) we do not know yet sufficient conditions for all the functions $\Psi_{i,j}(u_i, u_j)$ that are present in (18) so that, given the marginals represented by $\lambda_1(x_1), \dots, \lambda_k(x_k)$, (18) represents a valid k -variate survival function for $k = 4, 5, \dots$.

Nevertheless, we have reasons to propose the following hypothesis which, unfortunately, at the moment, we are not able to prove rigorously.

In the general k -variate case an eventual rigorous proof would, possibly, re-

quire to establish some common pattern that would comprise the k^{th} mixed derivative $\partial^k / \partial x_1 \cdots \partial x_k$ from the right-hand side of (18).

Such a pattern should be valid for any $k = 3, 4, \dots$, and it is more than finding the k^{th} derivative for any particular relatively small k .

As for the hypothesis we presume what follows:

Hypothesis: For any given univariate marginals of a random vector (X_1, \dots, X_k) represented by the hazard rates $\lambda_1(x_1), \dots, \lambda_k(x_k)$ consider the following expression:

$$\exp \left[- \int_0^{x_1} \lambda_1(t_1) dt_1 - \dots - \int_0^{x_k} \lambda_k(t_k) dt_k - \sum_{1 \leq i < j \leq k} c_{ij} \int_0^{x_i} \int_0^{x_j} \Psi_{i,j}(u_i, u_j) du_i du_j \right] \quad (21)$$

where for any pair of indexes i, j such that $1 \leq i < j \leq k$ the continuous functions $\Psi_{i,j}(u_i, u_j)$ satisfy

$$0 \leq \Psi_{i,j}(u_i, u_j) \leq \lambda_i(u_i) \lambda_j(u_j). \quad (22)$$

Moreover, let the nonnegative real constants c_{ij} satisfy:

$$\sum_{1 \leq i < j \leq k} c_{ij} \leq 1. \quad (23)$$

Then expression (21) represents a valid k -dimensional survival function of (X_1, \dots, X_k) . \square

If the above hypothesis holds, then from (22) one obtains the following specific k -variate model:

$$S(x_1, \dots, x_k) = \exp \left[- \int_0^{x_1} \lambda_1(t_1) dt_1 - \dots - \int_0^{x_k} \lambda_k(t_k) dt_k - \sum_{1 \leq i < j \leq k} c_{ij} \int_0^{x_i} \int_0^{x_j} \lambda_i(u_i) \lambda_j(u_j) du_i du_j \right] \quad (24)$$

where the nonnegative constants c_{ij} satisfy (23).

Model (24) is an extension of the similar bivariate model (9) as well as the tri-variate model (26) given in the next section.

If the Hypothesis holds, then we also can generalize (24) to the following:

$$S(x_1, \dots, x_k) = \exp \left[- \int_0^{x_1} \lambda_1(t_1) dt_1 - \dots - \int_0^{x_k} \lambda_k(t_k) dt_k - \sum_{1 \leq i < j \leq k} c_{ij} a_{ij}(x_i, x_j) \int_0^{x_i} \int_0^{x_j} \lambda_i(u_i) \lambda_j(u_j) du_i du_j \right] \quad (25)$$

for any set of continuous functions $a_{ij}(x_i, x_j)$ all satisfying $0 \leq a_{ij}(x_i, x_j) \leq 1$.

For example, one may choose $a_{ij}(x_i, x_j) = \exp[-b_{ij}x_i x_j]$ with a given set of real nonnegative constants $[b_{ij}]$.

However, if the hypothesis holds, model (24) seems to be the **most natural** in the class of k -dimensional survival functions in the continuous case. We expect many applications of (24) in multivariate survival analysis.

Although we do not have at our disposal a formal proof of the considered Hypothesis, an argument that may support it is that it holds, in particular, for k

= 2 (Section 2) and for $k = 3$ (next section). So the truthfulness of the hypothesis (for all k) is (unfortunately) based solely on that analogy. Nonetheless, we are strongly convinced it holds in all cases $k \geq 2$. In a case of occurrence an essential difficulty in finding a formal proof for dimensions higher than 3, in applications (only), some statistical arguments for the cases $k = 4, 5, \dots$ may possibly be applied. Another way out might be the use of CAS (Computer Algebra Systems) such as MAPLE or MATEMATICA for underlying computations.

In the case when possessing a k -variate “model” ($k \geq 4$) such as (21) is especially important for a given practical purpose one might, eventually, take a (slight) risk and adopt some model (21) together with condition (23), and then try to verify it statistically. This approach may turn out to be useful, at least from a practical point of view. However, such “purely statistical” arguments would not be equivalent to a proper analytical (mathematical) proof.

4. 3-Variate Survival Functions. A Closer Look

4.1. The Continuous “Canonical” Case

Consider the following 3-dimensional version of (24):

$$\begin{aligned}
 & S(x_1, x_2, x_3) \\
 &= \exp \left[- \int_0^{x_1} \lambda_1(t_1) dt_1 - \int_0^{x_2} \lambda_2(t_2) dt_2 - \int_0^{x_3} \lambda_3(t_3) dt_3 - c_{12} \int_0^{x_1} \int_0^{x_2} \lambda_1(u_1) \lambda_2(u_2) du_1 du_2 \right. \\
 & \quad \left. - c_{13} \int_0^{x_1} \int_0^{x_3} \lambda_1(u_1) \lambda_3(u_3) du_1 du_3 - c_{23} \int_0^{x_2} \int_0^{x_3} \lambda_2(u_2) \lambda_3(u_3) du_2 du_3 \right] \quad (26)
 \end{aligned}$$

where we assume that all three continuous hazard rates $\lambda_1(x_1), \lambda_2(x_2), \lambda_3(x_3)$ are never zero. This assumption may, possibly, be weakened. We adopt it only for simplification of our calculations.

The question now to be answered is, for which hazard rates and for which values of the constant parameters c_{12}, c_{13}, c_{23} does expression (26) represent a valid 3-variate survival function.

The answer to this question, together with proper restrictions, we formulate as the following:

Theorem 1. Given is a random vector (X_1, X_2, X_3) whose univariate marginal survival functions are represented by any given continuous hazard rates $\lambda_1(x_1), \lambda_2(x_2), \lambda_3(x_3)$, which never are zero.

The function $S(x_1, x_2, x_3)$ defined by formula (26) is a valid joint survival function of the random vector (X_1, X_2, X_3) , if the nonnegative coefficients c_{ij} ($1 \leq i < j \leq 3$) in (26) satisfy: $c_{12} + c_{13} + c_{23} \leq 1$.

Proof

First realize, that all of the three implicitly present in (26) candidates $G_{ij}(x_i, x_j)$ for the joiners are in the form

$$G_{ij}(x_i, x_j) = \exp \left[-c_{ij} \int_0^{x_i} \int_0^{x_j} \lambda_i(u_i) \lambda_j(u_j) du_i du_j \right] \quad (27)$$

for all $1 \leq i < j \leq 3$.

This form of the *functions* $G_{ij}(\cdot)$, whose arguments are only present as upper limits of the involved integrals, indicates that for all pairs (i, j) the following necessary conditions for the functions $G_{ij}(x_i, x_j)$ to be legitimate joiners

$$G_{ij}(0, x_j) = G_{ij}(x_i, 0) = 1, \text{ for all } x_i, x_j \tag{28}$$

are satisfied by elementary properties of Riemann integrals.

Therefore, we may return to our original notation for the joiners, *i.e.*, $G_{ij}(x_i, x_j) = J_{ij}(x_i, x_j)$.

Now consider the sufficient condition for (26) to be a survival function.

Since for the, here considered, “continuous case” the derivative $\partial^3/\partial x_1 \partial x_2 \partial x_3 S(x_1, x_2, x_3)$ exists and is continuous, the sufficient condition reduces to the following inequality

$$(-1) \partial^3/\partial x_1 \partial x_2 \partial x_3 S(x_1, x_2, x_3) \geq 0, \tag{29}$$

which must hold for all the triples $(x_1, x_2, x_3) \in R_+^3$.

Recall, $S(x_1, x_2, x_3)$ is determined by the right hand side of (26).

Now, we need to find proper additional conditions for the hazard rates $\lambda_1(x_1), \lambda_2(x_2), \lambda_3(x_3)$ and for the coefficients c_{12}, c_{13}, c_{23} in order for inequality (29) to hold.

After calculating the derivative $\partial^3/\partial x_1 \partial x_2 \partial x_3 (\cdot)$ from the right-hand side of (26) we set inequality equivalent to (29). Then, we simplify it by dividing both sides of the so obtained inequality by the common (always positive) factor

$$\begin{aligned} & \lambda_1(x_1) \lambda_2(x_2) \lambda_3(x_3) \cdot \exp \left[- \int_0^{x_1} \lambda_1(t_1) dt_1 - \int_0^{x_2} \lambda_2(t_2) dt_2 - \int_0^{x_3} \lambda_3(t_3) dt_3 \right. \\ & - c_{12} \int_0^{x_1} \int_0^{x_2} \lambda_1(u_1) \lambda_2(u_2) du_1 du_2 - c_{13} \int_0^{x_1} \int_0^{x_3} \lambda_1(u_1) \lambda_3(u_3) du_1 du_3 \\ & \left. - c_{23} \int_0^{x_2} \int_0^{x_3} \lambda_2(u_2) \lambda_3(u_3) du_2 du_3 \right] \end{aligned}$$

and set all the negative terms to the right-hand side of the inequality.

As a result we obtain an inequality equivalent to (29) as follows:

$$\begin{aligned} & \left[1 + c_{13} \int_0^{x_1} \lambda_1(u_1) du_1 + c_{23} \int_0^{x_2} \lambda_2(u_2) du_2 \right] \left[1 + c_{12} \int_0^{x_1} \lambda_1(u_1) du_1 + c_{23} \int_0^{x_3} \lambda_3(u_3) du_3 \right] \\ & \left[1 + c_{12} \int_0^{x_2} \lambda_2(u_2) du_2 + c_{13} \int_0^{x_3} \lambda_3(u_3) du_3 \right] \\ & \geq c_{12} \left[1 + c_{13} \int_0^{x_1} \lambda_1(u_1) du_1 + c_{23} \int_0^{x_2} \lambda_2(u_2) du_2 \right] \\ & + c_{13} \left[1 + c_{12} \int_0^{x_1} \lambda_1(u_1) du_1 + c_{23} \int_0^{x_3} \lambda_3(u_3) du_3 \right] \\ & + c_{23} \left[1 + c_{12} \int_0^{x_2} \lambda_2(u_2) du_2 + c_{13} \int_0^{x_3} \lambda_3(u_3) du_3 \right] \end{aligned} \tag{30}$$

(At this point realize that if for some $x_i \geq 0$ ($i = 1, 2, 3$), one admits $\lambda_i(x_i) = 0$, then at those x_i the inequality equivalent to (29) and (30) holds trivially as $0 \geq 0$.)

Inequality (30) reduces to a simple relation upon the following substitutions:

$$\begin{aligned}
 A &= c_{13} \int_0^{x_1} \lambda_1(u_1) du_1 + c_{23} \int_0^{x_2} \lambda_2(u_2) du_2 \\
 B &= c_{12} \int_0^{x_1} \lambda_1(u_1) du_1 + c_{23} \int_0^{x_3} \lambda_3(u_3) du_3 \\
 C &= c_{12} \int_0^{x_2} \lambda_2(u_2) du_2 + c_{13} \int_0^{x_3} \lambda_3(u_3) du_3
 \end{aligned}$$

Now (30) can be expressed as:

$$(1 + A)(1 + B)(1 + C) \geq c_{12}(1 + A) + c_{13}(1 + B) + c_{23}(1 + C) \tag{31}$$

where all the expressions A, B, C and all the constants c_{12}, c_{13}, c_{23} are nonnegative. Also, A, B, C are increasing functions of the variables x_1, x_2, x_3 , while they all are 0 at $(x_1, x_2, x_3) = (0, 0, 0)$.

It is easy to solve inequality (30) since it is equivalent to inequality (31). (31) is quite easy to solve and get to the conclusion that when it holds, (30) does too whenever the nonnegative constants c_{ij} satisfy:

$$c_{12} + c_{13} + c_{23} \leq 1. \tag{32}$$

The very important conclusion from the above form (31) of (30) and its set of solutions (given by (32)) is that both inequalities' truthfulness does not depend on either values or functional forms of the involved three continuous hazard rates $\lambda_1(x_1), \lambda_2(x_2), \lambda_3(x_3)$. Obviously, only if we admit the possibility that some *random events* ($X_i = +\infty$) ($i = 1, 2, 3$) may happen with positive probabilities; otherwise, the only additional assumption for the hazard rates we need to adopt is that, for each $i = 1, 2, 3$

$$\int_0^{\infty} \lambda_i(u_i) du_i = +\infty.$$

This ends the proof. \square

Thus, in the continuous case **all** triples of marginal distributions $(S_1(x_1), S_2(x_2), S_3(x_3))$ represented by the corresponding hazard rates, "form" their **(natural) joint survival function** (26) whenever (32) holds.

Recall, the necessary conditions (28) are a direct consequence of joiners' definition (27) for the continuous case.

Condition (32) is also necessary for (29) [or (30)] to hold. Otherwise, at the point $(x_1, x_2, x_3) = (0, 0, 0)$ or just at points such as, for example, $(0, x_2, x_3)$, (31) will be violated, and, by the continuity argument, this situation will persist in an open neighborhood of any such point. That would imply (31) was not true on a set of a positive Lebesgue measure in R_+^3 .

At the end of this subsection notice that setting any of the variables x_1, x_2, x_3 to 0, one directly obtains the bivariate survival function with respect to remaining two.

For example setting $x_3 = 0$ in (26) one obtains (9), with $c_{12} = a$.

4.2. A More General Continuous Case

One of the most general forms of “continuous” 3-dimensional survival functions, which also comprise the case considered in Section 4.1, can be defined as follows:

$$\begin{aligned}
 & S^g(x_1, x_2, x_3) \\
 &= \exp \left[- \int_0^{x_1} \lambda_1(t_1) dt_1 - \int_0^{x_2} \lambda_2(t_2) dt_2 - \int_0^{x_3} \lambda_3(t_3) dt_3 - c_{12} \int_0^{x_1} \int_0^{x_2} \Psi_{1,2}(u_1, u_2) du_1 du_2 \right. \\
 & \quad \left. - c_{13} \int_0^{x_1} \int_0^{x_3} \Psi_{1,3}(u_1, u_3) du_1 du_3 - c_{23} \int_0^{x_2} \int_0^{x_3} \Psi_{2,3}(u_2, u_3) du_2 du_3 \right] \tag{33}
 \end{aligned}$$

where, for all $1 \leq i < j \leq 3$, $\Psi_{ij}(u_i, u_j)$ are arbitrary continuous functions, and $0 \leq \Psi_{ij}(u_i, u_j) \leq \lambda_i(u_i) \lambda_j(u_j)$. (34)

Moreover, (32) holds.

We assumed nonnegativity of all $\Psi_{ij}(u_i, u_j)$, which, as in the case of the models (26), restricts the considerations to positive [but, possibly, to all positive] stochastic dependencies only. As a matter of fact that nonnegativity assumption is not necessary, and is adopted only for simplicity reasons.

We will prove the following theorem:

Theorem 2

Given that (32) and (34) hold, formula (33) defines a class of valid 3-variate survival functions.

Proof. The argumentation is mainly based on the already proven validity of (26) as defining survival functions. As in the case (26), we need to check if the following inequality (similar to (29)) holds:

$$(-1) \partial^3 / \partial x_1 \partial x_2 \partial x_3 S^g(x_1, x_2, x_3) \geq 0, \tag{35}$$

where $S^g(x_1, x_2, x_3)$ is given by (33).

After similar computations as in the previous case we arrive at the following inequality equivalent to (35):

$$\begin{aligned}
 & \left[\lambda_1(x_1) + c_{12} \int_0^{x_2} \Psi_{1,2}(x_1, u_2) du_2 + c_{13} \int_0^{x_3} \Psi_{1,3}(x_1, u_3) du_3 \right] \\
 & \left[\lambda_2(x_2) + c_{12} \int_0^{x_1} \Psi_{1,2}(u_1, x_2) du_1 + c_{23} \int_0^{x_3} \Psi_{2,3}(x_2, u_3) du_3 \right] \\
 & \left[\lambda_3(x_3) + c_{13} \int_0^{x_1} \Psi_{1,3}(u_1, x_3) du_1 + c_{23} \int_0^{x_2} \Psi_{2,3}(u_2, x_3) du_2 \right] \\
 & \geq c_{23} \Psi_{2,3}(x_2, x_3) \left[\lambda_1(x_1) + c_{12} \int_0^{x_2} \Psi_{1,2}(x_1, u_2) du_2 + c_{13} \int_0^{x_3} \Psi_{1,3}(x_1, u_3) du_3 \right] \\
 & + c_{13} \Psi_{1,3}(x_1, x_3) \left[\lambda_2(x_2) + c_{12} \int_0^{x_1} \Psi_{1,2}(u_1, x_2) du_1 + c_{23} \int_0^{x_3} \Psi_{2,3}(x_2, u_3) du_3 \right] \\
 & + c_{12} \Psi_{1,2}(x_1, x_2) \left[\lambda_3(x_3) + c_{13} \int_0^{x_1} \Psi_{1,3}(u_1, x_3) du_1 + c_{23} \int_0^{x_2} \Psi_{2,3}(u_2, x_3) du_2 \right] \tag{36}
 \end{aligned}$$

Now, all that remains to do is to prove inequality (36).

Using the assumption of positivity of the functions $\lambda_1(x_1), \lambda_2(x_2), \lambda_3(x_3)$, inequality (36) can be transformed into the following equivalent form:

$$\begin{aligned} & \lambda_1(x_1) \left[1 + \left(c_{12} \int_0^{x_2} \Psi_{1,2}(x_1, u_2) du_2 + c_{13} \int_0^{x_3} \Psi_{1,3}(x_1, u_3) du_3 \right) / \lambda_1(x_1) \right] \\ & \lambda_2(x_2) \left[1 + \left(c_{12} \int_0^{x_1} \Psi_{1,2}(u_1, x_2) du_1 + c_{23} \int_0^{x_3} \Psi_{2,3}(x_2, u_3) du_3 \right) / \lambda_2(x_2) \right] \\ & \lambda_3(x_3) \left[1 + \left(c_{13} \int_0^{x_1} \Psi_{1,3}(u_1, x_3) du_1 + c_{23} \int_0^{x_2} \Psi_{2,3}(u_2, x_3) du_2 \right) / \lambda_3(x_3) \right] \\ & \geq c_{23} \Psi_{2,3}(x_2, x_3) \lambda_1(x_1) \left[1 + \left(c_{12} \int_0^{x_2} \Psi_{1,2}(x_1, u_2) du_2 + c_{13} \int_0^{x_3} \Psi_{1,3}(x_1, u_3) du_3 \right) / \lambda_1(x_1) \right] \\ & + c_{13} \Psi_{1,3}(x_1, x_3) \lambda_2(x_2) \left[1 + \left(c_{12} \int_0^{x_1} \Psi_{1,2}(u_1, x_2) du_1 + c_{23} \int_0^{x_3} \Psi_{2,3}(x_2, u_3) du_3 \right) / \lambda_2(x_2) \right] \\ & + c_{12} \Psi_{1,2}(x_1, x_2) \lambda_3(x_3) \left[1 + \left(c_{13} \int_0^{x_1} \Psi_{1,3}(u_1, x_3) du_1 + c_{23} \int_0^{x_2} \Psi_{2,3}(u_2, x_3) du_2 \right) / \lambda_3(x_3) \right] \end{aligned} \tag{37}$$

From (34) we have that:

$$\begin{aligned} 0 & \leq \Psi_{2,3}(x_2, x_3) \lambda_1(x_1) / \lambda_1(x_1) \lambda_2(x_2) \lambda_3(x_3) = b_1(x_1, x_2, x_3) \leq 1, \\ 0 & \leq \Psi_{1,3}(x_1, x_3) \lambda_2(x_2) / \lambda_1(x_1) \lambda_2(x_2) \lambda_3(x_3) = b_2(x_1, x_2, x_3) \leq 1, \\ 0 & \leq \Psi_{1,2}(x_1, x_2) \lambda_3(x_3) / \lambda_1(x_1) \lambda_2(x_2) \lambda_3(x_3) = b_3(x_1, x_2, x_3) \leq 1. \end{aligned} \tag{38}$$

Also let us make the following substitutions:

$$\begin{aligned} & \left(c_{12} \int_0^{x_2} \Psi_{1,2}(x_1, u_2) du_2 + c_{13} \int_0^{x_3} \Psi_{1,3}(x_1, u_3) du_3 \right) / \lambda_1(x_1) = A^* \\ & \left(c_{12} \int_0^{x_1} \Psi_{1,2}(u_1, x_2) du_1 + c_{23} \int_0^{x_3} \Psi_{2,3}(x_2, u_3) du_3 \right) / \lambda_2(x_2) = B^* \\ & \left(c_{13} \int_0^{x_1} \Psi_{1,3}(u_1, x_3) du_1 + c_{23} \int_0^{x_2} \Psi_{2,3}(u_2, x_3) du_2 \right) / \lambda_3(x_3) = C^* \end{aligned} \tag{39}$$

where from the first inequalities in (34) [for all $1 \leq i < j \leq 3$] follows the nonnegativity of A^*, B^*, C^* .

Now, upon dividing both sides of (37) by the (always positive) product $\lambda_1(x_1) \lambda_2(x_2) \lambda_3(x_3)$ and applying substitutions (38) and (39), inequality (37) can be rewritten into the following equivalent form:

$$\begin{aligned} & (1 + A^*)(1 + B^*)(1 + C^*) \\ & \geq c_{23} b_1(x_1, x_2, x_3) (1 + A^*) + c_{13} b_2(x_1, x_2, x_3) (1 + B^*) \\ & + c_{12} b_3(x_1, x_2, x_3) (1 + C^*) \end{aligned} \tag{40}$$

The equivalence of (40) with (36) and (37) does not depend on the always nonnegative values of the, given by (39), expressions for A^*, B^*, C^* , even if they

differ from the expressions A, B, C present in (31) (in both inequalities (31) and (40) only nonnegativity of these symbols is relevant).

Suppose, we set in (40) $b_1(x_1, x_2, x_3) = b_2(x_1, x_2, x_3) = b_3(x_1, x_2, x_3) = 1$ for all x_1, x_2, x_3 . Then inequality (40) is essentially “the same” as inequality (31), and the common sufficient condition for them to hold is (32), now as applied to (36).

Assume that (32) holds. Then, the direction of inequality (40) must be preserved also when some or all $b_i(x_1, x_2, x_3)$ satisfy $b_i(x_1, x_2, x_3) < 1$, ($i = 1, 2, 3$). Therefore, (36) and so (35) holds, and the only sufficient condition for that is the logical conjunction of (32) and (34). This terminates the proof. \square

Thus, as was proven above, formula (33) satisfying conditions (32) and (34), represents the class of well-defined 3-variate survival functions.

Notice, the wide generality of that class of stochastic models even if it “only” comprises the “continuous” cases.

Notice too, that a vast majority of applications (such as reliability, for example) of multivariate survival analysis mostly deals with the “continuous” case.

As one can say, scheme (33) also provides a method for construction of a wide variety of tri-variate models for many practical situations. This method only relies on choosing (for all $1 \leq i < j \leq 3$), three proper 2-argument nonnegative functions, say $\Psi_{i,j}(x_i, x_j) \leq \lambda_i(x_i)\lambda_j(x_j)$, given the marginal distributions, in the continuous case represented by the hazard rates $\lambda_1(x_1), \lambda_2(x_2), \lambda_3(x_3)$.

Obviously, choices of the functions $\Psi_{i,j}(x_i, x_j)$ are dictated by the underlying “physical” structure of the modeled reality (recall that in this paper the word “physical” is meant to have a very general meaning including biological, social, and financial). For some guidance in choosing proper functions $\Psi_{i,j}(x_i, x_j)$, one may resort to the roots of such models which lie in the Aalen version of the Cox model [2] [16] [17]. Recall the close relations between our models and the Cox-Aalen approach to stochastic dependence, see [2].

In order yet to simplify such procedure of construction, we also may, starting with the case $\Psi_{i,j}(x_i, x_j) = \lambda_i(x_i)\lambda_j(x_j)$, preserve the separation of the variables x_i, x_j by choosing: $\Psi_{i,j}(x_i, x_j) = g_i(x_i)g_j(x_j)$ for some $0 \leq g_i(x_i) \leq \lambda_i(x_i)$ and $0 \leq g_j(x_j) \leq \lambda_j(x_j)$.

That foregoing assumption may facilitate the construction still preserving a significant amount of generality of the so obtained models.

Choosing different possibilities one can obtain, say up to ten or more versions of scheme (33) as the proper model’s candidates.

The next step in the modeling procedure is statistical verification of the admitted analytical models by testing their fit to the given data, after estimating all the underlying parameters [by the method of maximum likelihood estimates, for example]. Then, of course, we choose the one with the best fit to the given data. Another criterion for the right choice can be the degree of the model’s simplicity, sometimes dictated by the number of underlying parameters to be estimated.

However, the second, **statistical, part** of the modeling procedure is out of scope of this work, and is left as a set of open problems for future research.

5. Conclusions

The here and in [2] considered approach to the problem of construction of k -variate ($k \geq 3$) survival functions, given all the univariate marginals, turned out to be very fruitful. First of all, the form of the k -variate model, with an arbitrary k , as obtained in [2] is universal in the sense that every k -variate survival function is expressible in that way.

In this work we achieved the following:

1) The most general form of any k -variate model as obtained in [2] is a bit complicated as it describes all possible stochastic dependencies that might be encountered by means of all the joiners.

In this work, we simplified that model dramatically by assuming all the r -dimensional ($r \geq 3$) joiners are equal to 1 (this case was named the 3-independence). That yields a quite simple, nice, and still realistic form of the k -variate survival functions as given by formula (14). For the so defined “continuous” case this formula takes on the exponential form (21) which is easier for further analysis.

2) Formulas such as (14) and (21) only determine the form of k -variate models and not yet an actual model. To find a legitimate “active” model one needs to impose proper conditions on all the underlying functions $\Psi_{i,j}(x_i, x_j)$ and the real coefficients c_{ij} . The conditions for the general case were stated in the Hypothesis in Section 3. The formal proof was not explicitly given although the Hypothesis appears to be very convincing.

3) The conditions mentioned above were, however, found in Section 4 for the case $k = 3$ (in Section 2 the corresponding conditions when $k = 2$ were provided too). These were formulated as Theorem 1 and Theorem 2. The model found in Theorem 2 is very general and comprises every tri-variate “continuous” model such that the corresponding functions are nonnegative (positive dependence only) and the tri-variate joiners reduce to 1 (the bi-dependence case). The case in which the tri-variate joiner is nontrivial is postponed to our next publication. The special case of the model presented in Theorem 2 is given by Theorem 1. This “limit” case of the general one seems to be very natural, and one may expect many applications of it in practical situations.

4) The underlying **statistical problems** of the parameters estimation and testing the model’s fit to given data are a very important part of the emerging theory, but we consider it to be out of scope for our current interest.

Two other aspects of the created theory are the following:

5) A pretty close relation of the joiner theory to the **copula** methodology [9] is apparent. The joiner theory is not only competitive to the theory of copulas but, as it appears, it may become a part of a common theory of joiners and copulas. Having a joiner one can easily find the corresponding unique copula and every copula determines a corresponding unique joiner. Moreover, the joiner method may highly develop the copula approach in cases of dimensions higher than 2.

6) As mentioned in the Remark in Section 3, the joiners theory is not to be re-

stricted to investigation of distributions of k -dimensional random vectors only. As it was pointed out in that Remark, the method can as well be applied to the construction of **stochastic processes** with discrete time (the Kolmogorov-Daniels consistency conditions are always satisfied). The so constructed processes may possess very nice properties such as m -Markovianity (the Markovianity when $m = 1$) and stationarity. Easiness of such constructions indicates the significant value of the created “joiners theory”.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Filus, K. and Filus, L.Z. (2020) General Forms of Bi-Variate Survival Functions with Reliability Applications. Under Review for *Handbook of Reliability, Maintenance, and System Safety through Mathematical Modeling*.
- [2] Filus, J.K. and Filus, L.Z. (2020) Theoretical and Reliability Aspects of Multivariate Probability Distributions in Their Universal Form. In: Pham, H., Ed., *Springer Handbook of Engineering Statistics*, 2nd Edition.
- [3] Filus, J.K. and Filus, L.Z. (2020) A General (Universal) Form of Multivariate Survival Functions in Theoretical and Modeling Aspect of Multicomponent System Reliability Analysis. In: Ram, M. and Pham, H., Eds., *Advances in Reliability Analysis and its Applications. Springer Series in Reliability Engineering*, Springer, Cham, 319-342. https://doi.org/10.1007/978-3-030-31375-3_10
- [4] Filus, J.K. and Filus, L.Z. (2017) The Cox-Aalen Models as Framework for Construction of Bivariate Probability Distributions, Universal Representation. *Journal of Statistical Science and Applications*, **5**, 56-63. <https://doi.org/10.17265/2328-224X/2017.0304.002>
- [5] Freund, J.E. (1961) A Bivariate Extension of the Exponential Distribution. *Journal of the American Statistical Association*, **56**, 971-77. <https://doi.org/10.1080/01621459.1961.10482138>
- [6] Gumbel, E.J. (1960) Bivariate Exponential Distributions. *Journal of the American Statistical Association*, **55**, 698-707. <https://doi.org/10.1080/01621459.1960.10483368>
- [7] Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000) Continuous Multivariate Distributions. Vol. 1, 2nd Edition, John Wiley & Sons, Inc., New York. <https://doi.org/10.1002/0471722065>
- [8] Marshall, A.W. and Olkin, I. (1967) A Generalized Bivariate Exponential Distribution. *Journal of Applied Probability*, **4**, 291-303. <https://doi.org/10.2307/3212024>
- [9] Sklar, A. (1959) Fonctions de repartition a n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Universite de Paris*, **8**, 229-231.
- [10] Arnold, B.C., Castillo, E. and Sarabia, J.M. (2001) Conditionally Specified Distributions: An Introduction (with Discussion). *Statistical Science*, **16**, 249-274.
- [11] Arnold, B.C., Castillo, E. and Sarabia, J.M. (1999) Conditional Specification of Statistical Models. In: *Springer Series in Statistics*, Springer Verlag, New York.
- [12] Castillo, E. and Galambos, J. (1990) Bivariate Distributions with Weibull Condi-

- tionals. *Analysis Mathematica*, **16**, 3-9. <https://doi.org/10.1007/BF01906769>
- [13] Filus, J.K., Filus, L.Z., Arnold, B.C., Jordanova, P.K., Nunez Soza, L., Lu, Y., Stehlikova, S. and Stehlik, M. (2018) On Parameter Dependence and Related Topics: The Impact of Jerzy Filus from Genesis to Recent Developments. In: Vonta, I. and Ram, M., Eds., *Reliability Engineering: Theory and Applications*, CRC Press, Boca Raton, 143-172. <https://doi.org/10.1201/9781351130363-8>
- [14] Filus, J.K. and Filus, L.Z. (2013) A Method for Multivariate Probability Distributions Construction via Parameter Dependence. *Communications in Statistics—Theory and Methods*, **42**, 716-721. <https://doi.org/10.1080/03610926.2012.731549>
- [15] Filus, J.K., Filus, L.Z. and Arnold, B.C. (2010) Families of Multivariate Distributions Involving “Triangular” Transformations. *Communications in Statistics—Theory and Methods*, **39**, 107-116. <https://doi.org/10.1080/03610920802687793>
- [16] Aalen, O.O. (1989) A Linear Regression Model for the Analysis of the Life Times. *Statistics in Medicine*, **8**, 907-925. <https://doi.org/10.1002/sim.4780080803>
- [17] Cox, D.R. (1972) Regression Models and Life Tables (with Discussion). *Journal of the Royal Statistical Society B*, **74**, 187-220. <https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>
- [18] Arnold, B.C. (2017) Private Communication. December 2017.