Adv. Appl. Prob. **47,** 1157–1174 (2015)
Printed in Northern Ireland

ON DYNAMIC MUTUAL INFORMATION FOR BIVARIATE LIFETIMES

JAFAR AHMADI,* *Ferdowsi University of Mashhad*
ANTONIO DI CRESCENZO,** *Università degli Studi di Salerno*
MARIA LONGOBARDI,*** *Università di Napoli Federico II*

Abstract

We consider dynamic versions of the mutual information of lifetime distributions, with a focus on past lifetimes, residual lifetimes, and mixed lifetimes evaluated at different instants. This allows us to study multicomponent systems, by measuring the dependence in conditional lifetimes of two components having possibly different ages. We provide some bounds, and investigate the mutual information of residual lifetimes within the time-transformed exponential model (under both the assumptions of unbounded and truncated lifetimes). Moreover, with reference to the order statistics of a random sample, we evaluate explicitly the mutual information between the minimum and the maximum, conditional on inspection at different times, and show that it is distribution-free in a special case. Finally, we develop a copula-based approach aiming to express the dynamic mutual information for past and residual bivariate lifetimes in an alternative way.

Keywords: Entropy; mutual information; time-transformed exponential model; bivariate lifetimes; order statistics; copula

2010 Mathematics Subject Classification: Primary 94A17
Secondary 62N05; 60E99

1. Introduction and background

Information measures are largely used in applied contexts in order to describe useful notions related to stochastic models. The problem of measuring the information content in a dynamic setting arises in various fields, such as survival analysis, reliability, and mathematical finance, for example. Significant results in this area have been provided in Ebrahimi *et al.* [14], where the focus was directed on the joint, marginal, and conditional entropies, and the mutual information for residual life distributions in multivariate settings. In this paper we provide some further insight on the dynamic mutual information, with reference to past lifetimes, residual lifetimes, and mixed lifetimes evaluated at different ages.

In probability theory the mutual information of two random variables is a measure of their mutual dependence, and can be evaluated by means of the joint and marginal distributions. See Ebrahimi *et al.* [18] for a contribution dealing with the mutual information of certain classes of bivariate distributions, and Arellano-Valle *et al.* [2] for a recent investigation on the mutual information of multivariate skew-elliptical distributions. Other kinds of multivariate

Received 28 May 2014; revision received 13 November 2014.
* Postal address: Department of Statistics, Ordered and Spatial Data Center of Excellence, Ferdowsi University of Mashhad, P.O. Box 1159, Mashhad, 91775, Iran. Email address: ahmadi-j@um.ac.ir
** Postal address: Dipartimento di Matematica, Università degli Studi di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy. Email address: adicrescenzo@unisa.it
*** Postal address: Dipartimento di Matematica e Applicazioni, Università di Napoli Federico II, Via Cintia, 80126 Napoli, Italy. Email address: maria.longobardi@unina.it

information measures have been investigated by Ebrahimi *et al.* [15]. We also mention that a nonparametric and binless estimator for the mutual information of a d-dimensional random vector has been proposed recently by Giraudo *et al.* [19].

In view of suitable applications in the context of reliability theory, in this paper we consider both the dynamic extensions of the mutual information and the related entropies. Specifically, we aim to study the applications of mutual information to the cases of past, residual, and mixed distributions. In Section 2 we briefly recall the relevant mathematical concepts related to mutual information and entropy, and then introduce the bivariate distributions describing two lifetimes conditional on possibly different inspection times. In Section 3 we introduce the dynamic mutual information of past lifetimes. We obtain a bound for such a measure, which is suitable to describe stochastic models whose uncertainty is related to the past. Section 4 is concerning the mutual information of residual lifetimes. We provide a bound and a connection between past and residual mutual information. We also investigate such a measure within the time-transformed exponential model (both in the classical case of unbounded lifetimes and in the new setting involving truncated lifetimes). In Section 5 we study the dynamic mutual information for mixed lifetimes and apply it to ordered data. With reference to the order statistics $X_{i:n}$, $i = 1, 2, \dots, n$, we evaluate explicitly the mutual information between the minimum and the maximum $(X_{1:n}, X_{n:n})$ conditional on $(X_{1:n} \le s, X_{n:n} > t)$ for $s < t$ and show that it is distribution-free in a special case. This also allows us to describe the information content in n-component systems inspected at two different times. Finally, in Section 6 we discuss a copula-based approach, which allows us to express the dynamic mutual information for past and residual bivariate lifetimes in terms of copula and survival copula, respectively.

Throughout the paper we denote by $[Z \mid B]$ a random variable or a random vector whose distribution is identical to the conditional distribution of Z given B. Moreover, primes denote derivatives.

2. Preliminaries

Let (X, Y) be a random vector, where X and Y are nonnegative absolutely continuous random variables. We denote by $f(x, y)$ the joint probability density function (PDF) of (X, Y), and by $f_X(x)$ and $f_Y(y)$ the marginal densities of X and Y, respectively. It is well known that the mutual information of X and Y is defined as

$$M_{X,Y} = \int_0^{+\infty} \mathrm{d}x \int_0^{+\infty} f(x, y) \log \frac{f(x, y)}{f_X(x) f_Y(y)} \, \mathrm{d}y, \tag{1}$$

where 'log' means natural logarithm. The term $M_{X,Y}$ is a measure of dependence between X and Y. Indeed, (1) defines a premetric, since $M_{X,Y} \ge 0$, with $M_{X,Y} = 0$ if and only if X and Y are independent. Roughly speaking, it measures how far X and Y are from being independent, in the sense that high values of $M_{X,Y}$ correspond to a strong dependence between X and Y. Moreover, $M_{X,Y}$ is in general finite and is invariant under linear transformations. We recall that the mutual information can be expressed in terms of entropies as follows (see, for example, Ebrahimi *et al.* [16]):

$$M_{X,Y} = H_X + H_Y - H_{X,Y}, \tag{2}$$

where H_X is the differential entropy of X, defined by $H_X = -\int_0^{+\infty} f_X(x) \log f_X(x) \, \mathrm{d}x$, H_Y is similarly defined, and

$$H_{X,Y} = -\int_0^{+\infty} \mathrm{d}x \int_0^{+\infty} f(x, y) \log f(x, y) \, \mathrm{d}y \tag{3}$$

is the differential entropy of (X, Y). We recall that H_X measures the 'uniformity' of the distribution of X, i.e. how the distribution spreads over its domain, and is irrespective of the locations of concentration. High values of H_X correspond to a low concentration of the probability mass of X.

The reliability analysis of a system composed of two items involves the general setting by which they are inspected at possibly different times s and t. Assuming that the random variables X and Y describe the failure times of the two items, the following conditional random vectors thus deserve interest, for $s, t \geq 0$,

$$[(X, Y) \mid X \leq s, Y \leq t] \quad \text{if both items failed before inspection,} \tag{4}$$

$$[(X, Y) \mid X > s, Y > t] \quad \text{if no item failed before inspection,} \tag{5}$$

$$[(X, Y) \mid X \leq s, Y > t] \quad \text{if only the first item failed before inspection,} \tag{6}$$

$$[(X, Y) \mid X > s, Y \leq t] \quad \text{if only the second item failed before inspection.} \tag{7}$$

The probability of the conditional events considered above will be denoted as

$$F(s, t) = \mathbb{P}(X \leq s, Y \leq t), \qquad \overline{F}(s, t) = \mathbb{P}(X > s, Y > t),$$
$$F^{-,+}(s, t) = \mathbb{P}(X \leq s, Y > t), \qquad F^{+,-}(s, t) = \mathbb{P}(X > s, Y \leq t),$$

so that $F(s, t) + \overline{F}(s, t) + F^{-,+}(s, t) + F^{+,-}(s, t) = 1$. In order to introduce certain dynamic entropies, we now consider the following functions.

(i) The density of $[(X, Y) \mid X \leq s, Y \leq t]$ for all $s, t \geq 0$ such that $F(s, t) > 0$,

$$\tilde{f}_{X,Y}(x, y; s, t) = \frac{f(x, y)}{F(s, t)}, \qquad 0 \leq x \leq s,\ 0 \leq y \leq t. \tag{8}$$

(ii) The density of $[(X - s, Y - t) \mid X > s, Y > t]$ for all $s, t \geq 0$ such that $\overline{F}(s, t) > 0$,

$$f_{X,Y}(x, y; s, t) = \frac{f(x + s, y + t)}{\overline{F}(s, t)}, \qquad x \geq 0,\ y \geq 0. \tag{9}$$

(iii) The density of $[(X, Y - t) \mid X \leq s, Y > t]$ for all $s, t \geq 0$ such that $F^{-,+}(s, t) > 0$,

$$f^{-,+}_{X,Y}(x, y; s, t) = \frac{f(x, y + t)}{F^{-,+}(s, t)}, \qquad 0 \leq x \leq s,\ y \geq 0.$$

(iv) The density of $[(X - s, Y) \mid X > s, Y \leq t]$ for all $s, t \geq 0$ such that $F^{+,-}(s, t) > 0$,

$$f^{+,-}_{X,Y}(x, y; s, t) = \frac{f(x + s, y)}{F^{+,-}(s, t)}, \qquad x \geq 0,\ 0 \leq y \leq t.$$

Hence, in analogy with (3) we can now introduce the following entropies, for $s, t \geq 0$:

$$\tilde{H}_{X,Y}(s, t) = -\int_0^s \mathrm{d}x \int_0^t \tilde{f}_{X,Y}(x, y; s, t) \log \tilde{f}_{X,Y}(x, y; s, t)\, \mathrm{d}y, \tag{10}$$

$$H_{X,Y}(s, t) = -\int_0^{+\infty} \mathrm{d}x \int_0^{+\infty} f_{X,Y}(x, y; s, t) \log f_{X,Y}(x, y; s, t)\, \mathrm{d}y, \tag{11}$$

$$H^{-,+}_{X,Y}(s, t) = -\int_0^s \mathrm{d}x \int_0^{+\infty} f^{-,+}_{X,Y}(x, y; s, t) \log f^{-,+}_{X,Y}(x, y; s, t)\, \mathrm{d}y,$$

$$H^{+,-}_{X,Y}(s, t) = -\int_0^{+\infty} \mathrm{d}x \int_0^t f^{+,-}_{X,Y}(x, y; s, t) \log f^{+,-}_{X,Y}(x, y; s, t)\, \mathrm{d}y.$$

Remark 1. The entropy (3) can be expressed in terms of the entropies given above; indeed, for all $s, t \geq 0$,

$$\begin{aligned} H_{X,Y} = \mathcal{H}[F(s,t), \overline{F}(s,t), F^{-,+}(s,t), F^{+,-}(s,t)] + F(s,t)\tilde{H}_{X,Y}(s,t) \\ + \overline{F}(s,t)H_{X,Y}(s,t) + F^{-,+}(s,t)H^{-,+}_{X,Y}(s,t) + F^{+,-}(s,t)H^{+,-}_{X,Y}(s,t), \end{aligned} \tag{12}$$

where $\mathcal{H}[p_1, \ldots, p_n] := -\sum_{i=1}^{n} p_i \log p_i$ denotes the entropy of a discrete probability distribution.

We recall that (12) is the two-dimensional analogue of [8, Proposition 2.1]. It holds due to the partitioning property of the Shannon entropy (see, for example, [14, Equation (24)] for another application of such a property). It expresses that the uncertainty about the failure times of two items can be decomposed in five terms. The first term conveys the uncertainty of whether the items failed before or after their inspection times, the other terms give the uncertainties about the failure times in the domains specified in (4)–(7), given that the items failed in the corresponding regions. Note that (12) is in agreement with some remarks provided in [14, Section 4.4].

We are now able to study the dynamic mutual information for the cases introduced in this section.

3. Mutual information for past lifetimes

In various contexts the uncertainty is not necessarily related to the future but may refer to the past. For instance, if a system is observed at an inspection time t and is found failed, then the uncertainty relies on the past, i.e. on which instant in $(0, t)$ it failed. Several papers have been devoted to the investigation of information measures concerning past lifetimes. We recall, for instance, the univariate past entropy defined in [8]. Some properties and generalizations have also been investigated in [21], [23], [25], and [26].

In this section we introduce the mutual information for the bivariate past lifetimes defined in (4). To this aim we consider the marginal past lifetimes

$$[X \mid X \leq s, Y \leq t], \qquad [Y \mid X \leq s, Y \leq t], \quad s, t \geq 0, \tag{13}$$

having PDFs

$$\tilde{f}_X(x; s, t) := \frac{1}{F(s,t)} \frac{\partial}{\partial x} F(x,t) = \frac{1}{F(s,t)} \int_0^t f(x,y)\,\mathrm{d}y, \qquad 0 \leq x \leq s, \tag{14}$$

$$\tilde{f}_Y(y; s, t) := \frac{1}{F(s,t)} \frac{\partial}{\partial y} F(s,y) = \frac{1}{F(s,t)} \int_0^s f(x,y)\,\mathrm{d}x, \qquad 0 \leq y \leq t, \tag{15}$$

for $s, t \geq 0$ such that $F(s, t) > 0$. In analogy with (1), we are now able to define the following new information measure, named *bivariate dynamic past mutual information*:

$$\tilde{M}_{X,Y}(s,t) := \int_0^s \mathrm{d}x \int_0^t \tilde{f}_{X,Y}(x, y; s, t) \log \frac{\tilde{f}_{X,Y}(x, y; s, t)}{\tilde{f}_X(x; s, t)\tilde{f}_Y(y; s, t)}\,\mathrm{d}y \tag{16}$$

for $s, t \geq 0$ such that $F(s, t) > 0$, where the involved densities are given in (8), (14), and (15). This is a nonnegative function which measures the dependence between the past lifetimes of X and Y conditional on $\{X \leq s, Y \leq t\}$.

Remark 2. Similarly to (2), for $s, t \geq 0$ the following identity holds:

$$\tilde{M}_{X,Y}(s,t) = \tilde{H}_X(s,t) + \tilde{H}_Y(s,t) - \tilde{H}_{X,Y}(s,t),$$

where

$$\tilde{H}_X(s,t) = -\int_0^s \tilde{f}_X(x; s,t) \log \tilde{f}_X(x; s,t)\, \mathrm{d}x$$

and

$$\tilde{H}_Y(s,t) = -\int_0^t \tilde{f}_Y(y; s,t) \log \tilde{f}_Y(y; s,t)\, \mathrm{d}y$$

are the entropies of the marginal past lifetimes introduced in (13), and where $\tilde{H}_{X,Y}(s,t)$ is defined in (10).

Let us now obtain some bounds.

Proposition 1. *For $s, t \geq 0$ such that $F(s,t) > 0$, let*

$$\tilde{a}(x,y; s,t) := \frac{f(x,y)}{\int_0^t f(x,y)\,\mathrm{d}y \int_0^s f(x,y)\,\mathrm{d}x}, \qquad 0 \leq x \leq s,\ 0 \leq y \leq t. \tag{17}$$

If

$$\tilde{a}(x,y; s,t) \leq (\geq)\ \tilde{a}(s,t; s,t) \quad \textit{for all } 0 \leq x \leq s \textit{ and } 0 \leq y \leq t,$$

then the following upper [lower] bound holds:

$$\tilde{M}_{X,Y}(s,t) \leq (\geq) \log \tilde{a}(s,t; s,t) + \log F(s,t). \tag{18}$$

Proof. From (16), making use of (8), (14), and (15), we have

$$\tilde{M}_{X,Y}(s,t) = \frac{1}{F(s,t)} \int_0^s \mathrm{d}x \int_0^t f(x,y) \log \tilde{a}(x,y; s,t)\,\mathrm{d}y + \log F(s,t). \tag{19}$$

Hence, from (17) we immediately obtain (18).

Example 1. Let (X, Y) be a random vector with joint PDF and distribution function

$$f(x,y) = x + y, \qquad F(x,y) = \frac{xy(x+y)}{2}, \quad 0 \leq x \leq 1,\ 0 \leq y \leq 1.$$

Since, for $0 < s < 1$ and $0 < t < 1$,

$$\tilde{a}(x,y; s,t) = \frac{4(x+y)}{st(t+2x)(s+2y)}, \qquad 0 \leq x \leq 1,\ 0 \leq y \leq 1,$$

from (19) we have

$$\begin{aligned}
\tilde{M}_{X,Y}(s,t) = \log \frac{st(s+t)}{2} \\
&+ \frac{1}{st(s+t)} \bigg\{ st(s+t) \log \frac{4}{st} \\
&\qquad + \frac{1}{6}[-2s^3 \log s - 2t^3 \log t + 2(s+t)^3 \log(s+t) - 5st(s+t)] \\
&\qquad + \frac{t}{4}[2s(s+t) + t^2 \log t - (t+2s)^2 \log(t+2s)] \\
&\qquad + \frac{s}{4}[2t(s+t) + s^2 \log s - (s+2t)^2 \log(s+2t)] \bigg\}.
\end{aligned} \tag{20}$$

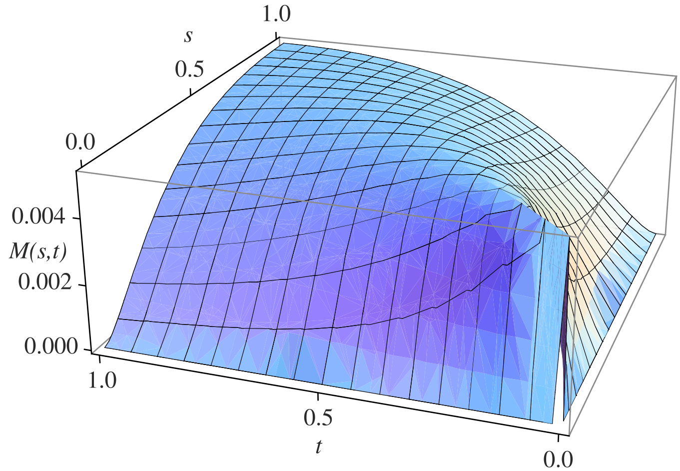

FIGURE 1: Plot of the past mutual information given in (20).

For any fixed $t \in (0, 1)$, it follows that $\tilde{M}_{X,Y}(s, t)$ is increasing for $s \in (0, t]$ and, thus, attains the maximum for $s = t$, with

$$\tilde{M}_{X,Y}(t, t) = \frac{2 + 40 \log 2 - 27 \log 3}{12} = 0.0053, \qquad t \in (0, 1).$$

The plot of $\tilde{M}_{X,Y}(s, t)$ is given in Figure 1. See [14, Example 1] for other results on the information content of the bivariate distribution considered in this example.

Let us now recall that the reversed hazard rate of a random lifetime X is given by $\tau_X(x) = -(\mathrm{d}/\mathrm{d}x) \log F_X(x) = f_X(x)/F_X(x)$ for all x such that $0 < F_X(x) < 1$, where $F_X(x) = \mathbb{P}(X \le x)$.

Remark 3. The argument of the logarithm in (16) can be viewed as a local dynamic measure of dependence between X and Y. Indeed, due to (8), (14), and (15), we have

$$\frac{\tilde{f}_{X,Y}(x, y; s, t)}{\tilde{f}_X(x; s, t)\tilde{f}_Y(y; s, t)} = \frac{\tau_{\tilde{X}_s}(x \mid Y = y)}{\tau_{\tilde{X}_s}(x \mid Y \le t)},$$

where $\tau_{\tilde{X}_s}(x \mid B)$ is the conditional reversed hazard rate of $\tilde{X}_s := [X \mid X \le s]$ given B.

4. Mutual information for residual lifetimes

The uncertainty about the remaining lifetime in reliability systems is often measured by means of the differential entropy of residual lifetimes; see [5], [12], and [13]. Recent contributions on the entropy of residual lifetimes are given in [1]. Other dynamic information measures involving conditional lifetimes have been proposed and studied in [3], [9], and [27]. For a random vector (X, Y) with nonnegative absolutely continuous components, Di Crescenzo *et al.* [10] studied the mutual information of the residual lifetimes $[X - t \mid X > t]$ and $[Y - t \mid Y > t]$ at the same age.

In this section, with reference to (5), we investigate the mutual information of the residual lifetimes at different ages, i.e.

$$[X - s \mid X > s, Y > t], \qquad [Y - t \mid X > s, Y > t], \quad s, t \ge 0. \tag{21}$$

For all $s, t \geq 0$ such that $\overline{F}(s,t) > 0$, the random variables (21) possess densities

$$f_X(x; s, t) = \frac{1}{\overline{F}(s,t)}\left[-\frac{\partial}{\partial u}\overline{F}(u,t)\right]_{u=x+s} = \frac{1}{\overline{F}(s,t)}\int_t^{+\infty} f(x+s, y)\,\mathrm{d}y, \qquad x \geq 0 \quad (22)$$

and

$$f_Y(y; s, t) = \frac{1}{\overline{F}(s,t)}\left[-\frac{\partial}{\partial v}\overline{F}(s,v)\right]_{v=y+t} = \frac{1}{\overline{F}(s,t)}\int_s^{+\infty} f(x, y+t)\,\mathrm{d}x, \qquad y \geq 0. \tag{23}$$

According to (1) we thus introduce the *bivariate dynamic residual mutual information*, for $s, t \geq 0$ such that $\overline{F}(s,t) > 0$,

$$M_{X,Y}(s,t) := \int_0^{+\infty} \mathrm{d}x \int_0^{+\infty} f_{X,Y}(x, y; s, t) \log \frac{f_{X,Y}(x, y; s, t)}{f_X(x; s, t) f_Y(y; s, t)}\,\mathrm{d}y, \tag{24}$$

the involved densities being defined in (9), (22), and (23). Since X and Y describe the random lifetimes of two systems, $M_{X,Y}(s,t)$ measures the dependence between their remaining lifetimes at different ages s and t. See the analogy between (24) and the mutual information of $[(X, Y) \mid X > s, Y > t]$ given in [14, Equation (1)]. We remark that other types of dynamic information measures for bivariate distributions have been studied by Sunoj and Linu [30].

Moreover, in agreement with (2), the mutual information $M_{X,Y}(s,t)$ satisfies the following identity (see [14, Equation (13)]):

$$M_{X,Y}(s,t) = H_X(s,t) + H_Y(s,t) - H_{X,Y}(s,t), \qquad s, t \geq 0, \tag{25}$$

where $H_{X,Y}(s,t)$ is defined in (11) and

$$\begin{aligned} H_X(s,t) &= -\int_0^{+\infty} f_X(x; s, t) \log f_X(x; s, t)\,\mathrm{d}x, \\ H_Y(s,t) &= -\int_0^{+\infty} f_Y(y; s, t) \log f_Y(y; s, t)\,\mathrm{d}y \end{aligned} \tag{26}$$

denote the entropies of the residual lifetimes (21) for $s, t \geq 0$. Various other results have been pinpointed in [14], such as the following property: if X and Y are exchangeable then $M_{X,Y}(s,t) = M_{X,Y}(t,s)$ for all $s, t \geq 0$.

We recall that the hazard rate of a random lifetime X is given by $h_X(x) = -(\mathrm{d}/\mathrm{d}x)\log[1 - F_X(x)] = f_X(x)/[1 - F_X(x)]$ for all x such that $0 < F_X(x) < 1$.

Remark 4. Similarly as in Remark 3, the argument of the logarithm in (24) can be viewed as a local dynamic measure of dependence between X and Y. Indeed, from (9), (22), and (23), we have

$$\frac{f_{X,Y}(x, y; s, t)}{f_X(x; s, t) f_Y(y; s, t)} = \frac{h_{X_s}(x \mid Y = y + t)}{h_{X_s}(x \mid Y > t)}, \tag{27}$$

where $h_{X_s}(x \mid B)$ is the conditional hazard rate of $X_s := [X - s \mid X > s]$ given B. Moreover, the right-hand side of (27) is a suitable extension of association measures that are often employed in reliability theory (see, for example [20] and the references therein).

The following result is analogous to Proposition 1.

Proposition 2. *For $s, t \geq 0$ such that $\overline{F}(s, t) > 0$, let*

$$a(x, y; s, t) := \frac{f(x, y)}{\int_t^{+\infty} f(x, v)\,\mathrm{d}v \int_s^{+\infty} f(u, y)\,\mathrm{d}u}, \qquad x \geq s,\ y \geq t. \tag{28}$$

If

$$a(x, y; s, t) \leq (\geq)\ a(s, t; s, t) \quad \text{for all } x \geq s,\ y \geq t, \tag{29}$$

then the following upper [lower] bound holds:

$$M_{X,Y}(s, t) \leq (\geq) \log a(s, t; s, t) + \log \overline{F}(s, t). \tag{30}$$

Proof. Due to (9), (22), and (23), from (24) we obtain the following alternative expression for $M_{X,Y}(s, t)$, $s, t \geq 0$:

$$M_{X,Y}(s, t) = \frac{1}{\overline{F}(s, t)} \int_s^{+\infty} \mathrm{d}x \int_t^{+\infty} f(x, y) \log a(x, y; s, t)\,\mathrm{d}y + \log \overline{F}(s, t). \tag{31}$$

The proof then immediately follows by use of (29) in the right-hand side of (31).

Example 2. Let (X, Y) be a random vector with joint PDF

$$f(x, y) = \frac{\theta}{\Gamma(0, 1/\theta)} \exp\biggl\{-\frac{1}{\theta}(1 + \theta x)(1 + \theta y)\biggr\}, \qquad x, y \geq 0,$$

with $\theta > 0$, and where $\Gamma(a, z) = \int_z^{+\infty} t^{a-1} \mathrm{e}^{-t}\,\mathrm{d}t$ is the incomplete gamma function. Since

$$\overline{F}(x, y) = \frac{\Gamma(0, (1/\theta)(1 + \theta x)(1 + \theta y))}{\Gamma(0, 1/\theta)}, \qquad x, y \geq 0,$$

from (9), we have, for $s, t \geq 0$,

$$f_{X,Y}(x, y; s, t) = \frac{\theta \exp\{-(1/\theta)[1 + \theta(x + s)][1 + \theta(y + t)]\}}{\Gamma(0, (1/\theta)(1 + \theta s)(1 + \theta t))}, \qquad x, y \geq 0.$$

Hence, recalling (28), after some calculations we obtain, for $x, y \geq 0$,

$$a(x, y; s, t) = \frac{1}{\theta}\Gamma\biggl(0, \frac{1}{\theta}\biggr)(1 + \theta x)(1 + \theta y) \exp\biggl\{\frac{1}{\theta}[1 + \theta(s + t) + \theta^2(tx + sy - xy)]\biggr\}.$$

This expression allows us to evaluate $M_{X,Y}(s, t)$ numerically, by use of (31). Other properties of dynamic measures concerning this case are given in [14, Section 4.4].

In the following proposition we show a relation between the bivariate dynamic residual and past mutual information of symmetric random vectors.

Proposition 3. *If the random vector (U, V) has bivariate density $f_{U,V}(x, y)$ such that, for a fixed $(x_0, y_0) \in \mathbb{R}_+^2$,*

$$f_{U,V}(x, y) = f(2x_0 - x, 2y_0 - y) \quad \text{for all } (x, y) \in \mathbb{R}_+^2, \tag{32}$$

then $\tilde{M}_{U,V}(s, t) = M_{X,Y}(2x_0 - s, 2y_0 - t)$ for all $s, t \geq 0$.

Proof. The proof follows from the definitions of $\tilde{M}_{U,V}$ and $M_{X,Y}$, since the distribution function of (U, V) satisfies $F_{U,V}(x, y) = \overline{F}(2x_0 - x, 2y_0 - y)$ for all $(x, y) \in \mathbb{R}_+^2$.

Example 3. Let (X, Y) be a random vector uniformly distributed over the domain $\mathcal{D} := \{(x, y)\colon x \geq 0, y \geq 0, \alpha x + \beta y \leq 1\}$ with $\alpha, \beta > 0$. Hence, the joint PDF and the joint survival function are given by

$$f(x, y) = 2\alpha\beta, \qquad \overline{F}(x, y) = (1 - \alpha x - \beta y)^2, \qquad (x, y) \in \mathcal{D},$$

so that, from (22), we have the density

$$f_X(x; s, t) = \frac{2\alpha[1 - \alpha(s + x) - \beta t]}{(1 - \alpha s - \beta t)^2}, \qquad 0 \leq x \leq \frac{1}{\alpha} - \frac{\beta}{\alpha} t - s, \quad (s, t) \in \mathcal{D}.$$

Due to (26), for $(s, t) \in \mathcal{D}$ the entropies of the residual lifetimes are

$$H_X(s, t) = \frac{1}{2} + \log \frac{1 - \alpha s - \beta t}{2\alpha}, \qquad H_Y(s, t) = \frac{1}{2} + \log \frac{1 - \alpha s - \beta t}{2\beta}. \tag{33}$$

From (9), we obtain, for $(s, t) \in \mathcal{D}$,

$$f_{X,Y}(x, y; s, t) = \frac{2\alpha\beta}{(1 - \alpha s - \beta t)^2}, \qquad (x + s, y + t) \in \mathcal{D}.$$

Hence, making use of (11) we obtain the entropy of $[(X - s, Y - t) \mid X > s, Y > t]$:

$$H_{X,Y}(s, t) = 2\log(1 - \alpha s - \beta t) - \log(2\alpha\beta), \qquad (s, t) \in \mathcal{D}. \tag{34}$$

In conclusion, recalling (25), (33), and (34) we establish that the dynamic residual mutual information of (X, Y) is constant:

$$M_{X,Y}(s, t) = 1 - \log 2 = 0.3069, \qquad (s, t) \in \mathcal{D}. \tag{35}$$

Note that in this case for $(s, t) \in \mathcal{D}$, we have

$$a(x, y; s, t) = \frac{1}{2(1 - \alpha x - \beta t)(1 - \alpha s - \beta y)} \geq a(s, t; s, t) = \frac{1}{2(1 - \alpha s - \beta t)^2};$$

however, now the bound given in (30) is not useful since the right-hand side of (30) is negative. Let (U, V) have density

$$f_{U,V}(x, y) = 2\alpha\beta \quad \text{for } (x, y) \in \tilde{D} := \left\{x \leq \frac{1}{\alpha}, y \leq \frac{1}{\beta}, \alpha x + \beta y \geq 1\right\},$$

and distribution function $F_{U,V}(x, y) = (\alpha x + \beta y - 1)^2$ for $(x, y) \in \tilde{D}$. Then, (U, V) is symmetric to (X, Y), in the sense that (32) holds for $(x_0, y_0) = (1/2\alpha, 1/2\beta)$. Hence, making use of Proposition 3 and recalling (35), we have $\tilde{M}_{U,V}(s, t) = 1 - \log 2 = 0.3069$ for $(s, t) \in \tilde{D}$.

It is worthwhile to remark that the residual mutual information is constant also in other cases. See [14, Section 3.1] for various comments on the memoryless property and related information notions. We recall that if the survival function of a nonnegative continuous vector variable (X, Y) satisfies $\overline{F}(x + t, y + t) = \overline{F}(x, y)\overline{F}(t, t)$ for all $x, y, t \geq 0$, then (X, Y) is said to possess the bivariate lack of memory (BLM) property; see, for example, [29]. It thus follows that if (X, Y) has the BLM property, then $M_{X,Y}(t, t)$ does not depend on t. For instance, the bivariate Block–Basu density and the bivariate Freund density have the BLM property. See also [28] for the weak multivariate lack of memory property within a stochastic model that will be discussed hereafter.

4.1. Dynamic mutual information for the time-transformed exponential model

We recall that a pair of random lifetimes (X, Y) is said to follow the time-transformed exponential (TTE) model if its joint survival function may be expressed in the following way:

$$\overline{F}(s,t) = \overline{W}[R_1(s) + R_2(t)] \quad \text{for all } s, t \geq 0, \tag{36}$$

where $\overline{W}\colon [0, +\infty) \to [0, 1]$ is a continuous, convex, and strictly decreasing survival function, such that $\overline{W}(0) = 1$ and $\lim_{r\to+\infty} \overline{W}(r) = 0$, and where $R_i\colon [0, +\infty) \to [0, +\infty)$ is a continuous and strictly increasing function, such that $R_i(0) = 0$ and $\lim_{t\to+\infty} R_i(t) = +\infty$ for $i = 1, 2$. Clearly, functions $\overline{W}$ and R_i, $i = 1, 2$, provide the time transform and the accumulated hazards, respectively. Note that the marginal survival functions are given by $\overline{F}_X(s) = \overline{W}[R_1(s)]$, $s \geq 0$, and $\overline{F}_Y(t) = \overline{W}[R_2(t)]$, $t \geq 0$. Moreover, if R_1 and R_2 are identical functions, then X and Y are exchangeable. The TTE model allows us to study the essential ageing properties of lifetimes (X, Y) by separating ageing property and dependence and, thus, it deserves wide interest in reliability theory and survival analysis. Various properties and applications of such a semiparametric model have been investigated recently in, for example, [4], [22], [24], [28], and [31].

Hereafter, we investigate the bivariate dynamic residual mutual information within the TTE model.

Proposition 4. *If the survival function of (X, Y) satisfies the TTE model as specified in (36), then, for all $s, t \geq 0$,*

$$\begin{aligned} M_{X,Y}(s,t) = {} & \frac{1}{\overline{W}[R_1(s) + R_2(t)]} \\ & \times \int_{R_1(s)}^{+\infty} \mathrm{d}u \int_{R_2(t)}^{+\infty} \overline{W}''[u+v] \log \frac{\overline{W}''[u+v]\overline{W}[R_1(s)+R_2(t)]}{\overline{W}'[u+R_2(t)]\overline{W}'[R_1(s)+v]}\, \mathrm{d}v. \end{aligned} \tag{37}$$

Proof. Let $s, t \geq 0$. From (36), it follows that

$$f(s,t) = \overline{W}''[R_1(s) + R_2(t)]R_1'(s)R_2'(t).$$

Hence, from (9), (22), and (23), we have

$$f_X(x; s, t) = -\frac{\overline{W}'[R_1(x+s) + R_2(t)]R_1'(x+s)}{\overline{W}[R_1(s) + R_2(t)]}, \qquad x \geq 0, \tag{38}$$

$$f_Y(y; s, t) = -\frac{\overline{W}'[R_1(s) + R_2(y+t)]R_2'(y+t)}{\overline{W}[R_1(s) + R_2(t)]}, \qquad y \geq 0, \tag{39}$$

and

$$f_{X,Y}(x, y; s, t) = \frac{\overline{W}''[R_1(x+s) + R_2(y+t)]R_1'(x+s)R_2'(y+t)}{\overline{W}[R_1(s) + R_2(t)]}, \qquad x, y \geq 0. \tag{40}$$

Finally, (37) follows by substituting the above densities in the right-hand side of (24), and by setting $u = R_1(x+s)$ and $v = R_2(y+t)$.

The following result can be obtained by means of straightforward calculations.

Corollary 1. *Let* (X, Y) *satisfy the assumptions of Proposition 4. If*

$$\overline{W}(x) = (1+x)^{-r}, \quad x \geq 0, \qquad R_1(s) = \alpha s, \quad s \geq 0, \qquad R_2(t) = \beta t, \quad t \geq 0,$$

with $r, \alpha, \beta > 0$*, then*

$$M_{X,Y}(s,t) = -\frac{1}{r+1} + \log\frac{r+1}{r}, \qquad s, t \geq 0.$$

From Corollary 1 we show that if (X, Y) has bivariate Lomax (Pareto type II) joint survival function then $M_{X,Y}(s,t)$ is constant (see also [14, Section 5.2]). Note that in this case $a(x, y; s, t)$ is not monotone; so that the bound (29) is not useful.

4.2. Dynamic mutual information for the truncated TTE model

We now consider a TTE model for *truncated* random lifetimes (X, Y). Specifically, we assume such that the nonnegative random variables X and Y are upper bounded through a suitable function. Unlike the previous section, we now assume that $\overline{W}(r)$ is a continuous, convex, and strictly decreasing one-dimensional survival function for all $r \in [0, \omega]$, where ω is a fixed positive real number, such that $\overline{W}(0) = 1$ and $\overline{W}(\omega) = 0$. Moreover, $R_1(\cdot)$ and $R_2(\cdot)$ are continuous and strictly increasing functions such that $R_1(0) = R_2(0) = 0$, and the set

$$D_\omega := \{(s,t) \in \mathbb{R}^2 \colon s \geq 0, t \geq 0, R_1(s) + R_2(t) \leq \omega\}$$

is not empty. Hence, there exists a continuous and strictly decreasing function $t = \ell_\omega(s)$, defined for $0 \leq s \leq R_1^{-1}(\omega)$, and such that $R_1(s) + R_2(\ell_\omega(s)) = \omega$ for all $s \in [0, R_1^{-1}(\omega)]$, with $\ell_\omega(0) = R_2^{-1}(\omega)$ and $\ell_\omega(R_1^{-1}(\omega)) = 0$. These assumptions thus lead to the following *truncated TTE model* for the joint survival function of (X, Y):

$$\overline{F}(s,t) = \overline{W}[R_1(s) + R_2(t)] \quad \text{for all } (s,t) \in D_\omega. \tag{41}$$

Similarly to Proposition 4, we thus have the following result for the dynamic residual mutual information within the above model.

Proposition 5. *If the joint survival function of* (X, Y) *satisfies the TTE model as specified in (41), with* $\overline{W}'(\omega) = 0$*, then, for all* $s, t \in D_\omega$*,*

$$M_{X,Y}(s,t) = \frac{1}{\overline{W}[R_1(s) + R_2(t)]} \times \int_{R_1(s)}^{\omega - R_2(t)} \mathrm{d}u \int_{R_2(t)}^{\omega - u} \overline{W}''[u+v] \log\frac{\overline{W}''[u+v]\overline{W}[R_1(s)+R_2(t)]}{\overline{W}'[u+R_2(t)]\overline{W}'[R_1(s)+v]}\,\mathrm{d}v.$$

Proof. Under the given assumptions the densities in (9), (22), and (23) can still be expressed respectively as in (38) for $0 \leq x \leq R_1^{-1}(\omega - R_2(t)) - s$, as in (39) for $0 \leq y \leq R_2^{-1}(\omega - R_1(s)) - t$, and as in (40) for all nonnegative x, y such that $R_1(x+s) + R_2(y+t) \leq \omega$. Note that the assumption $\overline{W}'(\omega) = 0$ is essential to ascertain that the integral of $f_{X,Y}(x, y; s, t)$ is unity. The proof thus proceeds similarly as that of Proposition 4.

The following result can be obtained via direct calculations.

Corollary 2. *Let (X, Y) satisfy the assumptions of Proposition 5. If*

$$\overline{W}(x) = \left(\frac{x}{\omega} - 1\right)^2, \qquad 0 \le x \le \omega,$$

then $M_{X,Y}(s,t) = 1 - \log 2 = 0.3069$, $(s,t) \in D_\omega$.

5. Dynamic mutual information for ordered data

The approach developed in the previous sections can also be adopted to study the mutual information in the presence of conditioning expressed as in (6) and (7). Here we restrict ourselves to consider models based on ordered data, with an application to order statistics. For $n \ge 2$, consider a system with n components, having independent and identically distributed random lifetimes. Assume that the failures of the components are observed upon a test. Suppose that the ith failure occurs before time s and $n - j + 1$ ($j > i$) components are still alive at time t, with $0 < s < t$. For $1 \le i < j \le n$, we can define the following random variables:

$$T_{i,j:n}(s,t) = [(X_{i:n}, X_{j:n}) \mid X_{i:n} \le s, X_{j:n} > t], \qquad 0 < s < t, \tag{42}$$

where $X_{r:n}$ denotes the rth order statistic. We recall that Ebrahimi *et al.* [17] defined and studied mutual information between consecutive ordinary order statistics.

Let us now define dynamic mutual information measures for order statistics. As a case study, we consider (42) for $i = 1$ and $j = n$, i.e. we assume that the first failure occurs before time s, and the last failure occurs after time t. Then the joint PDF of $T_{1,n:n}(s,t)$ and the marginal PDFs of $[X_{1:n} \mid X_{1:n} \le s, X_{n:n} > t]$ and $[X_{n:n} \mid X_{1:n} \le s, X_{n:n} > t]$ are needed. Let $f(x)$ and $F(x)$ denote respectively the common PDF and the distribution function of the components' lifetimes. Since (see, for example, [6])

$$f_{1,n:n}(x,y) = n(n-1)[F(y) - F(x)]^{n-2} f(x) f(y), \qquad 0 < x < y < +\infty,$$

for $0 < s < t$, we have

$$\begin{aligned}
\mathbb{P}(X_{1:n} \le s, X_{n:n} > t) &= \int_0^s \mathrm{d}x \int_t^{+\infty} f_{1,n:n}(x,y)\,\mathrm{d}y \\
&= \int_0^s n f(x)\{[1 - F(x)]^{n-1} - [F(t) - F(x)]^{n-1}\}\,\mathrm{d}x \\
&= 1 - [F(t)]^n + [F(t) - F(s)]^n - [1 - F(s)]^n.
\end{aligned} \tag{43}$$

Let

$$f^*_{1:n}(x; s, t) = \frac{(\partial/\partial x)\mathbb{P}(X_{1:n} \le x < t < X_{n:n})}{\mathbb{P}(X_{1:n} \le s, X_{n:n} > t)}$$

denote the PDF of $[X_{1:n} \mid X_{1:n} \le s, X_{n:n} > t]$. Hence, using (43), we obtain

$$f^*_{1:n}(x; s, t) = \frac{n\{[1 - F(x)]^{n-1} - [F(t) - F(x)]^{n-1}\} f(x)}{1 - [F(t)]^n + [F(t) - F(s)]^n - [1 - F(s)]^n}, \qquad 0 < x < s < t. \tag{44}$$

Similarly, denoting the PDF of $[X_{n:n} \mid X_{1:n} \le s, X_{n:n} > t]$ by $f^*_{n:n}(y; s, t)$, we have

$$f^*_{n:n}(y; s, t) = \frac{n\{[F(y)]^{n-1} - [F(y) - F(s)]^{n-1}\} f(y)}{1 - [F(t)]^n + [F(t) - F(s)]^n - [1 - F(s)]^n}, \qquad 0 < s < t < y. \tag{45}$$

Also, let $f^*_{1,n:n}(x, y; s, t)$ be the PDF of $T_{1,n:n}(s, t)$. Then, it is given by

$$f^*_{1,n:n}(x, y; s, t) = \frac{n(n-1)[F(y) - F(x)]^{n-2} f(x) f(y)}{1 - [F(t)]^n + [F(t) - F(s)]^n - [1 - F(s)]^n}, \qquad 0 < x < s < t < y. \tag{46}$$

By virtue of (44), (45), and (46), the dynamic mutual information of $T_{1,n:n}(s, t)$ can thus be defined as

$$M^*_{1,n:n}(s, t) = \int_0^s \mathrm{d}x \int_t^{+\infty} f^*_{1,n:n}(x, y; s, t) \log \frac{f^*_{1,n:n}(x, y; s, t)}{f^*_{1:n}(x; s, t) f^*_{n:n}(y; s, t)} \,\mathrm{d}y, \qquad 0 < s < t. \tag{47}$$

Obviously, (47) depends on $s, t, n, F(s)$, and $F(t)$. Also, $M^*_{1,2:2}(s, t) = 0$ for all $0 < s < t$. However, in agreement with [17, Theorem 3.3(a)], in the following we show that $M^*_{1,n:n}(s, t)$ is distribution-free under suitable assumptions.

According to the previous comments, s and t can be seen as inspection times for the n-component system. The knowledge of $[X_{1:n} \le s, X_{n:n} > t]$ thus means that, upon inspection, at least one failed component has been detected at time s, and at least one component is functioning at time t. We can fix s and t as quantiles of F, say as the pth and qth quantiles, respectively, i.e.

$$s = \xi_p = F^{-1}(p), \qquad t = \xi_q = F^{-1}(q), \quad 0 < p < q < 1, \tag{48}$$

where F^{-1} is the generalized inverse of F. Denote by $H_n(p, q)$ the joint probability (43) when s and t are chosen as in (48), i.e.

$$H_n(p, q) = 1 - q^n + (q - p)^n - (1 - p)^n, \qquad 0 < p < q < 1. \tag{49}$$

Moreover, in order to show that $M^*_{1,n:n}(s, t)$ is distribution-free, for $p, q \in (0, 1)$, we set

$$K_n(p, q) := \int_0^p [(1 - u)^{n-1} - (q - u)^{n-1}] \log((1 - u)^{n-1} - (q - u)^{n-1}) \,\mathrm{d}u. \tag{50}$$

Proposition 6. *Let $n \ge 2$. If s and t are chosen as in (48), with $0 < p < q < 1$, then the dynamic mutual information of $T_{1,n:n}(s, t)$ is given by*

$$\begin{aligned} M^*_{1,n:n}(\xi_p, \xi_q) &= \log\biggl[\frac{n-1}{n} H_n(p, q)\biggr] - \frac{(n-2)(2n-1)}{n(n-1)} \\ &\quad - \frac{n}{H_n(p, q)} \biggl\{K_n(p, q) + K_n(1 - q, 1 - p) + \frac{n-2}{n} \\ &\qquad \times [(1 - p)^n \log(1 - p) + q^n \log(q) - (q - p)^n \log(q - p)]\biggr\}, \end{aligned} \tag{51}$$

where H_n and K_n are given in (49) and (50), respectively.

Proof. Equation (51) follows from (47), and by using densities (44), (45), and (46).

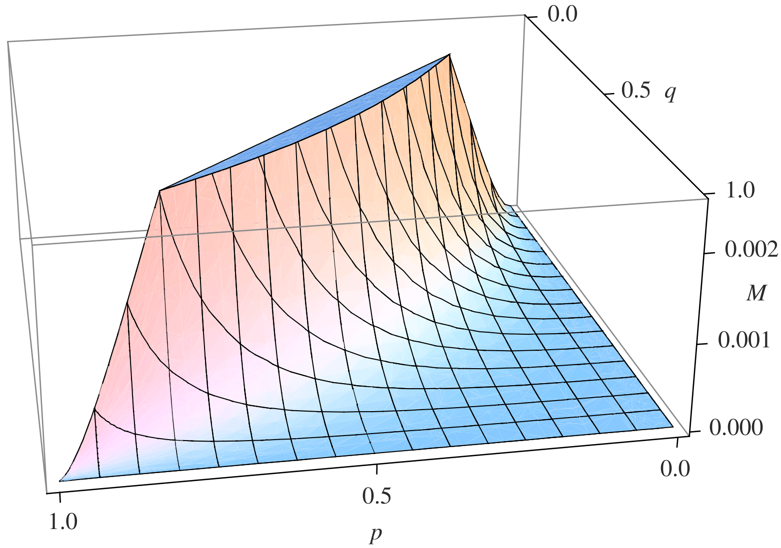

FIGURE 2: Plot of the mutual information given in (51) for $n = 3$.

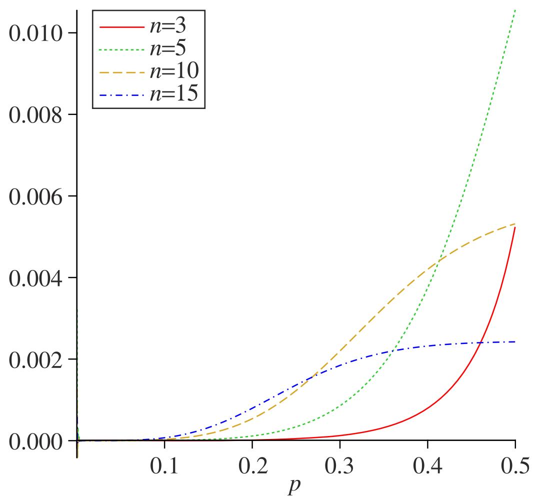

FIGURE 3: Plot of the mutual information given in (51) for $n = 3, 5, 10, 15$, and $q = 1 - p$.

If $n = 2$ the analysis of $T_{1,2:2}(s,t)$ is trivial. Indeed, from Proposition 6 it is not hard to see that $M^*_{1,2:2}(\xi_p, \xi_q) = 0$ for all $0 < p < q < 1$. Also, a closed-form expression for $M^*_{1,n:n}(\xi_p, \xi_q)$ can be obtained from (51) when $n = 3$; however, we omit it for being lengthy and tedious. We limit ourselves to show in Figure 2 the plot of $M^*_{1,3:3}(\xi_p, \xi_q)$ for $0 < p < q < 1$. Furthermore, in Figure 3 we show the plot of $M^*_{1,n:n}(\xi_p, \xi_{1-p})$ for some selected values of n and $0 < p < \frac{1}{2}$, in the special case $q = 1 - p$. From Figure 3, we confirm that the mutual information $M^*_{1,n:n}(\xi_p, \xi_{1-p})$ is increasing in p, as expected.

6. A copula-based approach

The copula function is an useful tool in studying the dependency in multivariate distributions; see, for example, [11]. Sklar's theorem asserts that, given a copula $C\colon [0,1]^2 \to [0,1]$, the joint cumulative distribution function of (X, Y) can be written in terms of the marginals as

$$F(x, y) = C(F_X(x), F_Y(y)), \qquad x, y \in \mathbb{R}, \tag{52}$$

the copula being unique if the marginals are continuous. The corresponding copula density is given by

$$c(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v) = \frac{\partial^2}{\partial u \partial v} F(F_X^{-1}(u), F_Y^{-1}(v)), \qquad u, v \in (0, 1),$$

where F_X^{-1} and F_Y^{-1} denote the generalized inverse of the marginals. Thus, the joint PDF of (X, Y) can be expressed as

$$f(x, y) = f_X(x) f_Y(y) c(F_X(x), F_Y(y)), \qquad x, y \in \mathbb{R}, \tag{53}$$

so that the mutual information can be written in terms of the copula density as (see, for example, [7])

$$M_{X,Y} = \int_0^1 \mathrm{d}u \int_0^1 c(u, v) \log c(u, v)\, \mathrm{d}v.$$

This confirms that the mutual information does not depend on the marginal distributions, and also that the copula entails all essential information on the dependence between X and Y.

Let us now represent the dynamic past mutual information in terms of the copula function. We first make use of (53) in (14) and perform the substitution $v = F_Y(y)$ in the integral. Moreover, similarly to (48), we set

$$s = \xi_p = F_X^{-1}(p), \qquad t = \xi_q = F_Y^{-1}(q), \quad p, q \in (0, 1), \tag{54}$$

so that the density of the marginal past lifetime $[X \mid X \le F_X^{-1}(p), Y \le F_Y^{-1}(q)]$ can be expressed as

$$\tilde{f}_X(x; \xi_p, \xi_q) = \frac{f_X(x)}{C(p, q)} \int_0^q c(F_X(x), v)\, \mathrm{d}v, \qquad 0 \le x \le F_X^{-1}(p), \tag{55}$$

the right-hand side of (55) being a weighted density of X. Similarly, from (15), it follows that the density of $[Y \mid X \le F_X^{-1}(p), Y \le F_Y^{-1}(q)]$ is given by

$$\tilde{f}_Y(y; \xi_p, \xi_q) = \frac{f_Y(y)}{C(p, q)} \int_0^p c(u, F_Y(y))\, \mathrm{d}u, \qquad 0 \le y \le F_Y^{-1}(q). \tag{56}$$

Finally, for the bivariate past lifetimes

$$[(X, Y) \mid X \le F_X^{-1}(p), Y \le F_Y^{-1}(q)], \qquad p, q \in (0, 1), \tag{57}$$

the joint PDF (8) becomes

$$\tilde{f}_{X,Y}(x, y; \xi_p, \xi_q) = f_X(x) f_Y(y) \frac{c(F_X(x), F_Y(y))}{C(p, q)} \tag{58}$$

for $0 \le x \le F_X^{-1}(p)$ and $0 \le y \le F_Y^{-1}(q)$.

Proposition 7. *For all $p, q \in (0, 1)$, the mutual information of the bivariate past lifetimes (57) is given by*

$$\tilde{M}_{X,Y}(\xi_p, \xi_q) = \log[C(p,q)] + \frac{1}{C(p,q)} \int_0^p \mathrm{d}u \int_0^q c(u,v) \log \frac{c(u,v)}{\int_0^q c(u,w)\,\mathrm{d}w \int_0^p c(z,v)\,\mathrm{d}z}\,\mathrm{d}v. \tag{59}$$

Proof. Due to (55), (56), and (58), it follows that the mutual information of (57) is

$$\tilde{M}_{X,Y}(\xi_p, \xi_q) = \int_0^{F_X^{-1}(p)} \times \mathrm{d}x \int_0^{F_Y^{-1}(q)} \frac{c(F_X(x), F_Y(y)) f_X(x) f_Y(y)}{C(p,q)} \times \log \frac{c(F_X(x), F_Y(y)) C(p,q)}{\int_0^q c(F_X(x), v)\,\mathrm{d}v \int_0^p c(u, F_Y(y))\,\mathrm{d}u}\,\mathrm{d}y, \qquad p, q \in (0,1).$$

Finally, setting $u = F_X(x)$ and $v = F_Y(y)$, we obtain (59).

Example 4. Let (X, Y) have copula

$$C(u,v) = \frac{u\,v}{u + v - uv}, \qquad u, v \in (0,1),$$

i.e. a special case of a Clayton copula. From Proposition 7, it follows that the mutual information of (57) is

$$\tilde{M}_{X,Y}(\xi_p, \xi_q) = -\tfrac{1}{2} + \log 2 = 0.1931, \qquad p, q \in (0,1).$$

Let us now consider the joint survival function $\overline{F}(x, y)$ and the corresponding marginal survival functions $\overline{F}_X(x) = \mathbb{P}(X > x)$ and $\overline{F}_Y(y) = \mathbb{P}(Y > y)$. Similarly as in (52), these functions are related by

$$\overline{F}(x,y) = \tilde{C}(\overline{F}_X(x), \overline{F}_Y(y)), \qquad x, y \in \mathbb{R}, \tag{60}$$

where $\tilde{C}(u,v) = 1 - u - v - C(u,v)$, $u, v \in (0,1)$, is the survival copula function. The survival copula density, given by

$$\tilde{c}(u,v) = \frac{\partial^2}{\partial u \partial v} \tilde{C}(u,v) = \frac{\partial^2}{\partial u \partial v} \overline{F}(\overline{F}_X^{-1}(u), \overline{F}_Y^{-1}(v)), \qquad u, v \in (0,1),$$

allows us to express the joint density of (X, Y) as

$$f(x,y) = f_X(x) f_Y(y) \tilde{c}(\overline{F}_X(x), \overline{F}_Y(y)), \qquad x, y \in \mathbb{R}. \tag{61}$$

We recall that the copula density and the survival copula density are related by the following identity: $c(u,v) = \tilde{c}(1-u, 1-v)$, $u, v \in (0,1)$.

In order to consider the residual mutual information we make use of (60) and (61) in (22), and perform the substitution $v = F_Y(y)$ in the integral. Moreover, by setting s and t as in (54), the density of the marginal residual lifetime $[X - F_X^{-1}(p) \mid X > F_X^{-1}(p), Y > F_Y^{-1}(q)]$ can be expressed as

$$f_X(x; \xi_p, \xi_q) = \frac{f_X(x + F_X^{-1}(p))}{\tilde{C}(1-p, 1-q)} \int_q^1 \tilde{c}(\overline{F}_X(x + F_X^{-1}(p)), 1 - v)\,\mathrm{d}v \quad \text{for } x \geq 0. \tag{62}$$

Similarly, from (23), it follows that the density of $[Y - F_Y^{-1}(q) \mid X > F_X^{-1}(p), Y > F_Y^{-1}(q)]$ is given by

$$f_Y(y; \xi_p, \xi_q) = \frac{f_Y(y + F_Y^{-1}(q))}{\tilde{C}(1-p, 1-q)} \int_p^1 \tilde{c}(1-u, \overline{F}_Y(y + F_Y^{-1}(q)))\,\mathrm{d}u \quad \text{for } y \geq 0. \tag{63}$$

Furthermore, the density of the joint residual lifetimes

$$[(X - F_X^{-1}(p), Y - F_Y^{-1}(q)) \mid X > F_X^{-1}(p), Y > F_Y^{-1}(q)], \qquad p, q \in (0,1), \tag{64}$$

is

$$f_{X,Y}(x, y; \xi_p, \xi_q) = f_X(x + F_X^{-1}(p)) f_Y(y + F_Y^{-1}(q)) \times \frac{\tilde{c}(\overline{F}_X(x + F_X^{-1}(p)), \overline{F}_Y(y + F_Y^{-1}(q)))}{\tilde{C}(1-p, 1-q)} \quad \text{for } x \geq 0,\ y \geq 0. \tag{65}$$

In conclusion, we obtain the dynamic mutual information for residual lifetimes in terms of the survival copula.

Proposition 8. *The mutual information of the bivariate residual lifetimes (64) for all* $p, q \in (0,1)$ *is given by*

$$M_{X,Y}(\xi_p, \xi_q) = \log[\tilde{C}(1-p, 1-q)] \tag{66}$$

$$+ \frac{1}{\tilde{C}(1-p, 1-q)} \times \int_0^{1-p} \mathrm{d}z \int_0^{1-q} \tilde{c}(z, w) \log \frac{\tilde{c}(z, w)}{\int_0^{1-q} \tilde{c}(z, v)\,\mathrm{d}v \int_0^{1-p} \tilde{c}(u, w)\,\mathrm{d}u} \mathrm{d}w. \tag{67}$$

Proof. Making use of (62), (63), and (65), for $p, q \in (0,1)$, we can write

$$M_{X,Y}(\xi_p, \xi_q) = \int_0^{+\infty} \mathrm{d}x \times \int_0^{+\infty} \frac{\tilde{c}(\overline{F}_X(x + F_X^{-1}(p)), \overline{F}_Y(y + F_Y^{-1}(q)))}{\tilde{C}(1-p, 1-q)} \times f_X(x + F_X^{-1}(p)) f_Y(y + F_Y^{-1}(q)) \times \log \frac{\tilde{C}(1-p, 1-q) \tilde{c}(\overline{F}_X(x + F_X^{-1}(p)), \overline{F}_Y(y + F_Y^{-1}(q)))}{\int_q^1 \tilde{c}(\overline{F}_X(x + F_X^{-1}(p)), 1 - v)\,\mathrm{d}v \int_p^1 \tilde{c}(1-u, \overline{F}_Y(y + F_Y^{-1}(q)))\,\mathrm{d}u} \mathrm{d}y.$$

Hence, setting $z = \overline{F}_X(x + F_X^{-1}(p))$ and $w = \overline{F}_Y(y + F_Y^{-1}(q))$ we obtain (67).

Acknowledgements

This work is partially supported by INdAM-GNCS, by FARO (Università di Napoli Federico II), and by Regione Campania (Legge 5).

References

[1] Abbasnejad, M., Arghami, N. R., Morgenthaler, S. and Mohtashami Borzadaran, G. R. (2010). On the dynamic survival entropy. *Statist. Prob. Lett.* **80,** 1962–1971.
[2] Arellano-Valle, R. B., Contreras-Reyes, J. E. and Genton, M. G. (2013). Shannon entropy and mutual information for multivariate skew-elliptical distributions. *Scand. J. Statist.* **40,** 42–62.
[3] Asadi, M. and Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. *J. Statist. Planning Infer.* **137,** 1931–1941.
[4] Bassan, B. and Spizzichino, F. (2005). Relations among univariate aging, bivariate aging and dependence for exchangeable lifetimes. *J. Multivariate Anal.* **93,** 313–339.
[5] Belzunce, F., Navarro, J., Ruiz, J. M. and del Aguila, Y. (2004). Some results on residual entropy function. *Metrika* **59,** 147–161.
[6] David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, 3rd edn. John Wiley, Hoboken, NJ.
[7] Davy, M. and Doucet A. (2003). Copulas: a new insight into positive time-frequency distributions. *IEEE Signal Process. Lett.* **10,** 215–218.
[8] Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *J. Appl. Prob.* **39,** 434–440.
[9] Di Crescenzo, A. and Longobardi, M. (2009). On cumulative entropies. *J. Statist. Planning. Infer.* **139,** 4072–4087.
[10] Di Crescenzo, A., Longobardi, M. and Nastro, A. (2004). On the mutual information between residual lifetime distributions. In: *Cybernetics and Systems 2004*, Vol. 1, Austrian Society for Cybernetic Studies, Vienna, pp. 142–145.
[11] Durante, F. and Sempi, C. (2015). *Principles of Copula Theory*. Chapman and Hall/CRC, Boca Raton, FL.
[12] Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. *Sankhyā A* **58,** 48–56.
[13] Ebrahimi, N. and Pellerey, F. (1995). New partial ordering of survival functions based on the notion of uncertainty. *J. Appl. Prob.* **32,** 202–211.
[14] Ebrahimi, N., Kirmani, S. N. U. A. and Soofi, E. S. (2007). Multivariate dynamic information. *J. Multivariate Anal.* **98,** 328–349.
[15] Ebrahimi, N., Soofi, E. S. and Soyer, R. (2008). Multivariate maximum entropy identification, transformation, and dependence. *J. Multivariate Anal.* **99,** 1217–1231.
[16] Ebrahimi, N., Soofi, E. S. and Soyer, R. (2010). Information measures in perspective. *Internat. Statist. Rev.* **78,** 383–412.
[17] Ebrahimi, N., Soofi, E. S. and Zahedi, H. (2004). Information properties of order statistics and spacings. *IEEE Trans. Inf. Theory* **50,** 177–183.
[18] Ebrahimi, N., Hamedani, G. G., Soofi, E. S. and Volkmer, H. (2010). A class of models for uncorrelated random variables. *J. Multivariate Anal.* **101,** 1859–1871.
[19] Giraudo, M. T., Sacerdote, L. and Sirovich, R. (2013). Non-parametric estimation of mutual information through the entropy of the linkage. *Entropy* **15,** 5154–5177.
[20] Gupta, R. C. (2008). Reliability studies of bivariate distributions with exponential conditionals. *Math. Comput. Modelling* **47,** 1009–1018.
[21] Kundu, C., Nanda, A. K. and Maiti, S. S. (2010). Some distributional results through past entropy. *J. Statist. Planning Infer.* **140,** 1280–1291.
[22] Li, X. and Lin, J. (2011). Stochastic orders in time transformed exponential models with applications. *Insurance Math. Econom.* **49,** 47–52.
[23] Misagh, F. and Yari, G. H. (2011). On weighted interval entropy. *Statist. Prob. Lett.* **81,** 188–194.
[24] Mulero, J., Pellerey, F. and Rodríguez-Griñolo, R. (2010). Stochastic comparisons for time transformed exponential models. *Insurance Math. Econom.* **46,** 328–333.
[25] Nanda, A. K. and Paul, P. (2006). Some properties of past entropy and their applications. *Metrika* **64,** 47–61.
[26] Nanda, A. K. and Paul, P. (2006). Some results on generalized past entropy. *J. Statist. Planning Infer.* **136,** 3659–3674.
[27] Navarro, J., del Aguila, Y. and Asadi, M. (2010). Some new results on the cumulative residual entropy. *J. Statist. Planning Infer.* **140,** 310–322.
[28] Pellerey, F. (2008). On univariate and bivariate aging for dependent lifetimes with Archimedean survival copulas. *Kybernetika* **44,** 795–806.
[29] Roy, D. (2002). On bivariate lack of memory property and a new definition. *Ann. Inst. Statist. Math.* **54,** 404–410.
[30] Sunoj, S. M. and Linu, M. N. (2012). Dynamic cumulative residual Renyi's entropy. *Statistics* **46,** 41–56.
[31] You, Y., Li, X. and Balakrishnan, N. (2014). On extremes of bivariate residual lifetimes from generalized Marshall–Olkin and time transformed exponential models. *Metrika* **77,** 1041–1056.