Varentropy of Past Lifetimes

Varentropy of past lifetimes

Francesco Buono
Dipartimento di Matematica e Applicazioni "Renato Caccioppoli"
Università degli Studi di Napoli Federico II
I-80126, Napoli, Italy
francesco.buono3@unina.it

Maria Longobardi
Dipartimento di Biologia
Università degli Studi di Napoli Federico II
malongob@unina.it
I-80126, Napoli, Italy

Franco Pellerey
Dipartimento di Scienze Matematiche
Politecnico di Torino
franco.pellerey@polito.it
I-10129, Torino, Italy

May 2, 2022

Abstract

In a variety of applicative fields the level of information in random quantities is commonly measured by means of the Shannon Entropy. In particular, in reliability theory and survival analysis, time-dependent generalizations of this measure of uncertainty have been considered to dynamically describe changes in the degree of information over time. The Residual Entropy and the Residual Varentropy, for example, have been considered in the specialized literature to measure the information and its variability in residual lifetimes. In a similar way, one can consider dynamic measures of information for past lifetimes, i.e. for random lifetimes of items when one assumes that their failures occur before a fixed inspection time. This paper provides a study of the Past Varentropy, defined as the dynamic measure of variability of information for past lifetimes. From this study emerges the interest on a particular family of lifetimes distributions, whose members satisfy the property to be the only ones having constant Past Varentropy.

Keywords: Past lifetimes; Reversed hazard rate function; Entropy, Varentropy.

AMS 2020 Subject Classification: Primary 62N05; Secondary: 60E05, 94A17.

1 Introduction

Let X be an absolutely continuous non-negative random variable representing the lifetime of an item, or of an individual. If f_X denotes its density, one can define the well-known *Shannon information measure* (or *Entropy*) as

$$H_e(X) = \mathbb{E}[IC(X)] = \mathbb{E}[-\log f_X(X)] = -\int_0^{+\infty} f_X(x)\log f_X(x)\mathrm{d}x, \tag{1.1}$$

where

$$IC(X) = -\log f_X(X)$$

denotes the *Information Content* of X, which can be understood as the self-information or "surprisal" associated with the possible outcomes of X (see Shannon, 1948). Actually, $H_e(X)$ measures the expected uncertainty contained in f_X related to the predictability of an outcome of X. We refer to Sun Han and Kobayashi (2002) or to Gray (2011) for two recent comprehensive monographs on information measures and their applications in a variety of fields (see also Nanda and Chowdhury, 2019). Note that, in statistics, one may think of the information content as the log likelihood function, that is of great interest in parameters estimation.

As already mentioned, it follows from (1.1) that the Shannon entropy represents the expectation of the (random) information content $IC(X)$. But for different purposes (see, e.g., Bobkov and Madiman, 2011), one can also consider its variance, in order to evaluate the concentration of the information content around the entropy $H_e(X)$. Thus, one can also be interested in the *Varentropy* of X (sometimes called *Minimal Coding Variance* of X, whenever X is discrete), defined as

$$\begin{aligned} V_e(X) &= \mathrm{Var}[IC(X)] = \mathrm{Var}[-\log f_X(X)] \\ &= \mathrm{Var}[\log f_X(X)] = \mathbb{E}\left[(\log f_X(X))^2\right] - [H_e(X)]^2 \\ &= \int_0^{+\infty} f_X(x)[\log f_X(x)]^2\mathrm{d}x - \left[\int_0^{+\infty} f_X(x)\log f_X(x)\mathrm{d}x\right]^2. \end{aligned} \tag{1.2}$$

In recent literature several papers deal with the varentropy and its properties and applications, such as Madiman and Wang (2014) and Arikan (2016) and references therein. For example, approximations of minimal rates for data compression in terms of entropy and varentropy are given in Kontoyiannis and Verdú (2014). Also, knowing the entropy and the varentropy one can define reference intervals for the information content $IC(X)$ of the form

$$\mathbb{E}[IC(X)] \pm k\sqrt{\mathrm{Var}[IC(X)]} = H_e(X) \pm k\sqrt{V_e(X)} \tag{1.3}$$

for suitable choices of k. In the statistics field, this interval can be used to evaluate the uncertainty about likelihood estimates.

It must be pointed out that the Shannon entropy, as well as the varentropy, provides a measure of information for the random lifetime of an item which is new, whenever X represents its lifetime. For such reason, different time dependent versions of this measure have been proposed in the context of reliability and survival analysis, where the behavior of residual lifetimes along time, or past lifetimes, are the main objects of the studies. The most well-known version of such dynamic ones is the Residual Entropy defined and studied in Muliere et al. (1993) and Ebrahimi (1996), whose definition is recalled here. Given the absolutely continuous random lifetime X, having support $\mathcal{S} \subseteq \mathbb{R}^+$, survival function $\overline{F}_X$ and density f_X, let $X_t = (X - t | X > t)$ denote the corresponding *Residual Lifetime* at time $t \in \mathcal{S}$, i.e. the variable whose density is given by

$$f_{X_t}(x) = \frac{f_X(x+t)}{\overline{F}_X(t)}, \qquad x : x + t \in \mathcal{S}.$$

The *Residual Entropy* of X is the function of time $t \in \mathcal{S}$ defined as

$$H_e(X_t) = \mathbb{E}[IC(X_t)] = \mathbb{E}[-\log f_{X_t}(X_t)] = -\int_t^{+\infty} \frac{f_X(x)}{\overline{F}_X(t)} \log \frac{f_X(x)}{\overline{F}_X(t)} \, \mathrm{d}x.$$

It must be pointed out that the entropy (and the varentropy) of a random lifetime Y is actually a functional of its density f_Y, depending only on the distribution of Y and not on the value assumed by Y. But the residual entropy (such as the residual varentropy defined below) can be considered as a function of the inspection time t: to every time t it corresponds a real value $H_e(X_t)$ (defined as a functional of the density f_{X_t}), and this is the reason why we treat them here as if they were functions of t.

In a similar manner it can be defined a dynamic version of the varentropy, useful to evaluate the concentration of the information content in residual lifetimes when the time increases. This is the *Residual Varentropy*, studied in details in Di Crescenzo and Paolillo (2020) and Paolillo et al. (2021), defined as

$$\begin{aligned} V_e(X_t) &= \mathrm{Var}[IC(X_t)] = \mathrm{Var}[-\log f_{X_t}(X_t)] \\ &= \mathrm{Var}[\log f_{X_t}(X_t)] = \mathbb{E}\left[(\log f_{X_t}(X_t))^2\right] - [H_e(X_t)]^2 \\ &= \int_t^{+\infty} \frac{f_X(x)}{\overline{F}_X(t)} \left[\log \frac{f_X(x)}{\overline{F}_X(t)}\right]^2 \mathrm{d}x - \left[\int_t^{+\infty} \frac{f_X(x)}{\overline{F}_X(t)} \log \frac{f_X(x)}{\overline{F}_X(t)} \mathrm{d}x\right]^2, \quad t \in \mathcal{S}. \end{aligned}$$

A large number of studies in reliability theory deal with past lifetime, that is the random variable conditioned on the fact that the failure occurs before a specified inspection time t. We refer the reader to Finkelstein (2002), Nanda et al. (2003) and Kayid and Izadkhah (2014), and references therein, for results concerning past lifetime in reliability analysis (see also Barlow and Proschan, 1996). In many situations, uncertainty can refer to the past instead of the future. In fact, if we consider a system which is failed or down at time t, it could be of interest to study the uncertainty about the time in $(0, t)$ in which it has failed. Moreover, past lifetime plays a central role in the analysis of right-censored data (see, e.g., Andersen et al., 1993). For past lifetimes, as well for residual lifetimes, it can be useful to provide dynamic versions of the entropy and the varentropy, whose definitions are similar to the ones described above. To this aim, recall that given the absolutely continuous random lifetime X, having support $\mathcal{S} \subseteq \mathbb{R}^+$, cumulative distribution F_X and density f_X, its *Past Lifetime* at time $t \in \mathcal{S}$ is the variable ${}_tX = (X | X \leq t)$ whose density is

$$f_{{}_tX}(x) = \frac{f_X(x)}{F_X(t)} \qquad x \in (0, t), \tag{1.4}$$

and whose mean, known as *Mean Past Lifetime*, is given by

$$\tilde{\mu}_X(t) = \int_0^t \left(1 - \frac{F_X(x)}{F_X(t)}\right) \mathrm{d}x = t - \frac{1}{F_X(t)} \int_0^t F_X(x) \mathrm{d}x, \quad t \in \mathcal{S}. \tag{1.5}$$

The corresponding *Past Entropy* and *Past Varentropy* can be thus defined as

$$H_e({}_tX) = \mathbb{E}[IC({}_tX)] = \mathbb{E}[-\log f_{{}_tX}({}_tX)] = -\int_0^t \frac{f_X(x)}{F_X(t)} \log \frac{f_X(x)}{F_X(t)} \, \mathrm{d}x \tag{1.6}$$

and

$$\begin{aligned} V_e({}_tX) &= \mathrm{Var}[IC({}_tX)] = \mathrm{Var}[-\log f_{{}_tX}({}_tX)] \\ &= \mathrm{Var}[\log f_{{}_tX}({}_tX)] = \mathbb{E}\left[(\log f_{{}_tX}({}_tX))^2\right] - [H_e({}_tX)]^2 \\ &= \int_0^t \frac{f(x)}{F(t)} \left[\log \frac{f_X(x)}{F_X(t)}\right]^2 \mathrm{d}x - \left[\int_0^t \frac{f_X(x)}{F_X(t)} \log \frac{f_X(x)}{F_X(t)} \mathrm{d}x\right]^2, \end{aligned} \tag{1.7}$$

for all $t \in \mathcal{S}$.

The past entropy $H_e({}_tX)$ has been studied in details in Di Crescenzo and Longobardi (2002), but there are no detailed studies describing specific properties of the past varentropy $V_e({}_tX)$. The purpose of this paper is to fill this lack, providing a list of useful formulas, properties and examples for the past varentropy. As pointed out in the following sections, the past varentropy satisfies some properties which are similar to those satisfied by the residual varentropy (described in Di Crescenzo and Paolillo, 2020), but one can also observe differences. For example, while there exist at least three families of lifetimes distributions with continuous densities for which the residual varentropy is constant, on the contrary there exists only one family for which such property is satisfied by the past entropy.

2 Main results

First, we provide an alternative simple formula for the past varentropy of a random lifetime X. To this aim, recall that the reversed hazard rate function of X is defined as

$$q_X(t) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} \mathbb{P}(X \geq t - \Delta t | X \leq t) = \frac{f_X(t)}{F_X(t)}, \tag{2.1}$$

for $t \in \mathcal{S}$. The reversed hazard rate function is the instantaneous failure rate occurring immediately before the time point t, i.e., that the failure occurs just before the time point t, given that the unit has not survived longer than time t (see Block and Savits (1998) and Finkelstein (2002) for more details about the reversed hazard rate function). We recall also the notion of inverse cumulative reversed hazard rate function, that is defined as

$$Q_X(t) = \int_t^{+\infty} q_X(x) \mathrm{d}x = -\log F_X(t), \tag{2.2}$$

(see for instance Li and Li, 2008). Also, note that the past entropy can be expressed as

$$\begin{aligned} H_e({}_tX) &= -Q_X(t) - \frac{1}{F_X(t)} \int_0^t f_X(x) \log f_X(x) \mathrm{d}x \\ &= 1 - \frac{1}{F_X(t)} \int_0^t f_X(x) \log q_X(x) \mathrm{d}x, \end{aligned} \tag{2.3}$$

as shown in Di Crescenzo and Longobardi (2002). Thus, through (1.7) and (2.3), one obtains

$$\begin{aligned}
V_e({}_tX) &= \int_0^t \frac{f(x)}{F(t)}\left[\log\frac{f_X(x)}{F_X(t)}\right]^2 \mathrm{d}x - [H_e({}_tX)]^2 \\
&= \frac{1}{F_X(t)}\int_0^t f_X(x)(\log f_X(x))^2\mathrm{d}x + (\log F_X(t))^2 - \frac{2\log F_X(t)}{F_X(t)}\int_0^t f_X(x)\log f_X(x)\mathrm{d}x \\
&\quad -[H_e({}_tX)]^2 \\
&= \frac{1}{F_X(t)}\int_0^t f_X(x)(\log f_X(x))^2\mathrm{d}x + (Q_X(t))^2 + \frac{2Q_X(t)}{F_X(t)}\int_0^t f_X(x)\log f_X(x)\mathrm{d}x \\
&\quad -[H_e({}_tX)]^2 \\
&= \frac{1}{F_X(t)}\int_0^t f_X(x)(\log f_X(x))^2\mathrm{d}x + (Q_X(t))^2 - 2Q_X(t)\left[Q_X(t)+H_e({}_tX)\right] - [H_e({}_tX)]^2 \\
&= \frac{1}{F_X(t)}\int_0^t f_X(x)(\log f_X(x))^2\mathrm{d}x - (Q_X(t)+H_e({}_tX))^2, \qquad (2.4)
\end{aligned}$$

for $t \in \mathcal{S}$.

Remark 2.1. *As well as, when t tends to the supremum of the support $\mathcal{S}$, u_X, the past entropy tends to Shannon entropy, also the past varentropy reduces to the varentropy, i.e., $\lim_{t\to u_X} V_e({}_tX) = V_e(X)$.*

Now, consider the case in which t tends to the infimum of the support, l_X. If the pdf f_X of X is differentiable and such that

$$\lim_{t\to l_X^+} f_X(t) \neq 0 \quad \textit{and} \quad \lim_{t\to l_X^+} f'_X(t) \neq +\infty, \qquad (2.5)$$

then $\lim_{t\to l_X^+} V_e({}_tX) = 0$. In fact, from (2.4) the past varentropy can be expressed as

$$V_e({}_tX) = \frac{F_X(t)\int_{l_X}^t f_X(x)\left(\log f_X(x)\right)^2\mathrm{d}x - \left(\int_{l_X}^t f_X(x)\log f_X(x)\mathrm{d}x\right)^2}{F_X^2(t)},$$

and by using L'Hôpital's rule twice, it readily follows

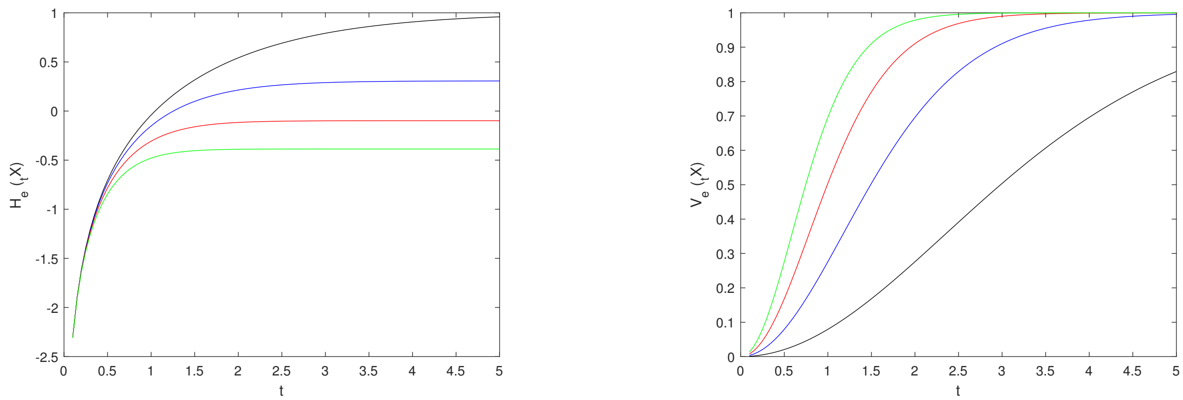
$$\begin{aligned}
\lim_{t\to l_X^+} V_e({}_tX) &= \lim_{t\to l_X^+}\left(\frac{f'_X(t)F_X(t)}{f_X^2(t)}\log f_X(t) - \frac{f'_X(t)}{f_X^2(t)}\int_{l_X}^t f_X(x)\log f_X(x)\mathrm{d}x\right) \\
&= \lim_{t\to l_X^+}\frac{f'_X(t)}{f_X^2(t)}\int_{l_X}^t \frac{F_X(x)f'_X(x)}{f_X(x)}\mathrm{d}x = 0,
\end{aligned}$$

where the last equality depends on the assumptions in (2.5).

By using (1.6) and (2.4) one can find the past entropy and past varentropy for some distributions of interest in reliability theory. Some examples are listed here.

- Let X be a random variable with uniform distribution over $(0, b)$, i.e., $X \sim U(0, b)$, $b > 0$. Hence, for $t \in (0, b)$ we have

$$H_e({}_tX) = \log t$$
$$V_e({}_tX) = 0.$$



(a) $H_e({}_tX)$ (b) $V_e({}_tX)$

Figure 1: Plots of past entropy and past varentropy of exponential distribution with parameter $\lambda = 1, 2, 3, 4$ (black, blue, red and green, respectively).

- Let X be a random variable with exponential distribution, i.e., $X \sim Exp(\lambda)$, for $\lambda > 0$. Then, for $t > 0$ we have

$$H_e({}_tX) = 1 + \log\left(\frac{1 - \mathrm{e}^{-\lambda t}}{\lambda}\right) - \frac{\lambda t \mathrm{e}^{-\lambda t}}{1 - \mathrm{e}^{-\lambda t}}$$
$$V_e({}_tX) = 1 - \frac{\lambda^2 t^2 \mathrm{e}^{-\lambda t}}{(1 - \mathrm{e}^{-\lambda t})^2}.$$

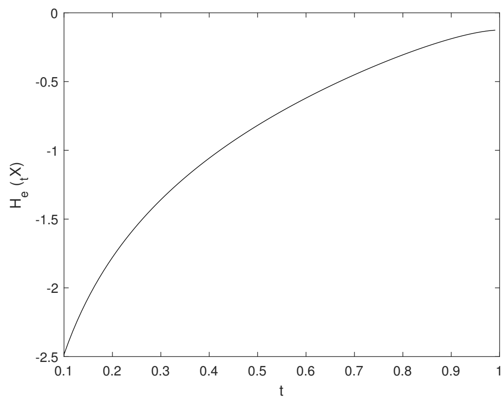
The plots of past entropy and past varentropy are shown in Figure 1 for different choices of λ. Observe that $\lim_{t \to 0^+} V_e({}_tX) = 0$ as expected since the exponential distribution satisfies the assumptions given in (2.5) for any value of the rate λ.

- Let X be a random variable such that $f_X(x) = 2x$ and $F_X(x) = x^2$, $x \in (0, 1)$. Hence, for $t \in (0, 1)$ we have

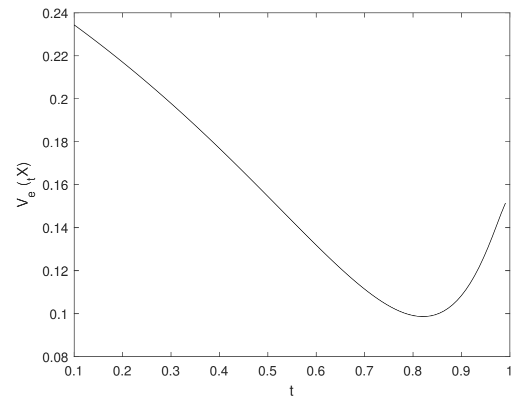
$$H_e({}_tX) = \frac{1}{2} + \log\frac{t}{2}$$
$$V_e({}_tX) = \frac{1}{4}$$

- Let X be a random variable $Beta(2, 2)$ distribution, i.e., such that $f_X(x) = 6x(1 - x)$ and $F_X(x) = 3x^2 - 2x^3$, $x \in (0, 1)$. Hence, for $t \in (0, 1)$ we have

$$H_e({}_tX) = \frac{1}{t^2/2 - t^3/3}\left[\left(\frac{t^2}{2} - \frac{t^3}{3}\right)\log\left(\frac{6(1-t)}{t(3-2t)}\right) + \frac{2}{9}t^3 - \frac{1}{3}t^2 - \frac{1}{6}t - \frac{1}{6}\log(1-t)\right]$$
$$\begin{aligned} V_e({}_tX) = \frac{1}{t^2/2 - t^3/3}\Bigg[&\left(\frac{t^2}{2} - \frac{t^3}{3}\right)\log^2\left(\frac{6(1-t)}{t(3-2t)}\right) + \frac{1}{3}\left(\frac{4}{3}t^3 - 2t^2 - t\right)\log\left(\frac{6(1-t)}{t(3-2t)}\right) \\ &+ \frac{1}{9}\left(-\frac{8}{3}t^3 + 4t^2 + 8t + 5\log(1-t)\right) - \frac{1}{3}\log\left(\frac{1}{t^2/2 - t^3/3}\right)\log(1-t) \\ &- \frac{1}{6}\log^2(1-t) - \frac{\pi^2}{18} + \frac{1}{3}Li_2(1-t)\Bigg] - [H_e({}_tX)]^2 \end{aligned}$$



(a) $H_e({}_tX)$



(b) $V_e({}_tX)$

Figure 2: Plots of past entropy and past varentropy of $X \sim Beta(2,2)$.

where Li_2 is the *Spence's function* or *dilogarithm function* (see, e.g., Morris, 1979). The plots of past entropy and past varentropy are shown in Figure 2.

- Let X be a random variable having cumulative distribution $F_X(x) = 1 - \left(\frac{b-x}{b}\right)^\alpha$, for $x \in (0,b) \subseteq \mathbb{R}^+$ and $\alpha > 0$. Then, for $t \in (0,b)$:

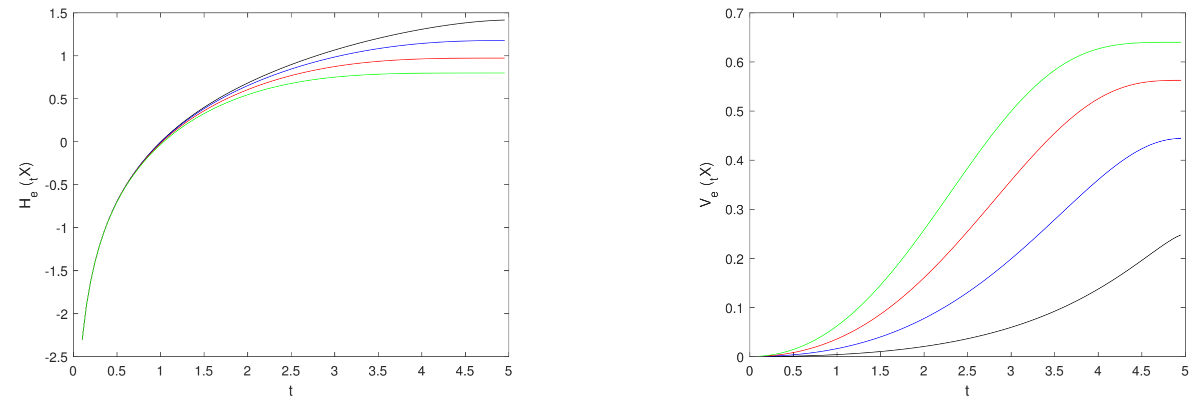
$$H_e({}_tX) = \frac{b^\alpha}{b^\alpha - (b-t)^\alpha} \log\left(\frac{\alpha b^{(\alpha-1)}}{b^\alpha - (b-t)^\alpha}\right) - \frac{(b-t)^\alpha}{b^\alpha - (b-t)^\alpha} \log\left(\frac{\alpha (b-t)^{(\alpha-1)}}{b^\alpha - (b-t)^\alpha}\right) - \frac{\alpha-1}{\alpha}$$

$$V_e({}_tX) = \left(\frac{\alpha-1}{\alpha}\right)^2 - \frac{b^\alpha (b-t)^\alpha}{[b^\alpha - (b-t)^\alpha]^2} \log^2\left[\left(\frac{b}{b-t}\right)^{\alpha-1}\right]$$

The plots of this past entropy and of the corresponding past varentropy are shown in Figure 3.

It is interesting to observe that the past varentropy is constant in two of the cases described above, increasing in one case, and non-monotone in the other one. Thus, monotonicity of the varentropy is not always guaranteed. Let us remark that, if the reversed hazard rate is decreasing for all t, then the past entropy is increasing for all t (see Di Crescenzo and Longobardi, 2002, Proposition 2.2). However, monotonicity of the reversed hazard rate is not a sufficient condition for monotonicity of the varentropy. In fact, for the $Beta(2,2)$ distribution the reversed hazard rate $q_X(t) = 6(1-t)/(3t-2t^2)$ is decreasing, while the varentropy is not monotone. For this reason, conditions for the monotonicity of $V_e({}_tX)$ and an implicit formula for the derivative of the past varentropy are now described.

Conditions for monotonicity of the past varentropy can be easily provided by using the results that appear in Paolillo et al. (2021). For it, we recall the definition of two stochastic comparisons between variables that are used in the proof. Given the random variables X_1 and X_2 having distributions F_1 and F_2, respectively, we say that X_1 *is smaller than* X_2 *in the concave order*, $X_1 \leq_c X_2$ in notation, if $F_2^{-1}(F_1(x))$ is convex on the support of F_1. We say that X_1 *is smaller*



(a) $H_e({}_tX)$ (b) $V_e({}_tX)$

Figure 3: Plots of past entropy and past varentropy of X with cdf $F_X(x) = 1 - \left(\frac{b-x}{b}\right)^\alpha$ for $b = 5$ and $\alpha = 2, 3, 4, 5$ (black, blue, red and green, respectively).

than X_2 *in the starshaped order*, $X_1 \leq_* X_2$ in notation, if $F_2^{-1}(F_1(x))/x$ is increasing on the support of F_1. Details and applications of these stochastic orders can be found in Shaked and Shanthikumar (2007).

Proposition 2.1. *Let* X *be a random lifetime with an absolutely continuous distribution* F_X *and a strictly decreasing [increasing] density function* f_X. *If the ratio*

$$\frac{f_X(F_X^{-1}(pF(s)))}{f_X(F_X^{-1}(pF(t)))} \tag{2.6}$$

is increasing in $p \in (0, 1)$ *for all* $s \leq t$, *then the corresponding past varentropy* $V_e({}_tX)$ *is increasing [decreasing] in* $t \in \mathcal{S}$.

Proof. Recall that, for any $t \in \mathcal{S}$, the past lifetime ${}_tX$ has density $f_{{}_tX}(x) = f_X(x)/F_X(t)$ and cumulative distribution $F_{{}_tX}(x) = F_X(x)/F_X(t)$, with $x \leq t$. Thus, the corresponding quantile function is $F_{{}_tX}^{-1}(p) = F_X^{-1}(pF_X(t))$, for $p \in (0, 1)$. Also observe that, for $s \leq t$,

$$\frac{f_{{}_sX}(F_{{}_sX}^{-1}(p))}{f_{{}_tX}(F_{{}_tX}^{-1}(p))} = \frac{f_X(F_X^{-1}(pF_X(s)))}{F_X(s)} \cdot \frac{F_X(t)}{f_X(F_X^{-1}(pF_X(t)))} = \frac{f_X(F_X^{-1}(pF(s)))}{f_X(F_X^{-1}(pF(t)))} \cdot \frac{F_X(t)}{F_X(s)},$$

where the latter is increasing in p by assumption (2.6). Then one has ${}_sX \leq_c {}_tX$ (see Remark 4.3 in Paolillo et al., 2021, or Section 4.2 in Shaked and Shanthikumar, 2007). Now observe that, since f_X is decreasing [increasing] by assumption, then also $f_{{}_sX}$ and $f_{{}_tX}$ are decreasing [increasing]. Thus, by the equivalence pointed out in Remark 4.6 in Paolillo et al. (2021), one also has that $f_{{}_sX}({}_sX) \leq_* f_{{}_tX}({}_tX)$ $[f_{{}_sX}({}_sX) \geq_* f_{{}_tX}({}_tX)]$, which, in turns, implies $V_e({}_sX) \leq V_e({}_tX)$ $[V_e({}_sX) \geq V_e({}_tX)]$ by Theorem 5.2 in the same paper. □

It is easy to verify, for example, that exponential distributions satisfy the assumptions of Proposition 2.1 for any value of the rate λ.

The following result provides an implicit formula for the derivative of the past varentropy, useful to describe distributions having constant varentropy.

Proposition 2.2. *For all $t \in \mathcal{S}$, the derivative of the past varentropy is*

$$V_e'({}_tX) = -q_X(t)\left[V_e({}_tX) - (H_e({}_tX) + \log q_X(t))^2\right].$$

Proof. First observe that by differentiating both sides of (2.3) we get the following expression for the derivative of the past entropy:

$$H_e'({}_tX) = q_X(t)[1 - H({}_tX) - \log q_X(t)]. \tag{2.7}$$

Consider now (2.4). By differentiating both sides we get

$$\begin{aligned} V_e'({}_tX) \quad &= \quad \frac{q_X(t)}{F_X(t)}\int_0^t f_X(x)(\log f_X(x))^2 \mathrm{d}x + q_X(t)(\log f_X(t))^2 \\ &\quad -2(Q_X(t) + H_e({}_tX))(-q_X(t) + H_e'({}_tX)), \end{aligned} \tag{2.8}$$

where $q_X(t)$ is defined in (2.1). Hence, recalling (2.7) and (2.4), from (2.8) we get

$$\begin{aligned} V_e'({}_tX) \quad &= \quad -q_X(t)\left[V_e({}_tX) + (Q_X(t) + H_e({}_tX))^2 - (\log f_X(t))^2\right. \\ &\quad \left. -2(Q_X(t) + H_e({}_tX))(H_e({}_tX) + \log q_X(t))\right], \end{aligned}$$

and, after straightforward calculations, one gets the statement. □

From (2.7) one can obtain conditions such that absolutely continuous distributions, having continuous densities, have a corresponding constant past varentropy. To this aim, consider first the case of random variables having support $\mathcal{S} = [0, 1]$.

Proposition 2.3. *Let X have support $\mathcal{S} = [0, 1]$. Then, its varentropy $V_e({}_tX)$ is constant if, and only if, X has cumulative distribution function*

$$F_X(x) = x^\alpha, \qquad x \in [0, 1], \tag{2.9}$$

for a parameter $\alpha > 0$. In this case, one has $V_e({}_tX) = (1 - 1/\alpha)^2$ for all $t \in [0, 1]$.

Proof. First observe that, if X has cumulative distribution defined as in (2.9), then the pdf is given by

$$f_X(x) = \alpha x^{\alpha-1}, \qquad x \in (0, 1),$$

and, for $t \in (0, 1)$, the past varentropy is defined as

$$V_e({}_tX) = \int_0^t \frac{\alpha x^{\alpha-1}}{t^\alpha}\left[\log\left(\frac{\alpha x^{\alpha-1}}{t^\alpha}\right)\right]^2 \mathrm{d}x - \left[\int_0^t \frac{\alpha x^{\alpha-1}}{t^\alpha}\log\left(\frac{\alpha x^{\alpha-1}}{t^\alpha}\right)\mathrm{d}x\right]^2.$$

By the change of variable $y = \left(\frac{x}{t}\right)^{\alpha}$, we get

$$\begin{aligned}
V_e({}_tX) &= \int_0^1 \left[\log\left(\frac{\alpha y^{(\alpha-1)/\alpha}}{t}\right)\right]^2 \mathrm{d}y - \left[\int_0^1 \log\left(\frac{\alpha y^{(\alpha-1)/\alpha}}{t}\right)\mathrm{d}y\right]^2 \\
&= \log^2\left(\frac{\alpha}{t}\right) - \int_0^1 2\left(\frac{\alpha-1}{\alpha}\right)\log\left(\frac{\alpha y^{(\alpha-1)/\alpha}}{t}\right)\mathrm{d}y - \left[\log\left(\frac{\alpha}{t}\right) - \frac{\alpha-1}{\alpha}\right]^2 \\
&= \log^2\left(\frac{\alpha}{t}\right) - 2\left(\frac{\alpha-1}{\alpha}\right)\log\left(\frac{\alpha}{t}\right) + 2\left(\frac{\alpha-1}{\alpha}\right)^2 - \left[\log\left(\frac{\alpha}{t}\right) - \frac{\alpha-1}{\alpha}\right]^2.
\end{aligned}$$

Thus $V_e({}_tX)$ is constant and equal to $(1-1/\alpha)^2$. It follows now, from Proposition 2.2, that

$$(H_e({}_tX) + \log q_X(t))^2 = (1-1/\alpha)^2,$$

and so

$$|H_e({}_tX) + \log q_X(t)| = |1-1/\alpha|, \qquad \forall t \in [0,1]. \tag{2.10}$$

Since the density f_X is continuous by assumption, then also q_X and $H_e(X_t)$ are continuous. Thus, $H_e({}_tX) + \log q_X(t)$ is continuous in $t \in [0,1]$, so that equality (2.10) implies

$$H_e({}_tX) + \log q_X(t) = c, \qquad \forall t \in [0,1]. \tag{2.11}$$

for some $c \in \mathbb{R}$.

As shown in Kundu et al. (2010), Theorem 2.1, there exist only three families of distribution for which (2.11) is satisfied. Two of them have infinite support on the left, i.e., of the form $(-\infty, b]$, for $b \in \mathbb{R}$ (thus they can not be distributions of random lifetimes), and the only one having support entirely contained in $\mathbb{R}^+$ (and in $[0,1]$ in particular) is the one defined in (2.9). Finally, since the family defined in (2.9) has constant past varentropy, the assertion follows. $\square$

To generalize the above result to random lifetimes having different supports, we can use the following proposition, that deals with the past varentropy under linear transformations. We recall that if $Y = aX + b$ for $a > 0$ and $b \geq 0$, then the past entropies of X and Y are related by

$$H_e({}_tY) = H_e\left({}_{\frac{t-b}{a}}X\right) + \log a \qquad \forall t, \tag{2.12}$$

(see Di Crescenzo and Longobardi, 2002).

Proposition 2.4. *Let $Y = aX + b$, with $a > 0$ and $b \geq 0$. Then, for their past varentropies, we have*

$$V_e({}_tY) = V_e\left({}_{\frac{t-b}{a}}X\right), \qquad \forall t. \tag{2.13}$$

Proof. From $Y = aX + b$ we know that $F_Y(x) = F_X\left(\frac{x-b}{a}\right)$ and $f_Y(x) = \frac{1}{a} f_X\left(\frac{x-b}{a}\right)$. Hence, from (1.7) and (2.12), we get

$$V_e({}_tY) = \int_0^{\frac{t-b}{a}} \frac{f_X(x)}{F_X\left(\frac{t-b}{a}\right)} \left(\log \frac{\frac{1}{a} f_X(x)}{F_X\left(\frac{t-b}{a}\right)}\right)^2 \mathrm{d}x - \left(H\left({}_{\frac{t-b}{a}}X\right) + \log a\right)^2. \tag{2.14}$$

By writing

$$\log \frac{\frac{1}{a} f_X\left(x\right)}{F_X\left(\frac{t-b}{a}\right)} = \log \frac{f_X\left(x\right)}{F_X\left(\frac{t-b}{a}\right)} - \log a,$$

and developing the two squares in (2.14), one easily obtains the statement. □

From Propositions 2.3 and 2.4 one immediately gets the following statement.

Corollary 2.1. *Let X be an absolutely continuous random lifetime with continuous density f_X. Then, its varentropy $V_e({}_tX)$ is constant if, and only if, X has cumulative distribution function in the family*

$$F_X(x) = \left(\frac{x-b}{a}\right)^{\alpha}, \qquad x \in [b, a+b], \tag{2.15}$$

for a parameter α such that $\alpha > 0$.

Apart for the property stated in Corollary 2.1, the family defined in (2.15) is the only one of lifetimes distribution having continuous density that satisfies the property stated in the next proposition. Recall first that the *Generalized Reversed Hazard Rate* of a random lifetime is defined, for $\gamma \in \mathbb{R}$, as

$$q_{\gamma,X}(t) = \frac{f_X(t)}{[F_X(t)]^{1-\gamma}}, \qquad t \in \mathcal{S} \tag{2.16}$$

(see Buono et al. (2021), where their applications in the study of properties of aging intensity functions are described). We remark that, by choosing $\gamma = 0$ in (2.16), we get $q_{0,X}(t) = q_X(t)$, i.e., $q_{0,X}$ is the usual reversed hazard rate function. Let us observe that for $\gamma = 1$ the generalized reversed hazard rate function is equal to the density function. This is reasonable since the density function gives a first rough illustration of the aging tendency of the random variable by its monotonicity.

Proposition 2.5. *Let X be a random lifetime having continuous density, and let $\gamma \in \mathbb{R}$. Its generalized reversed hazard rate function $q_{\gamma,X}(t)$, with parameter γ, is constant if, and only if, F_X is in the family defined in* (2.15) *and $\gamma = 1/\alpha$. Moreover, in this case one has*

$$q_{1/\alpha,X}(t) = \frac{f_X(t)}{[F_X(t)]^{1-1/\alpha}} = \mathrm{e}^{1-1/\alpha-H_e(X)}, \ \forall t \in [b, a+b]. \tag{2.17}$$

Proof. Let us suppose that there exists $c \in \mathbb{R}$ such that $q_{1-c,X}(t) = \mathrm{e}^{c-H_e(X)}$ for all $t \in \mathcal{S}$, being $\mathcal{S}$ the support of X. From (2.1) and (2.3) we have

$$\begin{aligned} H_e({}_tX) + \log q_X(t) &= \log f_X(t) - \frac{1}{F_X(t)} \int_{(0,t)\bigcap \mathcal{S}} f_X(x) \log f_X(x) \mathrm{d}x \\ &= \log f_X(t) + \frac{1}{F_X(t)} \left[H_e(X) + \int_{(t,+\infty)\bigcap \mathcal{S}} f_X(x) \log f_X(x) \mathrm{d}x \right]. \end{aligned}$$

Moreover, from the hypothesis, we get

$$\int_{(t,+\infty)\bigcap \mathcal{S}} f_X(x) \log f_X(x) \mathrm{d}x = -H_e(X)\overline{F}_X(t) - cF_X(t) \log F_X(t),$$

and so

$$H_e({}_tX) + \log q_X(t) = H_e(X) + \log \frac{f_X(t)}{[F_X(t)]^c} = c.$$

This last equality is satisfied only for distributions in the family described in (2.15), with $c = 1-1/\alpha$.

Conversely, if X has distributions in the family described in (2.15), then, with a direct calculation, one can verify that (2.17) holds. $\square$

A generalization of Proposition 2.1 will now be stated. For it, let ϕ be a differentiable and strictly monotonic function and let $Y = \phi(X)$ for a given X. It has been shown in Di Crescenzo and Longobardi (2002) that the past entropies of X and Y are related by the equations

$$H_e({}_tY) = \begin{cases} H_e\left({}_{\phi^{-1}(t)}X\right) + \mathbb{E}[\log \phi'(X)|X < \phi^{-1}(t)], & \text{if } \phi \text{ is strictly increasing,} \\ H_e\left(X_{\phi^{-1}(t)}\right) + \mathbb{E}[\log(-\phi'(X))|X > \phi^{-1}(t)], & \text{if } \phi \text{ is strictly decreasing.} \end{cases} \tag{2.18}$$

Similar results can be proved for the past varentropy.

Proposition 2.6. *Let $Y = \phi(X)$, where ϕ is a differentiable and strictly monotonic function. Then, if ϕ is strictly increasing, for the past varentropy of Y we have*

$$\begin{aligned} V_e({}_tY) &= V_e\left({}_{\phi^{-1}(t)}X\right) - 2\mathbb{E}\left[\log \frac{f_X(X)}{F_X(\phi^{-1}(t))} \log \phi'(X) \middle| X < \phi^{-1}(t)\right] \\ &\quad + \mathrm{Var}[\log \phi'(X)|X < \phi^{-1}(t)] - 2H_e\left({}_{\phi^{-1}(t)}X\right) \mathbb{E}[\log \phi'(X)|X < \phi^{-1}(t)], \end{aligned} \tag{2.19}$$

whereas, if ϕ is strictly decreasing

$$\begin{aligned} V_e({}_tY) &= V_e\left(X_{\phi^{-1}(t)}\right) - 2\mathbb{E}\left[\log \frac{f_X(X)}{\overline{F}_X(\phi^{-1}(t))} \log(-\phi'(X)) \middle| X > \phi^{-1}(t)\right] \\ &+ \mathrm{Var}[\log(-\phi'(X))|X > \phi^{-1}(t)] - 2H\left(X_{\phi^{-1}(t)}\right) \mathbb{E}[\log(-\phi'(X))|X > \phi^{-1}(t)]. \end{aligned} \tag{2.20}$$

Proof. Suppose first that ϕ is strictly increasing. From $Y = \phi(X)$ we know that $F_Y(x) = F_X\left(\phi^{-1}(x)\right)$ and $f_Y(x) = \frac{f_X(\phi^{-1}(x))}{\phi'(\phi^{-1}(x))}$. Hence, from (1.7) and (2.18), we get

$$\begin{aligned} V_e({}_tY) &= \int_0^{\phi^{-1}(t)} \frac{f_X(x)}{F_X\left(\phi^{-1}(t)\right)} \left(\log \frac{f_X(x)}{F_X\left(\phi^{-1}(t)\right)} - \log \phi'(x)\right)^2 \mathrm{d}x \\ &\quad - \left[H_e\left({}_{\phi^{-1}(t)}X\right) + \mathbb{E}[\log \phi'(X)|X < \phi^{-1}(t)]\right]^2. \end{aligned}$$

Now, by developing the two squares in the previous equality, and observing that

$$\begin{aligned} \int_0^{\phi^{-1}(t)} \frac{f_X(x)}{F_X\left(\phi^{-1}(t)\right)} \left(\log \phi'(x)\right)^2 \mathrm{d}x &- \mathbb{E}^2[\log \phi'(X)|X < \phi^{-1}(t)] \\ &= \mathrm{Var}[\log \phi'(X)|X < \phi^{-1}(t)], \end{aligned}$$

we obtain the result.

The proof is similar if ϕ is strictly decreasing. $\square$

Example 2.1. *The* Inverted Exponential *distribution (invExp), introduced as a lifetime model in Lin et al. (1989), has been considered by many authors in reliability studies (see, e.g., Krishna and Kumar, 2012, or Oguntunde et al., 2014, and references therein). The past varentropy of an inverted exponential distribution can be actually obtained by using Proposition 2.6. To this aim, consider $X \sim Exp(\lambda)$ and $Y = \phi(X) = 1/X$ so that ϕ is strictly decreasing and $Y \sim invExp(\lambda)$. We can use the result presented in* (2.20) *to evaluate the past varentropy of Y. In fact, we have*

$$\begin{aligned} V_e({}_tY) &= V_e\left(X_{1/t}\right) - 2\mathbb{E}\left[\log\frac{\lambda e^{-\lambda X}}{e^{-\lambda/t}}\log\left(\frac{1}{X^2}\right)\Big| X > \frac{1}{t}\right] \\ &\quad + \mathrm{Var}\left[\log\left(\frac{1}{X^2}\right)\Big| X > \frac{1}{t}\right] - 2H\left(X_{1/t}\right)\mathbb{E}\left[\log\left(\frac{1}{X^2}\right)\Big| X > \frac{1}{t}\right]. \end{aligned}$$

The residual entropy and the residual varentropy for the exponential distribution are given as

$$H_e(X_t) = 1 - \log\lambda, \quad V_e(X_t) = 1,$$

and then the past varentropy of Y is expressed as

$$\begin{aligned} V_e({}_tY) &= 1 - 2\mathbb{E}\left[\log\frac{\lambda e^{-\lambda X}}{e^{-\lambda/t}}\log\left(\frac{1}{X^2}\right)\Big| X > \frac{1}{t}\right] \\ &\quad + \mathrm{Var}\left[\log\left(\frac{1}{X^2}\right)\Big| X > \frac{1}{t}\right] - 2(1-\log\lambda)\mathbb{E}\left[\log\left(\frac{1}{X^2}\right)\Big| X > \frac{1}{t}\right]. \end{aligned}$$

With several calculations, the above expression reduces to

$$\begin{aligned} V_e({}_tY) &= -3 + \frac{4\lambda}{t}\log\frac{1}{t^2} + \frac{8t}{\lambda} + \left(8 - \frac{4\lambda}{t}\right)\frac{1}{e^{-\lambda/t}}Ei\left(-\frac{\lambda}{t}\right) - \frac{4}{e^{-2\lambda/t}}Ei^2\left(-\frac{\lambda}{t}\right) \\ &\quad + \frac{4}{e^{-\lambda/t}}\log\frac{1}{t^2}Ei\left(-\frac{\lambda}{t}\right) - \frac{4}{\lambda e^{-\lambda/t}}\int_{1/t}^{+\infty}\frac{\log x^2}{x^2}e^{-\lambda x}\mathrm{d}x, \end{aligned}$$

where $Ei(\cdot)$ is the exponential integral function (see, e.g., Gautschi and Gahill, 1972). The plot of this past varentropy is shown in Figure 4 for different choices of λ.

We conclude this section pointing out that there exists a strong relationship between the past varentropy and the residual varentropy of an absolutely continuous random lifetime X whenever its support $\mathcal{S}$ is finite. Without loss of generality let X assume values in $[0, u_X]$, and let f_X be its density function. Then consider a random lifetime $\tilde{X}$ whose density is the symmetric of f_X with respect to $u_X/2$, i.e., the lifetime having density $f_{\tilde{X}}(x) = f_X(u_X - x)$ for all $x \in [0, u_X]$. It is easy to observe that X and $\tilde{X}$ have the same information content, i.e., that $IC(X)$ and $IC(\tilde{X})$ have the same distribution. Let now $t \in [0, u_X]$, and consider the conditioned lifetimes $(X|X > t)$ and $(\tilde{X}|\tilde{X} \leq u_X - t)$. Again, it is easy to verify that the corresponding densities are symmetric, i.e., that $f_{(\tilde{X}|\tilde{X}\leq u_X - t)}(x) = f_{(X|X>t)}(u_X - x)$, so that $IC(X|X > t)$ and $IC(\tilde{X}|\tilde{X} \leq u_X - t)$ have the same distribution. It obviously follows that, for all $t \in [0, u_X]$, it holds

$$H_e(\tilde{X}|\tilde{X} \leq u_X - t) = H_e(X|X > t) \quad \text{and} \quad V_e(\tilde{X}|\tilde{X} \leq u_X - t) = V_e(X|X > t).$$

Thus properties and explicit expressions of the past entropy and past varentropy of a random lifetime with finite support can be obtained from properties and explicit expressions of the corresponding residual entropy and residual varentropy, after an appropriate transformation of the density.

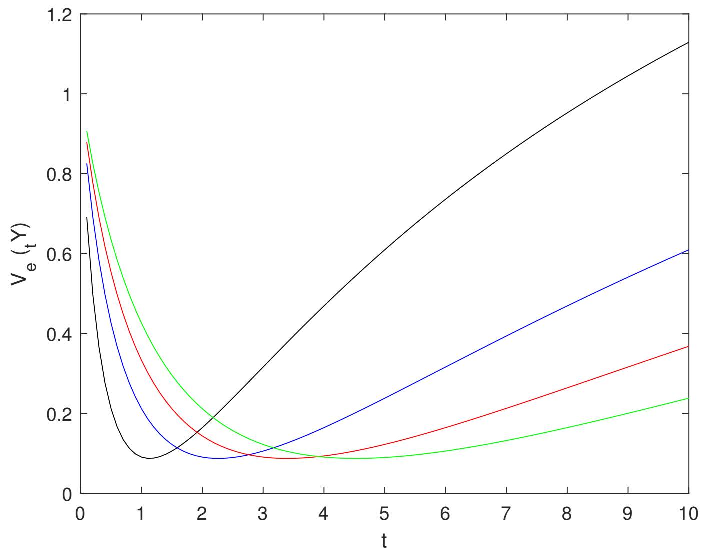


Figure 4: Plots of past varentropies of inverse exponential distributions with parameter $\lambda = 1, 2, 3, 4$ (black, blue, red and green, respectively).

3 Bounds for the past varentropy

A very simple upper bound for the past varentropy can be provided for a large class of distributions, as stated in the next proposition.

Proposition 3.1. *Let X be a non-negative random variable with support $\mathcal{S}$ and log-concave pdf $f_X(x)$. Then*

$$V_e({}_tX) \leq 1 \qquad \textit{for all} \;\; t \in \mathcal{S}.$$

Proof. We observe that if $f_X(x)$ is log-concave, then also $f_{{}_tX}(x) = \frac{f_X(x)}{F_X(t)}$ is log-concave. From Theorem 2.3 of Fradelizi et al. (2016), we know that if X has a log-concave pdf, then $V_e(X) \leq 1$ and the proof follows from this result. $\square$

For example, the density $f_X(x) = 6x(1-x),\ x \in [0,1]$ of $X \sim Beta(2,2)$ is logconcave, so that the past varentropy of X is always smaller than 1, as confirmed by its plot shown in Figure 2.

However, by comparing this bound with the plot of $V_e({}_tX)$, one can immediately observe that it is a really large bound. Better upper bounds can be provided, for any X, by using results available in the literature. For it recall that, for a random lifetime X, the corresponding *Inactivity Time* at t is defined as $X_{(t)} = (t - X | X \leq t) = t - {}_tX$, i.e. the random time whose density is

$$f_{X_{(t)}}(x) = \frac{f_X(t-x)}{F_X(t)}.$$

The following upper bound for $\mathrm{Var}[-\log f_{X_{(t)}}(X_{(t)})]$ has been proved in Goodarzi et al. (2016), Proposition 1, making use of an upper bound for variances proved in Cacoullos and Papathanasiou

(1985):

$$\mathrm{Var}[-\log f_{X_{(t)}}(X_{(t)})] \leq \mathbb{E}\left[\frac{\eta_X^2(t-X_{(t)})}{q_X(t-X_{(t)})}\left(m_X(t-X_{(t)})-m_X(t)+X_{(t)}\right)\right] \tag{3.1}$$

for all $t \in \mathcal{S}$, where $\eta_X(x) = -f_X'(x)/f_X(x)$ is the *eta function* and $m_X(x) = \mathbb{E}[X_{(x)}] = x - \tilde{\mu}_X(x)$ is the *mean inactivity time function*, and where $\tilde{\mu}_X(x)$ is defined in (1.5). Now observe that

$$\begin{aligned}
V_e(X_{(t)}) &= \mathrm{Var}[-\log f_{X_{(t)}}(X_{(t)})] \\
&= \mathbb{E}[\log^2 f_{X_{(t)}}(X_{(t)})] - \left[H_e(X_{(t)}))\right]^2 \\
&= \int_0^t \frac{f_X(t-x)}{F_X(t)}\log^2\left(\frac{f_X(t-x)}{F_X(t)}\right)\mathrm{d}x - \left[\int_0^t \frac{f_X(t-x)}{F_X(t)}\log\left(\frac{f_X(t-x)}{F_X(t)}\right)\mathrm{d}x\right]^2 \\
&= \int_0^t \frac{f_X(x)}{F_X(t)}\log^2\left(\frac{f_X(x)}{F_X(t)}\right)\mathrm{d}x - \left[\int_0^t \frac{f_X(x)}{F_X(t)}\log\left(\frac{f_X(x)}{F_X(t)}\right)\mathrm{d}x\right]^2 \\
&= \mathrm{Var}[-\log f_{{}_tX}({}_tX)] \\
&= V_e({}_tX).
\end{aligned}$$

Thus, recalling that $m_X(x) = x - \tilde{\mu}_X(x)$ and $X_{(t)} = t - {}_tX$, from (3.1) one gets the upper bound

$$V_e({}_tX) \leq \mathbb{E}\left[\frac{\eta^2({}_tX)}{q_X({}_tX)}\left(\tilde{\mu}(t) - \tilde{\mu}({}_tX)\right)\right] \quad \forall\, t \in \mathcal{S}.$$

A lower bound for the past varentropy can also be proved. For it, define first the *variance past lifetime function* $\tilde{\nu}_X^2$ as

$$\tilde{\nu}_X^2(t) = \mathrm{Var}({}_tX) = \mathrm{Var}(X|X \leq t) = \frac{1}{F_X(t)}\int_0^t x^2 f_X(x)\mathrm{d}x - (\tilde{\mu}_X(t))^2, \quad t \in \mathcal{S}.$$

Note that, for every $t \in \mathcal{S}$ the variance past lifetime function $\tilde{\nu}_X^2(t)$ is the same as the variance of the inactivity time $X_{(t)}$ (see, e.g., Kandil et al., 2011, for details and properties of the variance of the inactivity time function).

Proposition 3.2. *Let ${}_tX$ be the past lifetime of X at time t, and let the mean past lifetime $\tilde{\mu}_X(t)$ and the variance past lifetime $\tilde{\nu}_X^2(t)$ be finite for all $t \in \mathcal{S}$. Then*

$$V_e({}_tX) \geq \tilde{\nu}_X^2(t)\left[\mathbb{E}(\omega_t'({}_tX))\right]^2,$$

where the function $\omega_t(x)$ is defined by solving the equation

$$\tilde{\nu}_X^2(t)\omega_t(x)f_{{}_tX}(x) = \int_0^x (\tilde{\mu}_X(z) - z)f_{{}_tX}(z)\mathrm{d}z, \quad x \in \mathcal{S}. \tag{3.2}$$

Proof. Recall that if X is a random variable with pdf f_X, mean μ_X and variance σ_X^2, then

$$\mathrm{Var}[g(X)] \geq \sigma^2\left[\mathbb{E}(\omega(X)g'(X))\right]^2, \tag{3.3}$$

where $\omega(x)$ is defined by $\sigma^2\omega(x)f(x) = \int_0^x (\mu - z)f(z)\mathrm{d}z$ (see Cacoullos and Papathanasiou, 1989). Hence, in (3.3) choosing $g(x) = -\log f_{{}_tX}(x)$ and ${}_tX$ as X, one obtains

$$\mathrm{Var}(-\log f_{{}_tX}({}_tX)) \geq \tilde{\nu}_X^2(t)\left[\mathbb{E}\left(\omega_t({}_tX)\frac{f_{{}_tX}'({}_tX)}{f_{{}_tX}({}_tX)}\right)\right]^2. \tag{3.4}$$

By differentiating both sides of (3.2), one has

$$\omega(x)\frac{f'_{{}_TX}(x)}{f_{{}_TX}(x)} = \frac{\tilde{\mu}_X(x) - x}{\tilde{\nu}_X^2(t)} - \omega_t'(x),$$

and then, from (3.4),

$$\begin{aligned} V({}_tX) &= \mathrm{Var}(-\log f_{{}_tX}({}_tX)) \\ &\geq \tilde{\nu}_X^2(t)\left[\mathbb{E}\left(\frac{\tilde{\mu}_X(t) - {}_tX}{\tilde{\nu}_X^2(t)} - \omega'({}_tX)\right)\right]^2 \\ &= \tilde{\nu}_X^2(t)\left[\mathbb{E}(\omega_t'({}_tX))\right]^2. \end{aligned}$$

□

4 Past varentropy and parallel systems

When the past varentropy $V_e({}_tX)$ of a random lifetime X is available, then in some cases it is possible to easily compute the past varentropy of another lifetime Y whose distribution is a transformation of that one of X. An example is given by the scale model: the family of random variables $\{X^{(a)} : a > 0\}$ follows a *Scale model* if there exists a non-negative random variable X with cumulative distribution function F and density f such that $X^{(a)}$ has distribution $F^{(a)}(t) = F(at)$ for all t, where $a > 0$ is the parameter of the model. Some examples are the exponential, Weibull (with a fixed shape parameter) and Pareto (with a fixed shape parameter) distributions. In these cases, from Proposition 2.4 one immediately obtains that

$$V_e({}_tX^{(a)}) = V_e\left({}_{at}X\right), \qquad \forall t.$$

A more interesting case is when the family of random variables $\{X^{(a)} : a > 0\}$ follows a *Proportional Reversed Hazard Rate model*, i.e., if there exists a non-negative random variable X with cumulative distribution function F_X and density f_X such that

$$F^{(a)}(t) = \mathbb{P}(X^{(a)} \leq t) = [F_X(t)]^a, \qquad f^{(a)}(t) = a[F_X(t)]^{(a-1)} f_X(t), \quad t \in \mathcal{S}, \tag{4.1}$$

being $F^{(a)}$ and $f^{(a)}$ the cumulative distribution function and the density of $X^{(a)}$, respectively (see Gupta and Gupta (2007) for more details). We remark that the model takes the name from the fact that the reversed hazard rate functions of the random variables in the family are proportional to the reversed hazard rate function of X; in fact, letting $q^{(a)}$ be the reverse hazard rate of $X^{(a)}$,

$$q^{(a)}(t) = \frac{f^{(a)}(t)}{F^{(a)}(t)} = a\ \frac{f_X(t)}{F_X(t)} = a\ q_X(t) \qquad \forall t \in \mathcal{S}.$$

Moreover, we note that the inverse cumulative reversed hazard rate function is expressed as

$$Q_{X^{(a)}}(t) = -\log F^{(a)}(t) = aQ_X(t).$$

The proportional reversed hazard rate model finds applications, for example, in analysis of parallel systems. In fact, if we have a system composed by n units in parallel and characterized by i.i.d.

lifetimes $X_1, \dots, X_n$ with distribution $F_X(t)$, then the lifetime of the system is given by $X^{(n)} = \max\{X_1, \dots, X_n\}$. Then, we have $F_{X^{(n)}}(t) = [F_X(t)]^n$, i.e., the system satisfies the proportional reversed hazard rate model (4.1) with $a = n$. The purpose of the next examples is to highlight the behavior of the past varentropy when it refers to the lifetime of a parallel system with i.i.d. components.

To this aim, let us first evaluate the past entropy of $X^{(a)}$ and the past varentropy of $X^{(a)}$ for an arbitrary $a > 0$. One has

$$\begin{aligned} H_e\left({}_tX^{(a)}\right) &= -Q_{X^{(a)}}(t) - \frac{1}{[F_X(t)]^a}\int_0^t f^{(a)}(x)\log f^{(a)}(x)\mathrm{d}x \\ &= -aQ_X(t) - \frac{1}{[F_X(t)]^a}\int_0^{[F_X(t)]^a} \gamma(y;a)\mathrm{d}y, \end{aligned}$$

with the change of variable $y = [F_X(x)]^a$, and where $\gamma(y;a) = \log\left[ay^{1-1/a} f_X(F_X^{-1}(y^{1/a}))\right]$. Hence, we obtain the past varentropy of $X^{(a)}$ as

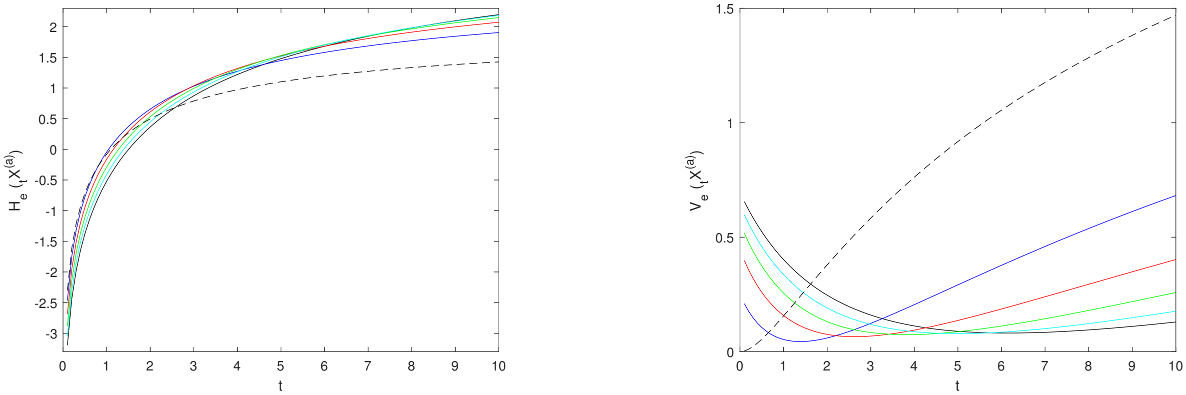
$$\begin{aligned} V_e\left({}_tX^{(a)}\right) &= \frac{1}{[F_X(t)]^a}\int_0^t f^{(a)}(x)(\log f^{(a)}(x))^2\mathrm{d}x - \left[\frac{1}{[F_X(t)]^a}\int_0^t f^{(a)}(x)\log f^{(a)}(x)\mathrm{d}x\right]^2 \\ &= \frac{1}{[F_X(t)]^a}\int_0^{[F_X(t)]^a} [\gamma(y;a)]^2\mathrm{d}y - \left[\frac{1}{[F_X(t)]^a}\int_0^{[F_X(t)]^a} \gamma(y;a)\mathrm{d}y\right]^2. \end{aligned}$$

Let us now consider, as an example, the case where X has a modified Pareto distribution with $F_X(t) = t/(1+t)$ and density $f_X(t) = 1/(1+t)^2$, for $t \geq 0$. In this case, $\gamma(y;a) = \log\left[ay^{1-1/a}(1-y^{1/a})^2\right]$, so that

$$\begin{aligned} H_e\left({}_tX^{(a)}\right) &= a\log\left(\frac{t}{1+t}\right) - \frac{1}{[t/(1+t)]^a}\int_0^{[t/(1+t)]^a} \gamma(y;a)\mathrm{d}y, \\ V_e\left({}_tX^{(a)}\right) &= \frac{1}{[t/(1+t)]^a}\int_0^{[t/(1+t)]^a} [\gamma(y;a)]^2\mathrm{d}y - \left[\frac{1}{[t/(1+t)]^a}\int_0^{[t/(1+t)]^a} \gamma(y;a)\mathrm{d}y\right]^2. \end{aligned}$$

When a is an integer, i.e., when $X^{(a)}$ represent the lifetime of a parallel system of a number a of i.i.d. components, one obtains the past entropies and past varentropies shown in Figure 5 (for different integer values of a). It is interesting to observe that both the past entropies and the past varentropies intersect each other for different values of a: for small values of the time t one has the smaller past entropies and larger past varentropies when the number of components in parallel is large, and viceversa for large values of the time t. It means, for example, that in the long run (for large values of the inspection time t) the uncertainty of the information content of the past lifetime of a parallel system reduces as the number of components in the system increases (and viceversa for small t).

The same can be observed when X has an exponential distribution with parameter λ. In this



(a) $H_e({}_tX^{(a)})$ (b) $V_e({}_tX^{(a)})$

Figure 5: Plots of past entropy and the past varentropy of modified Pareto PRHR model for $a = 1$ (dashed line) and $a = 2, 3, 4, 5, 6$ (blue, red, green, cyan and black, respectively).

case, $\gamma(y; a) = \log\left[\lambda a y^{1-1/a}(1 - y^{1/a})\right]$, so that

$$H_e\left({}_tX^{(a)}\right) = a\log\left(1 - \mathrm{e}^{-\lambda t}\right) - \frac{1}{[1-\mathrm{e}^{-\lambda t}]^a}\int_0^{[1-\mathrm{e}^{-\lambda t}]^a} \gamma(y; a)\mathrm{d}y,$$

$$V_e\left({}_tX^{(a)}\right) = \frac{1}{[1-\mathrm{e}^{-\lambda t}]^a}\int_0^{[1-\mathrm{e}^{-\lambda t}]^a} [\gamma(y; a)]^2\mathrm{d}y - \left[\frac{1}{[1-\mathrm{e}^{-\lambda t}]^a}\int_0^{[1-\mathrm{e}^{-\lambda t}]^a} \gamma(y; a)\mathrm{d}y\right]^2.$$

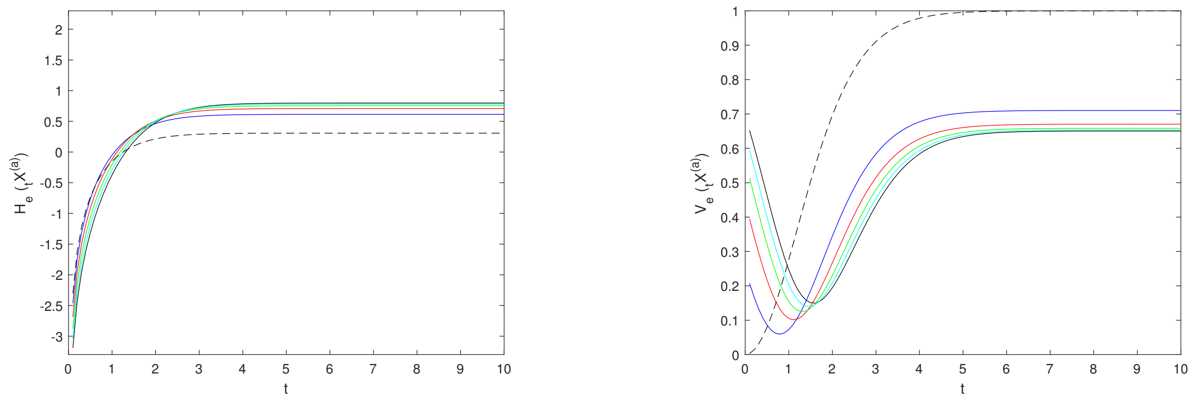
The plots of $H_e\left({}_tX^{(a)}\right)$ and $V_e\left({}_tX^{(a)}\right)$, for different integer values of a and with $\lambda = 2$, are shown in Figure 6. As for the modified Pareto case, both the past entropies and the past varentropies intersect each other for different values of a, having a similar behavior.

It must be observed that this behavior differs from what is shown in Example 4.1 in Di Crescenzo and Paolillo (2020), where the residual varentropies for a proportional hazard model with an underlying generalized exponential distribution do not intersect for different values of the parameter a.

This behavior seems to be confirmed by other similar analysis we performed. But there exists a family for which the past varentropies do not intersect, which is the family discussed in Proposition 2.3, whose varentropies are constant. In fact, let X_α be a lifetime having support $\mathcal{S} = [0, 1]$ and distribution $F_\alpha(x) = x^\alpha$, for $x \in \mathcal{S}$. Then, the corresponding parallel system with n i.i.d. components has distribution $F_{n\alpha}(x) = x^{n\alpha}$ for $x \in \mathcal{S}$, which is still in the family of distribution having constant past varentropy. Thus, in particular, one has $V_e({}_tX_{n\alpha}) = (1 - 1/(n\alpha))^2$ for all $t \in \mathcal{S}$, and obviously these past varentropies do not intersect as n varies in $\mathbb{N}^+$. This is another interesting property of such a family of distributions.

Conclusions

In this paper, we have introduced and studied the past varentropy. It is related to the past entropy, which is a measure of information about the past lifetimes. In particular, the past varentropy



(a) $H_e({}_tX^{(a)})$ (b) $V_e({}_tX^{(a)})$

Figure 6: Plots of past entropy and the past varentropy of exponential PRHR model for $a = 1$ (dashed line) and $a = 2, 3, 4, 5, 6$ (blue, red, green, cyan and black, respectively).

provides the variability of the information content of past lifetimes. We have given some examples of past varentropy and obtained conditions ensuring that it is monotone or constant. Moreover, its behavior under linear or monotonic transformations has been studied. A relationship between past varentropy and residual varentropy has been also provided, and upper and lower bounds for the past varentropy have been described. Finally, the behavior of the past varentropy under construction of parallel systems of i.i.d. components (and, more generally, under proportional reversed hazard rate models) has been discussed.

Acknowledgements

The authors would like to thank the reviewers for their constructive comments that greatly improved the paper.

During the preparation of the final revised version of this paper, the authors have been informed that some of their results can also be found in the recent paper

- Raqab, M. Z., Bayoud, H. A., & Qiu, G. (2022). Varentropy of inactivity time of a random variable and its related applications. IMA Journal of Mathematical Control and Information, 39(1), 132-154,

which has already appeared but was submitted after the submission to this journal of this work.

The authors are members of the research group GNAMPA of INdAM (Istituto Nazionale di Alta Matematica). Francesco Buono and Maria Longobardi are partially supported by MIUR - PRIN 2017, project "Stochastic Models for Complex Systems", no. 2017 JFFHSH. The present work was developed within the activities of the project 000009_ALTRI_CDA_75_2021_FRA_LINEA_B funded by "Programma per il finanziamento della ricerca di Ateneo - Linea B" of the University of Naples Federico II.

References

[1] Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N. (1993) *Statistical models based on counting processes*, Springer Verlag, New York.

[2] Arikan, E. (2016). varentropy decreases under polar transform. *IEEE Transactions on Information Theory*, **62**, 3390–3400.

[3] Asadi, M., Berred, A. (2012) Properties and estimation of the mean past lifetime. *Statistics*, 46, 405–417.

[4] Barlow, R. E., and F.J. Proschan (1996). *Mathematical Theory of Reliability.* Philadelphia: Society for Industrial and Applied Mathematics.

[5] Block, H. W., Savits, T. H. (1998). The Reversed Hazard Rate Function. *Probability in the Engineering and Informational Sciences* **12**: 69–90.

[6] Bobkov, S., Madiman, M. (2011). Concentration of the information in data with log-concave distributions.. *Annals of Probability*, **39**, 1528–1543.

[7] Buono, F., Longobardi, M., Szymkowiak, M. (2021). On Generalized Reversed Aging Intensity Functions. Ricerche mat, https://doi.org/10.1007/s11587-021-00560-w.

[8] Cacoullos, T., Papathanasiou, V., (1985). On upper bounds for the variance of functions of random variables. *Statistics & Probability Letters*, **3**, 175–184.

[9] Cacoullos, T., Papathanasiou, V. (1989). Characterizations of distributions by variance bounds. *Statistics & Probability Letters*, **7**(5), 351–356.

[10] Di Crescenzo, A., Longobardi, M. (2002). entropy-based measure of uncertainty in past lifetime distributions. *Journal of Applied Probability*, **39**, 434–440.

[11] Di Crescenzo, A., Paolillo, L. (2021). Analysis and applications of the residual varentropy of random lifetimes. *Probability in the Engineering and Informational Sciences*, **35**(3), 680–698.

[12] Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. *Sankhya: Series A*, **58**, 48–56.

[13] Finkelstein, M. S. (2002). On the Reversed Hazard Rate. *Reliability Engineering & System Safety* **78**: 71–75.

[14] Fradelizi, M., Madiman, M., Wang, L. (2016). Optimal Concentration of Information Content for Log-Concave Densities. *In: Houdré C., Mason D., Reynaud-Bouret P., Rosinski J. (eds) High Dimensional Probability VII. Progress in Probability*, **71**, 45–60.

[15] Gautschi, W., Gahill, W. F. (1972). Exponential Integral and Related Functions. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. (M. Abramowitz and I. A. Stegun, eds.). New York: Dover.

[16] Goodarzi, F., Amini, M., Mohtashami Borzadaran, G.R. (2016). On upper bounds for the variance of functions of the inactivity time. *Statistics & Probability Letters*, **117**, 62–71.

[17] Gray, R.M. (2011). *Entropy and information theory.* Springer, New York.

[18] Gupta, R. C., Gupta, R. D. (2007). Proportional reversed hazard rate model and its applications. *Journal of Statistical Planning and Inference* **137**: 3525–3536.

[19] Kandil, A. M., Kayid, M., Mahdy, M. (2011). Variance Inactivity Time Function and its Reliability Properties. *The 45th Annual Conference on Statistics, Computer Science and operations Research ISRR*, Cairo-Egypt, 94–113.

[20] Kayid, M., Izadkhah, S. (2014). Mean Inactivity Time Function, Associated Orderings, and Classes of Life Distributions. *IEEE Transactions on Reliability*, **63**(2), 593–602.

[21] Kontoyiannis, I., Verdú, S. (2014). Optimal lossless data compression: non-asymptotics and asymptotics. *IEEE Transactions on Information Theory* 60: 777–795.

[22] Krishna, H. and Kumar, K (2013). Reliability estimation in generalized inverted exponential distribution with progressively type II censored sample. *Journal of Statistical Computation and Simulation*, **83** (6), 1007–1019.

[23] Kundu, C., Nanda, A., Maiti, S. (2010). Some distributional results through past entropy. *Journal of Statistical Planning and Inference* **140**: 1280–1291.

[24] Li, X., Li, Z. (2008). A Mixture Model of Proportional Reversed Hazard Rate. *Communications in Statistics - Theory and Methods*, **37**(18), 2953–2963, DOI: 10.1080/03610920802050935.

[25] Lin, C.T., Duran, B.S. and Lewis, T.O. (1989). Inverted gamma as life distribution. *Microelectronics Reliability.* **29**(4), 619–626.

[26] Madiman, M., Wang, L. (2014). An optimal varentropy bound for log-concave distributions. *International Conference on Signal Processing and Communications* (SPCOM), Bangalore, doi: 10.1109/SPCOM.2014.6983953.

[27] Morris, R. (1979). The Dilogarithm Function of a Real Argument. *Mathematics of Computation*, **33**, 778–787.

[28] Muliere, P., Parmigiani, G., Polson, N. G. (1993). A Note on the Residual entropy Function. *Probability in the Engineering and Informational Sciences*, **7**, 413–420.

[29] Nanda, A.K., Chowdhury, S. (2019). Shannon's entropy and its generalizations towards statistics, reliability and information science during 1948-2018. arXiv:1901.09779v1.

[30] Nanda, A.K., Singh, H., Misra, N., Paul, P. (2003). Reliability properties of reversed residual lifetime. *Communications in Statistics - Theory and Methods* **32**, 2031–2042.

[31] Oguntunde P.E., Adejumo A.O. and Balogun, O.S. (2014). Statistical Properties of the Exponentiated Generalized Inverted Exponential Distribution. Applied Mathematics **4**(2), 47–55.

[32] Paolillo, L., Di Crescenzo, A., Suárez-Llorens, A. (2021). Stochastic Comparisons, Differential Entropy and Varentropy for Distributions Induced by Probability Density Functions. *arXiv:2103.11038v1.*

[33] Shaked, M., Shantikumar, J.G. (2007). *Stochastic Orders*, Springer.

[34] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423.

[35] Sun Han. T., Kobayashi, K. (2002). *Mathematics of Information and Coding.* American Mathematical Society. ISBN:978-0-8218-0534-3.