

Article

Interval Entropy and Informative Distance

Fakhroddin Misagh ^{1,*} and Gholamhossein Yari ²

¹ Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, 14778-93855, Iran

² School of Mathematics, Iran University of Science and Technology, Tehran, 16846-13114, Iran; E-Mail: yari@iust.ac.ir

* Author to whom correspondence should be addressed; E-Mail: misagh@iaut.ac.ir.

Received: 20 December 2011; in revised form: 4 February 2012 / Accepted: 7 February 2012 / Published: 2 March 2012

Abstract: The Shannon interval entropy function as a useful dynamic measure of uncertainty for two sided truncated random variables has been proposed in the literature of reliability. In this paper, we show that interval entropy can uniquely determine the distribution function. Furthermore, we propose a measure of discrepancy between two lifetime distributions at the interval of time in base of Kullback-Leibler discrimination information. We study various properties of this measure, including its connection with residual and past measures of discrepancy and interval entropy, and we obtain its upper and lower bounds.

Keywords: uncertainty; discrepancy; characterization

AMS Classification: 62N05, 62B10

1. Introduction

Recently, information theory has attracted the attention of statisticians. Sunoj *et al.* [1] have explored the use of information measures for doubly truncated random variables which plays a significant role in studying the various aspects of a system when it fails between two time points. In reliability theory and survival analysis, the residual entropy was considered in Ebrahimi and Pellerey [2], which basically measures the expected uncertainty contained in remaining lifetime of a system. The residual entropy has been used to measure the wear and tear of components and to

characterize, classify and order distributions of lifetimes by Belzunce *et al.* [3] and Ebrahimi [4]. The notion of past entropy, which can be viewed as the entropy of the inactivity time of a system, was introduced in Di Crescenzo and Longobardi [5].

Ebrahimi and Kirmani [6] introduced the residual discrimination measure and studied the minimum discrimination principle. Di Crescenzo and Longobardi [7] have considered the past discrepancy measure and presented a characterization of the proportional reversed hazards model. Furthermore, the use of information measures for doubly truncated random variables was explored by Misagh and Yari [8,9]. In this paper, continuing their work, we propose a new measure of discrepancy between two doubly truncated life distributions. The remaining of this paper is organized as follows: in Section 2, some results, including uniqueness of interval entropy and its invariance property are presented. Section 3 is devoted to definitions of dynamic measures of discrimination, including residual and past lifetimes and also the notion of interval discrimination measure is introduced. In Section 4 we present some results and properties of interval entropy and interval discrimination measures. Some conclusions are given in Section 4. Throughout this paper we consider absolutely continuous random variables.

2. Interval Entropy

Let X be a non-negative random variable describing a system failure time. We denote the probability density function of X as $f(x)$, the failure distribution as $F(x) = P(X \leq x)$ and the survival function as $\bar{F}(x) = P(X > x)$. The Shannon [10] information measure of uncertainty is defined as:

$$H(X) = -E(\log f(X)) = -\int_0^{\infty} f(x) \log f(x) dx \quad (1)$$

where $\log$ denotes the natural logarithm. Ebrahimi and Pellerey [2] considered the residual entropy of the non-negative random variable X at time t as:

$$H_X(t) = -\int_t^{\infty} \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)}{\bar{F}(t)} dx \quad (2)$$

Given that a system has survived up to time t , $H_X(t)$ essentially measures the uncertainty represented by the remaining lifetime. The residual entropy has been used to measure the wear and tear of systems and to characterize, classify and order distributions of lifetimes. See Belzunce *et al.* [3], Ebrahimi [4] and Ebrahimi and Kirmani [6]. Di Crescenzo and Longobardi [5] introduced the notion of past entropy and motivated its use in real-life situations. They also discussed its relationship with the residual entropy. Formally, the past entropy of X at time t is defined as follows:

$$\bar{H}_X(t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} dx \quad (3)$$

Given that the system X has failed at time t , $\bar{H}_X(t)$ measures the uncertainty regarding its past lifetime. Now Recall that the probability density function of $(X|t_1 < X < t_2)$ for all $0 < t_1 < t_2$ is given by $f(x)/(F(t_2) - F(t_1))$. Sunoj *et al.* [1] considered the notion of interval entropy of X in the interval (t_1, t_2) as the uncertainty contained in $(X|t_1 < X < t_2)$ which is denoted by:

$$IH(t_1, t_2) = -\int_{t_1}^{t_2} \frac{f(x)}{F(t_2) - F(t_1)} \log \frac{f(x)}{F(t_2) - F(t_1)} dx \quad (4)$$

We can rewrite the interval entropy as:

$$\begin{aligned} IH(t_1, t_2) = & 1 - \frac{1}{F(t_2) - F(t_1)} \int_{t_1}^{t_2} f(x) \log r(x) dx \\ & + \frac{1}{F(t_2) - F(t_1)} \{ \bar{F}(t_2) \log \bar{F}(t_2) - F(t_1) \log F(t_1) \\ & + [F(t_2) - F(t_1)] \log [F(t_2) - F(t_1)] \} \end{aligned}$$

where $r(x) = f(x)/\bar{F}(x)$ is the hazard function of X . Note that interval entropy can be negative and also it can be $-\infty$ or $+\infty$. Given that a system has survived up to time t_1 , and has been found to be down at time t_2 , $IH(t_1, t_2)$ measures the uncertainty about its lifetimes between t_1 and t_2 . Misagh and Yari [9] introduced a shift-dependent version of $IH(t_1, t_2)$. The entropy (4) has been used to characterize and ordering random lifetime distributions. See Misagh and Yari [8] and Sunoj *et al.* [1].

The general characterization problem is to obtain when the interval entropy uniquely determines the distribution function. The following proposition attempts to solve this problem. We first give definition of general failure rate (GFR) functions extracted from Navarro and Ruiz [11].

Definition 2.1. The GFRs of a random variable X having density function $f(x)$ and cumulative distribution function $F(x)$ are given by $h_1^X(t_1, t_2) = \frac{f(t_1)}{F(t_2) - F(t_1)}$ and $h_2^X(t_1, t_2) = \frac{f(t_2)}{F(t_2) - F(t_1)}$.

Remark 2.1. GFR functions determine distribution function uniquely. See Navarro and Ruiz [11].

Proposition 2.1. Let X be a non-negative random variable, and assume $IH(t_1, t_2)$ be increasing with respect to t_1 and decreasing with respect to t_2 , then $IH(t_1, t_2)$ uniquely determines $F(x)$.

Proof. By differentiating $IH(t_1, t_2)$ with respect to t_j , we have:

$$\frac{\partial IH_X(t_1, t_2)}{\partial t_1} = h_1(t_1, t_2) [IH_X(t_1, t_2) - 1 + \log h_1(t_1, t_2)]$$

and:

$$\frac{\partial IH_X(t_1, t_2)}{\partial t_2} = -h_2(t_1, t_2) [IH_X(t_1, t_2) - 1 + \log h_2(t_1, t_2)]$$

Thus, for fixed t_1 and arbitrary t_2 , $h_1(t_1, t_2)$ is a positive solution of the following equation:

$$g(x_{t_2}) = x_{t_2} [IH_X(t_1, t_2) - 1 + \log x_{t_2}] - \frac{\partial IH(t_1, t_2)}{\partial t_1} = 0 \quad (5)$$

Similarly, for fixed t_2 and arbitrary t_1 , we have $h_2(t_1, t_2)$ as a positive solution of the following equation:

$$\gamma(y_{t_1}) = y_{t_1} [IH_X(t_1, t_2) - 1 + \log y_{t_1}] + \frac{\partial IH(t_1, t_2)}{\partial t_2} = 0 \quad (6)$$

By differentiating g and γ with respect to x_{t_2} and y_{t_1} , we get $\frac{\partial g(x_{t_2})}{\partial x_{t_2}} = \log x_{t_2} + IH(t_1, t_2)$, and $\frac{\partial \gamma(y_{t_1})}{\partial y_{t_1}} = \log y_{t_1} + IH(t_1, t_2)$. Furthermore, second-order derivatives of g and γ with respect to x_{t_2}

and y_{t_1} are $\frac{1}{x_t} > 0$ and $\frac{1}{y_t} > 0$ respectively. Then the functions g and γ are minimized at points $x_{t_2} = e^{-IH(t_1, t_2)}$ and $y_{t_1} = e^{-IH(t_1, t_2)}$ respectively. In addition, $g(0) = -\frac{\partial IH(t_1, t_2)}{\partial t_1} < 0, g(\infty) = \infty$ and $\gamma(0) = -\frac{\partial IH(t_1, t_2)}{\partial t_1} < 0, \gamma(\infty) = \infty$. So, both functions g and γ first decrease and then increase with respect to x_{t_2} and y_{t_1} respectively. which conclude that Equations (5) and (6) has unique roots $h_1(t_1, t_2)$ and $h_2(t_1, t_2)$ respectively. Now, $IH(t_1, t_2)$ uniquely determines GFRs and by virtue of Remark 2.1, the distribution function.

The effect of monotone transformations on the residual and past entropy has been discussed in Ebrahimi and Kirmani [6] and Di Crescenzo and Longobardi [5] respectively. Following proposition gives similar results for interval entropy.

Proposition 2.2. Suppose X be a non-negative random variable with cumulative distribution function F and survival function $\bar{F}$; let $Y = \varphi(X)$, with φ , strictly increasing and differentiable function. Then for all $0 < t_1 < t_2 < \infty$:

$$\begin{aligned} IH_Y(t_1, t_2) &= IH_X(\varphi^{-1}(t_1), \varphi^{-1}(t_2)) \\ &+ \frac{1}{F(\varphi^{-1}(t_2)) - F(\varphi^{-1}(t_1))} \{ E(\log \varphi'(X)) \\ &- F(\varphi^{-1}(t_1)) E(\log \varphi'(X) | X < \varphi^{-1}(t_1)) \\ &- \bar{F}(\varphi^{-1}(t_2)) E(\log \varphi'(X) | X > \varphi^{-1}(t_2)) \} \end{aligned}$$

Proof. Recalling (1), The Shannon information of X and Y can be expressed as:

$$H(Y) = H(X) + E(\log \varphi'(X)) \quad (7)$$

From Theorem 2 of Ebrahimi and Kirmani [6] and Proposition 2.4 of Di Crescenzo and Longobardi [5], we have:

$$H_Y(t_1) = H_X(\varphi^{-1}(t_2)) + E(\log \varphi'(X) | X > \varphi^{-1}(t_1)) \quad (8)$$

and:

$$\bar{H}_Y(t_1) = \bar{H}_X(\varphi^{-1}(t_1)) + E(\log \varphi'(X) | X < \varphi^{-1}(t_1)) \quad (9)$$

Due to Equation 2.8 of Sunoj *et al.* [1], there holds:

$$\begin{aligned} H(Y) &= \mathcal{H}(G(t_1), \bar{G}(t_2), 1 - G(t_1) - \bar{G}(t_2)) \\ &+ G(t_1)\bar{H}_Y(t_1) + \bar{G}(t_2)H_Y(t_2) \\ &+ [1 - G(t_1) - \bar{G}(t_2)] IH_Y(t_1, t_2) \end{aligned} \quad (10)$$

where G and $\bar{G}$ denote distribution and survival functions of Y respectively. Substituting $H(Y)$, $\bar{H}_Y(t_1)$ and $H_Y(t_1)$ in (7), (8) and (9) into terms of (10), we get:

$$\begin{aligned} H(X) + E(\log \varphi'(X)) &= (F(\varphi^{-1}(t_2)) - F(\varphi^{-1}(t_1))) IH_Y(t_1, t_2) \\ &- F(\varphi^{-1}(t_1)) E(\log \varphi'(X) | X < \varphi^{-1}(t_1)) \end{aligned}$$

$$\begin{aligned}
& -\bar{F}(\varphi^{-1}(t_2))E\left(\log \varphi'(X) \mid X > \varphi^{-1}(t_2)\right) \\
& +F(\varphi^{-1}(t_1))\bar{H}_X(\varphi^{-1}(t_1)) \\
& +\bar{F}(\varphi^{-1}(t_2))H_X(\varphi^{-1}(t_2)) \\
& +\mathcal{H}(F(\varphi^{-1}(t_1)), \bar{F}(\varphi^{-1}(t_2)), F(\varphi^{-1}(t_2)) - F(\varphi^{-1}(t_1)))
\end{aligned} \tag{11}$$

Three terms of the right hand side of (11) are equal to:

$$[F(\varphi^{-1}(t_2)) - F(\varphi^{-1}(t_1))]IH_X(\varphi^{-1}(t_1), \varphi^{-1}(t_2))$$

and the proof is complete.

Remark 2.2. Suppose $\varphi(X) = F(X)$, then the function φ satisfies the assumptions of Proposition 2.2 and uniformly distributed over $(0,1)$, then:

$$\begin{aligned}
& IH_{F(X)}(t_1, t_2) = IH_X(F^{-1}(t_1), F^{-1}(t_2)) \\
& -\frac{1}{t_2-t_1} \{ H(X) + t_1 E(\log f(X) \mid X < F^{-1}(t_1)) + t_2 E(\log f(X) \mid X > F^{-1}(t_2)) \}
\end{aligned}$$

Remark 2.3. For all $0 < \theta < t_1$, we get $IH_{X+\theta}(t_1, t_2) = IH_X(t_1 - \theta, t_2 - \theta)$.

Remark 2.4. Let $Y = aX$ where $a > 0$, then, we have $IH_{aX}(t_1, t_2) = IH_X\left(\frac{t_1}{a}, \frac{t_2}{a}\right) + \log a$.

3. Informative Distance

In this section, we review some basic definitions and facts for measures of discrimination between two residual and past lifetime distributions. We introduce a measure of discrepancy between two random variables at an interval of time.

Let X and Y are two non-negative random variables describing times to failure of two systems. We denote the probability density functions of X and Y as $f(x)$ and $g(y)$, failure distributions as $F(x) = P(X \leq x)$ and $G(y) = G(Y \leq y)$ and the survival functions as $\bar{F}(x) = P(X > x)$ and $\bar{G}(y) = G(Y > y)$ respectively, with $F(0) = G(0) = 1$. Kullback-Leibler [12] informative distance between F and G is defined by:

$$I_{X,Y} = \int_0^\infty f(x) \log \frac{f(x)}{g(x)} dx \tag{12}$$

where $\log$ denotes natural logarithm. $I_{X,Y}$ is known as relative entropy and it is shift and scale invariant. However it is not metric, since symmetrization and triangle inequality does not hold. We point out the Jensen-Shannon divergence (JSD) which is based on the Kullback-Leibler divergence, with the notable differences that it is always a finite value and its square root is a metric. See Nielsen [13] and Amari *et al.* [14]. The application of $I_{X,Y}$ as an informative distance in residual and past lifetimes has increasingly studied in recent years. In particular, Ebrahimi and Kirmani [6] considered the residual Kullback-Leibler discrimination information of non-negative lifetimes of the systems X and Y at time t as:

$$I_{X,Y}(t) = \int_t^\infty \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)/\bar{F}(t)}{g(x)/\bar{G}(t)} dx \tag{13}$$

Given that both systems have survived up to time t , $I_{X,Y}(t)$ identifies with the relative entropy of remaining lifetimes ($X|X > t$) and ($Y|Y > t$). Furthermore, the Kullback-Leibler distance for two past lifetimes was studied in Di Crescenzo and Longobardi [7] which is dual to (13) in the sense that it is an informative distance between past lifetimes ($X|X < t$) and ($Y|Y < t$). Formally, the past Kullback-Leibler distance of non-negative random lifetimes of the systems X and Y at time t is defined as:

$$\bar{I}_{X,Y}(t) = \int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)/F(t)}{g(x)/G(t)} dx \quad (14)$$

Given that at time t , both systems have been found to be down, $\bar{I}_{X,Y}(t)$ measures the informative distance between their past lives.

Along a similar line, we define a new discrepancy measure that completes studying informative distance between two random lifetimes.

Definition 3.1. The interval distance between random lifetimes X and Y at interval (t_1, t_2) is the Kullback-Leibler discrimination measure between the truncated lives ($X|t_1 < X < t_2$) and ($Y|t_1 < Y < t_2$):

$$ID_{X,Y}(t_1, t_2) = \int_{t_1}^{t_2} \frac{f(x)}{F(t_2)-F(t_1)} \log \frac{f(x)/[F(t_2)-F(t_1)]}{g(x)/[G(t_2)-G(t_1)]} dx \quad (15)$$

Remark 3.1. Clearly $ID_{X,Y}(0, t) = \bar{I}_{X,Y}(t)$, $ID_{X,Y}(t, \infty) = I_{X,Y}(t)$ and $ID_{X,Y}(0, \infty) = I_{X,Y}$.

Given that both systems X and Y have survived up to time t_1 and have seen to be down at time t_2 , $ID_{X,Y}(t_1, t_2)$ measures the discrepancy between their unknown failure times in the interval (t_1, t_2) . $ID_{X,Y}(t_1, t_2)$ satisfies all properties of Kullback-Leibler discrimination measure and can be rewritten as:

$$ID_{X,Y}(t_1, t_2) = - \int_{t_1}^{t_2} \frac{f(x)}{F(t_2)-F(t_1)} \log \frac{g(x)}{G(t_2)-G(t_1)} dx - IH_X(t_1, t_2) \quad (16)$$

where $IH_X(t_1, t_2)$ is the interval entropy of X in (4).

An alternative way of writing (16) is the following:

$$ID_{X,Y}(t_1, t_2) = \log \frac{G(t_2)-G(t_1)}{F(t_2)-F(t_1)} + \frac{1}{F(t_2)-F(t_1)} \int_{t_1}^{t_2} f(x) \log \frac{f(x)}{g(x)} dx \quad (17)$$

The following example clarifies the effectiveness of the interval discrimination measure.

Example 3.1. Suppose X and Y be random lifetimes of two systems with joint density function:

$$f(x, y) = \frac{1}{4}, 0 < x < 2, 0 < y < 4 - 2x$$

and that the marginal densities of X and Y are $f(x) = \frac{1}{2}(2-x)$, $0 < x < 2$ and $g(y) = \frac{1}{8}(4-y)$, $0 < y < 4$, respectively. Because X and Y , belongs to different domains, using relative entropy to measure the informative distance between X and Y is not interpretable. The interval distance between X and Y in the intervals $(0, 1.5)$ and $(1.5, 2)$ are 0.01 and 0.16 respectively. Hence, the informative distance between X and Y in the interval $(1.5, 2)$ is greater than of it in the interval $(0, 1.5)$.

In the following proposition we decompose the Kullback-Leibler discrimination function in terms of residual, past and interval discrepancy measure. The proof is straightforward.

Proposition 3.1. Let X and Y are two non-negative random lifetimes of two systems. For all $0 \leq t_1 < t_2 < \infty$, the Kullback-Leibler discrimination measure is decomposed as follows:

$$I_{X,Y} = [F(t_2) - F(t_1)]ID_{X,Y}(t_1, t_2) + F(t_1)\bar{I}_{X,Y}(t_1) + \bar{F}(t_2)I_{X,Y}(t_2) + I_{U,V}(t_1, t_2) \quad (18)$$

where:

$$I_{U,V}(t_1, t_2) = F(t_1) \log \frac{F(t_1)}{G(t_1)} + \bar{F}(t_2) \log \frac{\bar{F}(t_2)}{\bar{G}(t_2)} + [F(t_2) - F(t_1)] \log \frac{F(t_2) - F(t_1)}{G(t_2) - G(t_1)}$$

is the Kullback-Leibler distance between two trivalent discrete random variables.

Proposition 3.1 admits the following interpretation: the Kullback-Leibler discrepancy measure between random lifetimes of systems X and Y is composed from four parts: (i) the discrepancy between the past lives of two systems at time t_1 ; (ii) the discrepancy between residual lifetimes of X and Y that have both survived up to time t_2 ; (iii) the discrepancy between the lifetimes of both systems in the interval (t_1, t_2) ; (iv) the discrepancy between two random variables which determines if the systems have been found to be failing before t_1 , between t_1 and t_2 or after t_2 .

4. Some Results on Interval Based Measures

In this section we study the properties of $ID(t_1, t_2)$ and point out certain similarities with those of $I_{X,Y}(t)$ and $\bar{I}_{X,Y}(t)$. The following proposition gives lower and upper bounds for the interval distance. We first give definition of likelihood ratio ordering.

Definition 4.1. X is said to be larger than Y in likelihood ratio ($X \geq^{LR} Y$) if $\frac{f(x)}{g(x)}$ is increasing in x over the union of the supports of X and Y .

Several results regarding the ordering in Definition 4.1. was provided in Ebrahimi and Pellerey [2].

Proposition 4.1. Let X and Y are random variables with common support $(0, \infty)$. Then:

(i) $X \geq^{LR} Y$ implies:

$$\log \frac{h_1^X(t_1, t_2)}{h_1^Y(t_1, t_2)} \leq ID_{X,Y}(t_1, t_2) \leq \log \frac{h_2^X(t_1, t_2)}{h_2^Y(t_1, t_2)} \quad (19)$$

when $\frac{f(x)}{g(x)}$ is decreasing in $x > 0$, then the inequalities in (19) are reversed.

(ii) Decreasing $g(x)$ in $x > 0$, implies:

$$\log \frac{1}{h_1^Y(t_1, t_2)} \leq ID_{X,Y}(t_1, t_2) + IH_X(t_1, t_2) \leq \log \frac{1}{h_2^Y(t_1, t_2)} \quad (20)$$

for increasing $g(x)$ then the inequalities in (20) are reversed.

Proof. Because of increasing $\frac{f(x)}{g(x)}$ in $x > 0$, from (15), we have:

$$ID_{X,Y}(t_1, t_2) \leq \int_{t_1}^{t_2} \frac{f(x)}{F(t_2) - F(t_1)} \log \frac{f(t_2)/[F(t_2) - F(t_1)]}{g(t_2)/[G(t_2) - G(t_1)]} dx = \log \frac{h_2^X(t_1, t_2)}{h_2^Y(t_1, t_2)}$$

and:

$$ID_{X,Y}(t_1, t_2) \geq \int_{t_1}^{t_2} \frac{f(x)}{F(t_2) - F(t_1)} \log \frac{f(t_1)/[F(t_2) - F(t_1)]}{g(t_1)/[G(t_2) - G(t_1)]} dx = \log \frac{h_1^X(t_1, t_2)}{h_1^Y(t_1, t_2)}$$

which gives (19). When $\frac{f(x)}{g(x)}$ is decreasing, the proof is similar. Furthermore, for all $t_1 < x < t_2$ decreasing $g(x)$ in $x > 0$ implies $g(t_2) < g(x) < g(t_1)$, then from (16) we get:

$$ID_{X,Y}(t_1, t_2) \leq -\log h_2^Y(t_1, t_2) - IH_X(t_1, t_2)$$

and:

$$ID_{X,Y}(t_1, t_2) \geq -\log h_1^Y(t_1, t_2) - IH_X(t_1, t_2)$$

so that (20) holds. When $g(x)$ is increasing the proof is similar.

Remark 4.1. Consider X and Y are two non-negative random variables corresponding to weighted exponential distributions with positive rates λ and μ respectively and with common positive real weight function $\omega(\cdot)$. The densities of X and Y are $f(x) = \frac{\omega(x)e^{-\lambda x}}{h(\lambda)}$, and $g(x) = \frac{\omega(x)e^{-\mu x}}{h(\mu)}$ respectively, where $h(\cdot)$ denotes the Laplace transform of $\omega(\cdot)$ given by $h(\theta) = \int_0^\infty \omega(x)e^{-\theta x} dx$, $\theta > 0$, therefore, for $\lambda \neq \mu$ the interval distance between X and Y at interval (t_1, t_2) is the following:

$$ID_{X,Y}(t_1, t_2) = \log \frac{G(t_2) - G(t_1)}{F(t_2) - F(t_1)} + \log \frac{h(\mu)}{h(\lambda)} - (\lambda - \mu)E(X|t_1 < X < t_2) \quad (21)$$

Remark 4.2. Let X be a non-negative random lifetime with density function $f(x)$ and cumulative distribution function $F(t) = P(X \leq t)$. Then the density function and cumulative distribution function for the weighted random variable X_ω associated to a positive real function $\omega(\cdot)$ are $f_\omega(x) = \frac{\omega(x)}{E(\omega(X))}f(x)$, and $F_\omega(t) = \frac{E(\omega(X)|X \leq t)}{E(\omega(X))}F(t)$, respectively, where $E(\omega(X)) = \int_0^\infty \omega(x)f(x)dx$.

Then, from (17) we have:

$$ID_{X,X_\omega}(t_1, t_2) = \log E(\omega(X)|t_1 < X < t_2) - E(\log(\omega(X))|t_1 < X < t_2) \quad (22)$$

A similar expression is available in Maya and Sunoj [15] for past life time. Due to (22) and from non-negativity of $ID_{X,X_\omega}(t_1, t_2)$ we have:

$$\log E(\omega(X)|t_1 < X < t_2) \geq E(\log(\omega(X))|t_1 < X < t_2)$$

which is a direct result of Markov inequality for concave functions.

Example 4.1. For $\omega(x) = x^{n-1}$ and $h(\theta) = (n-1)!/\theta^n$ the distributions of random variables in Remark 4.1 called Erlang distributions with scale parameters λ and μ and with common shape parameter. The conditional mean of $(X|t_1 < X < t_2)$ is the following:

$$E(X|t_1 < X < t_2) = \frac{1}{F(t_2) - F(t_1)} \int_{t_1}^{t_2} x \frac{x^{n-1}\lambda^n e^{-\lambda x}}{(n-1)!} dx = \frac{\gamma(n+1, \lambda t_2) - \gamma(n+1, \lambda t_1)}{\lambda(n-1)!(F(t_2) - F(t_1))}$$

where $\gamma(\alpha, x) = \int_0^x e^{-u}u^{\alpha-1}du$ is the incomplete Gamma function. From (21) we obtain:

$$ID_{X,Y}(t_1, t_2) = \log \frac{\gamma(n, \mu t_2) - \gamma(n, \mu t_1)}{\gamma(n, \lambda t_2) - \gamma(n, \lambda t_1)} - n \log \frac{\lambda}{\mu} + (\lambda - \mu) \frac{\gamma(n+1, \lambda t_2) - \gamma(n+1, \lambda t_1)}{\lambda(n-1)!(F(t_2) - F(t_1))}$$

In the following proposition, sufficient condition for $ID_{X_1,Y}(t_1, t_2)$ to be smaller than $ID_{X_2,Y}(t_1, t_2)$ is provided.

Proposition 4.2. Consider three non-negative random variables X_1 , X_2 and Y with probability density functions f_1 , f_2 and g respectively. $X_1 \stackrel{LR}{\geq} Y$ implies $ID_{X_1,Y}(t_1, t_2) \leq ID_{X_2,Y}(t_1, t_2)$.

Proof. From (17) we have:

$$\begin{aligned} ID_{X_1,Y}(t_1, t_2) - ID_{X_2,Y}(t_1, t_2) \\ &= -ID_{X_2,X_1}(t_1, t_2) + \int_{t_1}^{t_2} \left(\frac{f_1(x)}{F_1(t_2) - F_1(t_1)} - \frac{f_2(x)}{F_2(t_2) - F_2(t_1)} \right) \log \frac{f_1(x)}{g(x)} dx \\ &\leq \int_{t_1}^{t_2} \left(\frac{f_1(x)}{F_1(t_2) - F_1(t_1)} - \frac{f_2(x)}{F_2(t_2) - F_2(t_1)} \right) \log \frac{f_1(x)}{g(x)} dx \\ &\leq \log \frac{f_1(t_2)}{g(t_2)} \int_{t_1}^{t_2} \left(\frac{f_1(x)}{F_1(t_2) - F_1(t_1)} - \frac{f_2(x)}{F_2(t_2) - F_2(t_1)} \right) dx = 0 \end{aligned}$$

where the first inequality comes from the fact that $ID_{X_2,X_1}(t_1, t_2) \geq 0$ and the second one follows from the increasing $\frac{f_1(x)}{g(x)}$ in $x > 0$.

Example 4.2. Let $\{N(t), t \geq 0\}$ be a non-homogeneous Poisson process with a differentiable mean function $M(t) = E(N(t))$ such that $M(t)$ tends to ∞ as t tends to ∞ . Let $R_n, n = 1, 2, 3$ denote the time of the occurrence of the n -th event in such a process. Then R_n has the following density function:

$$f_n(x) = \frac{(M(x))^{n-1}}{(n-1)!} f_1(x), x > 0$$

where:

$$f_1(x) = -\frac{d}{dx} \exp(-M(x)), x > 0$$

clearly $f_n(x)/f_1(x)$ is increasing in $x > 0$. It follows from Proposition 4.2 that for all $m \leq n$:

$$ID_{X_n,X_1}(t_1, t_2) \leq ID_{X_m,X_1}(t_1, t_2)$$

Proposition 4.3. Let X and Y were random variables with common support $(0, \infty)$. Let φ be a continuous and increasing function, then:

$$ID_{\varphi(X),\varphi(Y)}(t_1, t_2) = ID_{X,Y}(\varphi^{-1}(t_1), \varphi^{-1}(t_2))$$

Proof. The proof is straightforward.

The following remarks clarify the invariance of interval discrimination measure under location and scale transformation.

Remark 4.3. For all $0 \leq \theta < t_1$, we get $ID_{X+\theta,Y+\theta}(t_1, t_2) = ID_{X,Y}(t_1 - \theta, t_2 - \theta)$.

Remark 4.4. Let $Y = aX$ where $a > 0$, $ID_{aX,aY}(t_1, t_2) = ID_{X,Y}\left(\frac{t_1}{a}, \frac{t_2}{a}\right)$.

4. Conclusions

In this paper, we presented two novel measures of information which are based on a time interval and are more general than the well-known Shannon's differential entropy and Kullback-Leibler divergence measure. These new measures are consistent in that they are valid in both past and residual lifetimes. We call these measures of information the interval entropy and the informative distance. We obtain the requirements that interval entropy can uniquely determine the distribution function. We presented several propositions and remarks, some of which parallel those for Shannon entropy and Kullback-Leibler divergence and others that are more general. The advantages of interval entropy and informative distance were outlined as well. We believe that interval basic measures will have many applications in reliability, stochastic processes and other areas in the near future. The results presented here are by no means comprehensive but hopefully will pave the way for studying the entropy in a different and more general setting.

Acknowledgment

We are very grateful to the anonymous referees for important and constructive comments, which improved the presentation significantly. We also thank Assistant Editor Sarah Shao.

References and Notes

1. Sunoj, S.M.; Sankaran, P.G.; Maya, S.S. Characterizations of life distributions using conditional expectations of doubly (Interval) truncated random variables. *Comm. Stat. Theor. Meth.* **2009**, *38*, 1441–1452.
2. Ebrahimi, N.; Pellerey, F. New partial ordering of survival functions based on the notion of uncertainty. *J. Appl. Prob.* **1995**, *32*, 202–211.
3. Belzunce, F.; Navarro, J.; Ruiz, J.M.; Yolanda, A. Some results on residual entropy function. *Metrika* **2004**, *59*, 147–161.
4. Ebrahimi, N. Testing whether lifetime distribution is increasing uncertainty. *J. Statist. Plann. Infer.* **1997**, *64*, 9–19.
5. Di Crescenzo, A.; Longobardi, M. Entropy based measure of uncertainty in past lifetime distributions. *J. Appl. Prob.* **2002**, *39*, 434–440.
6. Ebrahimi, N.; Kirmani, S.N.U.A. Some results on ordering of survival functions through uncertainty. *Statist. Prob. Lett.* **1996**, *29*, 167–176.
7. Di Crescenzo, A.; Longobardi, M. A measure of discrimination between past lifetime distributions. *Stat. Prob. Lett.* **2004**, *67*, 173–182.
8. Misagh, F.; Yari, G.H. A novel entropy-based measure of uncertainty to lifetime distributions characterizations. In Proceedings of the First International Conference on Mathematics and Statistics, AUS-ICMS'10, American University of Sharjah, Sharjah, UAE, 18–21 March 2010; ISBN 978-9948-427-17-9, ID: 100196.
9. Misagh, F.; Yari, G.H. On weighted interval entropy. *Statist. Prob. Lett.* **2011**, *29*, 167–176.
10. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.

11. Navarro, J.; Ruiz, J.M. Failure rate functions for doubly truncated random variables. *IEEE Trans. Reliab.* **1996**, *45*, 685–690.
12. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Statist.* **1951**, *25*, 745–751.
13. Nielsen, F. A family of statistical symmetric divergences based on Jensen’s inequality. *arXiv* **2011**, arXiv:1009.4004v2[cs.CV].
14. Amari, S.-I.; Barndorff-Nielsen, O.E.; Kass, R.E.; Lauritzen, S.L.; Rao, C.R. Differential geometry in statistical inference. *IMS Lectures Notes* **1987**, *10*, 217.
15. Maya, S.S.; Sunoj, S.M. Some dynamic generalized information measures in the context of weighted models. *Statistica Anno.* **2008**, *LXVIII 1*, 71–84.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).