JIRSS (2017)
Vol. 17, No. 2, pp 21-32
DOI:10.22034/jirss.2017.16.02

Some Results on Weighted Cumulative Entropy

Maliheh Mirali [1] **and Simindokht Baratpour** [2]

[1]Department of Statistics, International Campus, Ferdowsi University of Mashhad, Mashhad, Iran.
[2]Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran.

Received: 14/09/2016, Revision received: 01/04/2017, Published online: 27/11/2017

Abstract. Considering Rao *et al.* (2004) and Di Crescenzo and Longobardi (2009) studies, Misagh *et al.* (2011) proposed a weighted information which is based on the cumulative entropy called *Weighted Cumulative Entropy* (WCE). The above-mentioned model is a *Shift-dependent Uncertainty Measure*. In this paper, we examine some of the properties of WCE and obtain some bounds for that. In order to estimate this information measure, we put forward empirical WCE. Furthermore, in some theorems, we have some characterization results. We explore that, if the WCE of the first (last) order statistic are equal for two distributions, then this two distributions will be equal.

Keywords. Cumulative entropy, Cumulative residual entropy, Survival function, Weighted Shannon entropy.

MSC: 94A17; 62N05, 60E15.

1 Introduction

Information coding and transmission is a prominent part in description of the behaviour of biological and engineering systems. Entropy, which was introduced by Shannon (1948) and Wiener (1948, 2nd Ed. 1961), plays an undeniable essential role in the field of information

Corresponding Author: S. Baratpour (baratpour@um.ac.ir)
M. Mirali (malihemirali@yahoo.com).

theory. As an example, it can be used in dealing with information in the context of theoretical neurobiology (see for instance, Johnson and Glantz (2004)) .

Let X be a non-negative random variable with a continuous cumulative distribution function (cdf) $F_X(x)$ and the probability density function (pdf) $f_X(x)$. Then, the differential entropy, commonly known as Shannon entropy, is defined as

$$H(X) = -E[\log f_X(X)] = -\int_0^{\infty} f_X(x)\log f_X(x)dx,$$

where log means natural logarithm. Note that $H(X)$ is a location free *shift-independent* measure.

Di Crescenzo and Longobardi (2006) have proposed the weighted entropy, which is another measure of uncertainty as a suitable generalization or modification of the classical entropy, which is a *Length biased* shift-dependent information measure, defined as

$$H^w(X) = -E[X\log f_X(X)] = -\int_0^{+\infty} x f_X(x)\log f_X(x)dx.$$

Despite the merits and vast application of Shannon's entropy in many fields of research, some limitations in measuring the randomness of certain systems lead to proposing a new measure of uncertainty, called cumulative residual entropy (CRE), by Rao *et al.* (2004), which is defined as

$$\mathcal{E}(X) = -\int_0^{\infty} \bar{F}_X(x)\log \bar{F}_X(x)dx,$$

where $\bar{F}_X(x) = 1 - F_X(x)$ is the survival function of X.

Navarro *et al.* (2010) studied several features of the CRE. Asadi and Zohrevand (2007) also considered a dynamic version of the CRE. Moreover, Psarrakos and Navarro (2013) have considered a dynamic generalized cumulative residual entropy (GCRE). Another study on this topic is Kapodistria and Psarrakos (2012) in which some new connections of the CRE and the residual lifetime were introduced. Also a cumulative version of Renyi's entropy was studied in Sunoj and Linu (2012).

Misagh *et al.* (2011) proposed a weighted information which is based on the cumulative residual entropy, called weighted cumulative residual entropy (WCRE). This is a *Length biased* Shift-dependent information measure that assigns larger weights to larger values of a random variable. As pointed out by Misagh *et al.* (2011), in some practical situations of reliability and neurobiology, a shift-dependent measure of uncertainty is desirable. Also, an important feature of the human visual system is that it can recognize objects in a scale- and translation-invariant manner. However, achieving this desirable behavior using biologically realistic network is a challenge. The notion of weighted entropy addresses this requirement. Mirali *et al.* (2015) investigated some properties of WCRE, conditional and empirical WCRE. This measure is defined as

$$\mathcal{E}^w(X) = -\int_0^{\infty} x\bar{F}_X(x)\log \bar{F}_X(x)dx. \tag{1.1}$$

Mirali and Baratpour (2015) proposed a dynamic form of WCRE and studied some properties of this measure. Also, some characterization results were provided. A new information measure similar to CRE has been proposed by Di Crescenzo and Longobardi (2009). It is named cumulative entropy and is defined as

$$\mathcal{CE}(X) = -\int_0^{+\infty} F_X(x) \log F_X(x) dx.$$

They discussed its main properties, including a connection to the reliability theory, and estimated it based on the empirical cumulative entropy.

In this paper, some characterization results based on order statistics are obtained. Let $X_1, X_2, \cdots, X_n$ be iid observations from (cdf) $F_X(x)$ and pdf $f_X(x)$. The order statistics of this sample are denoted as $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ (See Arnold et al. (1992)).

Rest of the paper is organized as follows. In Section 2, the definition of the weighted cumulative entropy (WCE) and a description of its properties are given in the form of several theorems. An estimate of this measure together with the central limit theorem are also provided. In Section 3, we present WCRE and WCE of the first and the last order statistics and conclude with some characterization results.

2 Weighted Cumulative Entropy

Misagh *et al.* (2011) introduced weighted cumulative entropy (WCE) which is based on the cumulative entropy. In this section, we investigate some properties of this measure. The connection between the cumulative entropy and the reliability theory will be also considered.

Definition 2.1. Let X be a non-negative continuous random variable with cdf $F_X(x)$. The WCE of X is defined by

$$\mathcal{CE}^{\omega}(X) = -\int_0^{\infty} x F_X(x) \log F_X(x) dx. \tag{2.1}$$

Note that the WCRE and WCE have some advantages with respect to the weighted entropy. Weighted entropy is based on the density function which, in general, may or may not exist. Weighted entropy may be negative, but WCRE and WCE are always non-negative. Despite WCRE and WCE, the weighted entropy cannot be estimated by the empirical distribution function.

WCE has some properties similar to the WCRE which have been introduced by Mirali *et al.* (2015) Therefore, their proofs are omitted.

- If $Y = aX + b$, $a > 0$, $b \geq 0$, then

$$\mathcal{CE}^{\omega}(Y) = a^2 \mathcal{CE}^{\omega}(X) + ab\mathcal{CE}^{\omega}(X).$$

- If $E(X^2) < \infty$, then $\mathcal{CE}^{\omega}(X) < \infty$.

- If $H(X)$ is the Shannon entropy of X, then

$$\mathcal{CE}^{w}(X) \geq C \exp[E\Big(\log(X)\Big)] \cdot \exp H(X), \tag{2.2}$$

where $C = \exp \int_0^1 \log(x|\log x|)dx = 0.2065$.

- $\mathcal{CE}^{\omega}(X) = 0$, if and only if X is degenerate.
- $\mathcal{CE}^{w}(X) \geq \int_0^{\infty} x F_X(x)\bar{F}_X(x)dx$.
- If $\mathcal{CE}^{w}(X) < \infty$, then $\mathcal{CE}^{w}(X) = E[T(X)]$, where

$$T(t) = \int_t^{\infty} x \int_x^{\infty} r_X(u)dudx,$$

and $r_X(t) = \frac{f_X(t)}{F_X(t)}$ is the reversed hazard rate function.

- Let X be a continuous random variable with cdf $F_X(x)$ that takes values in $[0, b]$ with finite b. Then, $\mathcal{CE}^{\omega}(X) \leq b\mathcal{CE}(X)$ and $\mathcal{CE}^{\omega}(X) \leq -\frac{1}{2}\Big(b^2 - E(X^2)\Big)\log\Big(1 - \frac{E(X^2)}{b^2}\Big)$.
- Let X be a continuous random variable with cdf $F_X(x)$ that takes values in $[a, \infty)$ with finite a>0. Then, $\mathcal{CE}^{\omega}(X) \geq a\,\mathcal{CE}(X)$.
- Let X and Y be continuous random variables with cdfs $F_X(x)$ and $G_X(x)$, respectively, that take values in $[a, +\infty)$, $a > 0$. Then, we have

$$\mathcal{CE}^{w}(X) \geq \frac{a}{2}E(X - Y).$$

- For any continuous, non-negative and independent random variables X and Y, we have

$$\mathcal{CE}^{\omega}(X + Y) \geq \max\Big\{\mathcal{CE}^{\omega}(X) + E(X)\mathcal{CE}(X), \mathcal{CE}^{\omega}(Y) + E(Y)\mathcal{CE}(Y)\Big\}.$$

2.1 Connection to the Reliability Theory

In this part, the connection between the cumulative entropy and the reliability theory will be considered. If X is the lifetime of a system, then the inactivity time of the system is denoted by $[t - X\Big|X \leq t],\ t \geq 0$. The inactivity time is thus the duration of the time occurring between an inspection time t and the failure time X, given that at time t the system has been found to be down. For all $t \geq 0$, such that $F_X(t) > 0$, the mean inactivity time is given by

$$\tilde{\mu}(t) = E[t - X\Big|X \leq t] = \frac{1}{F_X(t)} \int_0^t F_X(x)dx.$$

This function has been used in various contexts of the reliability theory involving stochastic orders and characterization of random lifetime (see for instance, Ahmad and Kayid (2005), Ahmad et al. (2005), Kayid and Ahmad (2004), Li and Lu (2003), Misra *et al.* (2008) and Nanda *et al.* (2003)). We indicate that WCE has a direct relation to the second moment of the inactivity time: for all $t \geq 0$ and $F_X(t) > 0$, it is defined as

$$\eta_2(t) = E\Big((t-X)^2 \Big| X \leq t\Big). \tag{2.3}$$

It can easily be shown that

$$\begin{aligned}\eta_2(t) &= \frac{2}{F_X(t)} \int_0^t (t-x) F_X(x) dx \\ &= 2t\tilde{\mu}(t) - \frac{2}{F_X(t)} \int_0^t x F_X(x) dx.\end{aligned} \tag{2.4}$$

Theorem 2.1. *Let X be a non-negative continuous random variable with mean inactivity time function $\tilde{\mu}(t)$ and weighted cumulative entropy $\mathcal{CE}^w(X) < \infty$. Then,*

$$\mathcal{CE}^w(X) = E\Big(X\tilde{\mu}(X)\Big) - \frac{1}{2}E\Big(\eta_2(X)\Big),$$

where $\eta_2(t)$ is defined by (2.3).

Proof. By (2.4), we have

$$\begin{aligned}\frac{1}{2}E\Big(\eta_2(X)\Big) &= -\int_0^\infty \Big[\int_0^t x\frac{F_X(x)}{F_X(t)} dx\Big] f_X(t) dt + E\Big(X\tilde{\mu}(X)\Big) \\ &= -\int_0^\infty x F_X(x) \int_x^\infty r_X(t) dt dx + E\Big(X\tilde{\mu}(X)\Big) \\ &= \int_0^\infty x F_X(x) \log F_X(x) dx + E\Big(X\tilde{\mu}(X)\Big) \\ &= -\mathcal{CE}^w(X) + E\Big(X\tilde{\mu}(X)\Big).\end{aligned}$$

and the proof is completed. □

2.2 Empirical Weighted Cumulative Entropy

In this subsection, an estimate of the WCE is constructed by means of the empirical WCE. Let $X_1, X_2, \cdots, X_n$ be a non-negative, continuous, independent and identically distributed random sample from a population having the distribution function $F_X(x)$. By using the plug-in method, we define the empirical weighted cumulative entropy as

$$\mathcal{CE}_n^w(X) = -\int_0^\infty x F_n(x) \log F_n(x) dx,$$

where $F_n(x)$ is the empirical distribution function. Denoting the order statistics of the random sample by $(0 = X_{(0)}) \leq X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$, we have

$$\mathcal{CE}_n^w(X) = -\sum_{i=1}^{n-1} \frac{1}{2} U_i \frac{i}{n} \log \frac{i}{n}, \tag{2.5}$$

where we use the sample spacings

$$U_i = [X_{(i+1)}^2 - X_{(i)}^2], \quad i = 1, 2, \ldots, n-1,$$

and recall that

$$F_n(x) = \frac{i}{n}, \quad X_{(i)} \leq x < X_{(i+1)}, \quad i = 1, 2, \ldots, n-1.$$

Theorem 2.2. *Let $X_1, X_2, \ldots, X_n$ be a random sample from Rayleigh distribution with pdf $f_\lambda(x) = 2\lambda x e^{-\lambda x^2}$, $x > 0$, $\lambda > 0$. Then,*

$$Z_n := \frac{\mathcal{CE}_n^w(X) - E[\mathcal{CE}_n^w(X)]}{\sqrt{\mathrm{var}[\mathcal{CE}_n^w(X)]}}$$

converges in distribution to a standard normal variable as $n \to \infty$.

Proof. Relation (2.5) can be written as

$$\mathcal{CE}_n^w(X) = \sum_{i=1}^{n-1} Y_i,$$

where $Y_i = -\frac{1}{2} U_i \frac{i}{n} \log \frac{i}{n}$, $i = 1, \ldots, n-1$. According to the fact that, if X has a Rayleigh distribution with parameter λ, then $Z = X^2$ has an exponential distribution with parameter λ, we have

$$E[Y_i] = -\frac{i}{2n} (\log \frac{i}{n}) \frac{1}{\lambda(n-i)},$$
$$\mathrm{var}[Y_i] = \frac{i^2}{4n^2} \Big(\log \frac{i}{n} \Big)^2 \frac{1}{\lambda^2 (n-i)^2}.$$

Since $E[|Z - E(Z)|^3] = 2e^{-1}(6-e)[E(Z)]^3$ for any exponentially distributed random variable Z, by setting $\alpha_{i,k} = E[|Y_i - E(Y_i)|^k]$, the approximations

$$\sum_{i=1}^{n} \alpha_{i,2} = \frac{1}{4\lambda^2 n^2} \sum_{i=1}^{n} \Big(\frac{i \log \frac{i}{n}}{n-i} \Big)^2 \approx \frac{c_2}{4\lambda^2 n},$$
$$\sum_{i=1}^{n} \alpha_{i,3} = \frac{-(6-e)}{4en^3\lambda^3} \sum_{i=1}^{n} \Big(\frac{i \log \frac{i}{n}}{n-i} \Big) \approx \frac{-(6-e)c_3}{4en^2\lambda^3},$$

hold for large n, where

$$c_k := \int_0^1 \Big(\frac{x}{1-x}\log x\Big)^k dx = \begin{cases} 0.481640522, & k=2 \\ -0.385766882, & k=3. \end{cases}$$

Since,

$$\frac{(\alpha_{1,3}+\alpha_{2,3}+\cdots+\alpha_{n,3})^{\frac{1}{3}}}{(\alpha_{1,2}+\alpha_{2,2}+\cdots+\alpha_{n,2})^{\frac{1}{2}}} \approx n^{-\frac{1}{6}} \to 0 \quad \text{as} \quad n \to \infty,$$

Lyapunov's condition of the central limit theorem is satisfied (see Gut (2005)). This completes the proof. □

3 Some Characterization Results Based on Order Statistics

In this section, for some characterization results, the following Lemma is needed.

Lemma 3.1. *If η is a continuous function on $[0,1]$, such that $\int_0^1 x^n\eta(x)dx = 0$, for $n \geq 0$, then $\eta(x) = 0$, for all $x \in [0,1]$.*

In fact, Lemma 3.1 is a corollary of Stone-Weierstrass Theorem (see Aliprantis and Burkinshaw (1981)). In the following, we will achieve the WCRE of the first and the last order statistics. The WCRE of the first order statistic is

$$\mathcal{E}^w(X_{(1)}) = -\int_0^\infty x\bar{F}_{X_{(1)}}(x)\log \bar{F}_{X_{(1)}}(x)dx, \tag{3.1}$$

where $\bar{F}_{X_{(1)}}(x) = \bar{F}_X^n(x)$. By changing the variable to $u = \bar{F}_X(x)$ in (3.1),

$$\mathcal{E}^w(X_{(1)}) = -n\int_0^1 F_X^{-1}(1-u)u^n \frac{\log u}{f_X\big(F_X^{-1}(1-u)\big)}du, \tag{3.2}$$

where $F_X^{-1}(x)$ is the inverse function of $F_X(x)$. Also, it can be shown that

$$\mathcal{E}^w(X_{(n)}) = -\int_0^1 \frac{F_X^{-1}(u)(1-u^n)\log(1-u^n)}{f_X\big(F_X^{-1}(u)\big)}du. \tag{3.3}$$

Theorem 3.1. *Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be random samples from non-negative continuous cdfs $F(x)$ and $G(x)$ and pdfs $f(x)$ and $g(x)$, respectively, with common support $[0,\infty)$. Then, $F(x) = G(x)$ if and only if $\mathcal{E}^w(X_{(1)}) = \mathcal{E}^w(Y_{(1)})$, for all n.*

Proof. The necessity is trivial. Therefore, it remains to prove the sufficiency part. If $\mathcal{E}^w(X_{(1)}) = \mathcal{E}^w(Y_{(1)})$, then we have

$$\int_0^1 u^n \log u\Big[\frac{F^{-1}(1-u)}{f\big(F^{-1}(1-u)\big)} - \frac{G^{-1}(1-u)}{g\big(G^{-1}(1-u)\big)}\Big]du = 0.$$

By Lemma 3.1, we conclude that $\frac{F^{-1}(1-u)}{f\left(F^{-1}(1-u)\right)} = \frac{G^{-1}(1-u)}{g\left(G^{-1}(1-u)\right)}$, $0 \leq u \leq 1$. By assuming that $v = 1-u$, we have $F^{-1}(v)\frac{d}{dv}F^{-1}(v) = G^{-1}(v)\frac{d}{dv}G^{-1}(v)$, $0 \leq v \leq 1$. Since $\frac{d}{dv}F^{-1}(v) = \frac{1}{f\left(F^{-1}(v)\right)}$, it will be concluded that $F^{-1}(v) = G^{-1}(v)$, $0 \leq v \leq 1$. Thus, the proof is completed. □

WCE of the last order statistic is

$$\mathcal{CE}^{\omega}(X_{(n)}) = -n\int_0^{\infty} xF_X^n(x)\log F_X(x)dx, \tag{3.4}$$

where $F_X^n(x) = F_{X_{(n)}}(x)$. With a change of variable, $u = F_X(x)$, we are able to write

$$\mathcal{CE}^{\omega}(X_{(n)}) = -n\int_0^1 F_X^{-1}(u)u^n\frac{\log u}{f_X\big(F_X^{-1}(u)\big)}du. \tag{3.5}$$

Also, by using $\bar{F}_{X_{(1)}}(x) = 1 - \bar{F}_X^n(x)$, we have

$$\mathcal{CE}^{\omega}(X_{(1)}) = -\int_0^{\infty} x\Big(1-\bar{F}_X^n(x)\Big)\log\Big(1-\bar{F}_X^n(x)\Big)dx. \tag{3.6}$$

With another change of variable, $u = \bar{F}_X^n(x)$ in (3.6), we have

$$\mathcal{CE}^{\omega}(X_{(1)}) = -\int_0^1 F_X^{-1}(1-u^{\frac{1}{n}})(1-u)\log(1-u)\frac{u^{\frac{1}{n}-1}}{nf_X\Big(F_X^{-1}(1-u^{\frac{1}{n}})\Big)}du.$$

Theorem 3.2. *Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be random samples from non-negative continuous cdfs $F(x)$ and $G(x)$ and pdfs $f(x)$ and $g(x)$, respectively, with common support $[0,\infty)$. Then $F(x) = G(x)$ if and only if $\mathcal{CE}^w(X_{(n)}) = \mathcal{CE}^w(Y_{(n)})$, for all n.*

Proof. The necessity is trivial. Therefore, it remains to prove the sufficiency part. If $\mathcal{CE}^w(X_{(n)}) = \mathcal{CE}^w(Y_{(n)})$, for all n, then we have

$$\int_0^1 u^n \log u\Big(\frac{F^{-1}(u)}{f\big(F^{-1}(u)\big)} - \frac{G^{-1}(u)}{g\big(G^{-1}(u)\big)}\Big)du = 0.$$

By using Lemma 3.1, we have

$$\frac{F^{-1}(u)}{f\big(F^{-1}(u)\big)} = \frac{G^{-1}(u)}{g\big(G^{-1}(u)\big)}, \tag{3.7}$$

for all n. The rest of the proof is similar to the proof of Theorem 3.1. □

Now, we evaluate the WCRE of $X_{(1)}$ and WCE of $X_{(n)}$ for some distributions.

Examples 3.1. Suppose that X has an exponential distribution with mean $\frac{1}{\lambda}$. Then,

(i) $\mathcal{E}^w(X) = \frac{2}{\lambda^2}$, $\mathcal{E}^w(X_{(1)}) = \frac{2}{(n\lambda)^2}$, $\mathcal{E}(X_{(1)}) = \frac{1}{n\lambda}$ and $E(X_{(1)}^2) = \frac{2}{(n\lambda)^2}$. Thus, we have $\mathcal{E}^w(X_{(1)}) = E(X_{(1)}^2)$, $\mathcal{E}^w(X) = n^2 E(X_{(1)}^2) = n^2\mathcal{E}^w(X_{(1)})$ and $\mathcal{E}^w(X_{(1)}) = 2\Big(\mathcal{E}(X_{(1)})\Big)^2$.

(ii) $C\mathcal{E}^w(X_{(n)}) = \frac{n}{\lambda^2}[C + \sum_{k=1}^{n} C_n^k(-1)^k \frac{1}{k}\Big[\frac{1}{k}\sum_{j=1}^{n}\frac{1}{j} + \text{dilog}(0) + \sum_{j=1}^{n}\frac{1}{j^2}\Big]$,

where $C = \int_0^1 \frac{\log x}{x}\log(1-x)dx = 1.202056903$ and $\text{dilog}(0) = \int_1^0 \frac{\log x}{1-x}dx = \frac{\pi^2}{6}$.

Examples 3.2. Let X have a Pareto distribution with pdf $f(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}}$, $x \geq \beta$, $\beta > 0$ and $\alpha > 0$. Then,

(i) By (1.1) and (3.2), we have

$$\mathcal{E}^w(X) = \begin{cases} \frac{\alpha\beta^2}{(\alpha-2)^2}, & \alpha > 2 \\ \infty, & \alpha \leq 2, \end{cases}$$

and

$$\mathcal{E}^w(X_{(1)}) = \begin{cases} \frac{n\alpha\beta^2}{(n\alpha-2)^2}, & \alpha > \frac{2}{n} \\ \infty, & \alpha \leq \frac{2}{n}. \end{cases}$$

For $\alpha > \frac{2}{n}$, we derive $\mathcal{E}^w(X_{(1)}) = CnE^2(X_{(1)}^2)$, where $E(X_{(1)}^2) = \frac{2\beta^2}{n\alpha-2}$ and $C = \frac{\alpha}{4\beta^2}$.
Let $\Delta_1 = \mathcal{E}^w(X) - \mathcal{E}^w(X_{(1)})$. Then, according to $\Delta_1 \geq 0$ for $\alpha > 2$ and all n, the uncertainty of X is more than the uncertainty of $X_{(1)}$, for all n.

(ii) By (3.5), we have

$$C\mathcal{E}^w(X_{(n)}) = -\frac{n\beta^2}{\alpha}\int_0^1 u^n \log u(1-u)^{-(1+\frac{2}{\alpha})}du.$$

Using binomial series for $(1-u)^{-(1+\frac{2}{\alpha})}$, we conclude that

$$C\mathcal{E}^w(X_{(n)}) = \frac{n\beta^2}{\alpha}\sum_{j=0}^{\infty} a_j(-1)^j \frac{1}{(n+j+1)^2}, \tag{3.8}$$

where $a_0 = 1$ and $a_j = \frac{[-(1+\frac{2}{\alpha})][-(1+\frac{2}{\alpha})-1]\cdots[-(1+\frac{2}{\alpha})-j+1]}{j!}$.

By taking $n = 1$ in (3.8), we conclude that $C\mathcal{E}^w(X) = \frac{\beta^2}{\alpha}\sum_{j=0}^{\infty} a_j(-1)^j \frac{1}{(j+2)^2}$.

Examples 3.3. Suppose X has Rayleigh distribution with density function $f_\lambda(x) = 2\lambda x e^{-\lambda x^2}$, $x > 0$, $\lambda > 0$.

Using (1.1) and (3.1), we have

$$\mathcal{E}^w(X) = \frac{1}{2\lambda}, \quad \mathcal{E}^w(X_{(1)}) = \frac{1}{2n\lambda}.$$

Thus, we conclude that $n\mathcal{E}^w(X_{(1)}) = \mathcal{E}^w(X)$.

Let $\Delta_2 = \mathcal{E}^w(X) - \mathcal{E}^w(X_{(1)})$, then according to $\Delta_2 \geq 0$, for all n, we conclude that the uncertainty of X is more than the uncertainty of $X_{(1)}$. Also, we can obtain $\mathcal{E}^w(X_{(1)}) = \frac{1}{2}E(X^2_{(1)})$.

(ii) By (3.5), we have

$$\begin{aligned} \mathcal{CE}^w(X_{(n)}) &= \frac{n}{2\lambda}\Big[\text{dilog}(0) - \sum_{j=1}^{n} \frac{1}{i^2}\Big] \\ &= \frac{n}{2\lambda}\Big[\frac{\pi^2}{6} - \sum_{i=1}^{n} \frac{1}{i^2}\Big], \end{aligned}$$

where $\text{dilog}(0) = \int_1^0 \frac{\log x}{1-x}dx = \frac{\pi^2}{6}$. Also, $\mathcal{CE}^w(X) = \frac{1}{2\lambda}\Big[\frac{\pi^2}{6} - 1\Big]$.

Let $\Delta_3 = \mathcal{CE}^w(X_{(n)}) - \mathcal{CE}^w(X)$, then according to $\Delta_3 \geq 0$ for all n, we conclude that the uncertainty of $X_{(n)}$ is more than the uncertainty of X.

Theorem 3.3. *Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be random samples from non-negative continuous distributions $F(x)$ and $G(x)$, respectively. If $F(x_0) = G(x_0)$, then $F(x) = G(x)$, for $x > x_0$, if and only if*

$$\mathcal{E}^w(X_{(j)} \big| X_{(j-1)} = x_0) = \mathcal{E}^w(Y_{(j)} \big| Y_{(j-1)} = x_0).$$

Proof. If $F(x) = G(x)$ for $x > x_0$, then by assumption $F(x_0) = G(x_0)$, $F(x)$ and $G(x)$, truncated on the left at x_0, are equal for $x > x_0$, that is,

$$\frac{F(x) - F(x_0)}{1 - F(x_0)} = \frac{G(x) - G(x_0)}{1 - G(x_0)}, \quad x > x_0.$$

By Theorem 2.4.1 of Arnold et al. (1992), the distribution of $X_{(j)}$ given that $X_{(j-1)} = x_0$ is the same as the distribution of the first order statistic obtained from a sample of size $n - j + 1$ from a population whose distribution $F(x)$ is truncated on the left at x_0. By Theorem 3.1, we conclude that

$$\mathcal{E}^w(X_{(j)} \big| X_{(j-1)} = x_0) = \mathcal{E}^w(Y_{(j)} \big| Y_{(j-1)} = x_0).$$

Conversely, let $\mathcal{E}^w(X_{(j)} \big| X_{(j-1)} = x_0) = \mathcal{E}^w(Y_{(j)} \big| Y_{(j-1)} = x_0)$, that is WCRE of the first order statistic for two distributions $F(x)$ and $G(x)$ truncated on the left at x_0 are equal. Thus, by Theorem 3.1, this two truncated distributions are equal, which leads to $F(x) = G(x)$ for $x > x_0$. □

Theorem 3.4. *Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be random samples from non-negative continuous distributions $F(x)$ and $G(x)$, respectively. If $F(x_0) = G(x_0)$, then $F(x) = G(x)$, for $x < x_0$, if and only if*

$$\mathcal{CE}^w(X_{(j)}\big|X_{(j+1)} = x_0) = \mathcal{CE}^w(Y_{(j)}\big|Y_{(j+1)} = x_0).$$

Proof. Using Theorem 2.4.2 of Arnold et al. (1992) and Theorem 3.2, the proof is similar to the Theorem 3.3. □

4 Conclusion

In this paper, we considered a weighted entropy, called WCE, which is based on the cdf. The properties of this measure and its connection with reliability were investigated. An estimate of the WCE was constructed by means of the empirical cdf. Also, some characterization results were given, in patrticular, some results on the first and last order statistics.

Acknowledgements

The authors are grateful to the referees for the very useful and constructive comments on the earlier version of this paper, which have led to a substantial improvement of the paper.

References

Ahmad, I. A. and Kayid, M. (2005). Characterizations of the RHR and MIT orderings and the DRHR and IMIT classes of life distributions. *Probability in the Engineering and Informational Sciences*, **19**, 447-461.

Ahmad, I. A., Kayid, M. and Pellery, F. (2005). Further results involving the MIT order and IMIT class. *Probability in the Engineering and Informational Sciences*, **19**, 377-395.

Aliprantis, C. D. and Burkinshaw, O. (1981). *Principles of Real Analysis.* Elsevier North Holland Inc.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics.* New York: John Wiley & Sons.

Asadi, M. and Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. *Journal of Statistical Planning and Inference*. **137**, 1931-1941.

Di Crescenzo, A. and Longobardi, M. (2006). On weighted residual and past entropies. *arXiv preprint math/0703489*. **64**, 255-266.

Di Crescenzo, A. and Longobardi, M. (2009). On cumulative entropies. *Journal of Statistical Planning and Inference* **139**, 4072-4087.

Gut, A. (2005). *Probability: A graduate course.* Springer, New York.

Johnson, D. H. and Glantz, R. M. (2004). When does interval coding occur?. *Neurocomputing* **58**, 13-18.

Kapodistria, S. and Psarrakos, G. (2012). Some extensions of the residual lifetime and its connection to the cumulative residual entropy. *Probability in the Engineering and Informational Sciences* **26**, 129-146.

Kayid, M. and Ahmad, I. A. (2004). On the mean inactivity time ordering with reliability applications. *Probability in the Engineering and Informational Sciences* **18**, 395-409.

Li, X. and Lu, J. (2003). Stochastic comparisons on residual life and inactivity time of series and parallel systems. *Probability in the Engineering and Informational Sciences* **17**, 267-275.

Mirali, M. and Baratpour, S. (2015). Dynamic version of weighted cumulative residual entropy. *Communications in Statistics-Theory and Methods*, doi: 10.1080/03610926.2016.1257711.

Mirali, M., Baratpour, S. and Fakoor, V. (2015). On weighted cumulative residual entropy. *Communications in Statistics-Theory and Methods*, **46**, 2857-2869.

Misagh, F., Panahi, Y., Yari, G. H. and Shahi, R. (2011). Weighted cumulative entropy and its estimation. *In Quality and Reliability (ICQR), 2011, IEEE International conference on* (pp. 477-480), IEEE.

Misra, N., Gupta, N. and Dhariyal, I. D. (2008). Stochastic properties of residual life and inactivity time at a random time. *Stochastic Models* **24**, 89-102.

Nanda, A. K., Singh, H., Misra, N. and Paul, P. (2003). *Reliability properties of reversed residual lifetime. Communications in Statistics-Theory and Methods* **32**, 2031-2042. With correction in Comm. Statist. Theory Methods (2004), **33**, 991-992.

Navarro, J., Aguila, Y. D. and Asadi, M. (2010). *Some new results on the cumulative residual entropy. Journal of Statistical Planning and Inference* **140**, 310-322.

Psarrakos, G. and Navarro, J. (2013). Generalized cumulative residual entropy and record values. *Metrika* **76** (5) , 623-640.

Rao, M., Chen, Y., Vemuri, B. C. and Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE transactions on Information Theory* **50** (6), 1220-1228.

Shannon, C. E. (1948). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review* **27**, 379-432.

Sunoj, S. M. and Linu, M. N. (2012). Dynamic cumulative residual Renyi's entropy. *Statistics* **46**, 41-56.

Wiener, N. (1948, 2nd Ed. 1961). *Cybernetics*, The MIT Press and Wiley, New York.