

Article

Not peer-reviewed version

Estimation of Weighted Extropy with a Focus on Its Use in Reliability Modeling

[Muhammed Rasheed Irshad](#) , Archana Krishnakumar , [Radhakumari Maya](#) , [Maria Longobardi](#) *

Posted Date: 3 January 2024

doi: 10.20944/preprints202401.0275.v1

Keywords: Weighted extropy; Log kernel estimation; Kernel estimation; Empirical estimation; simulation.



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Estimation of Weighted Extropy with a Focus on Its Use in Reliability Modeling

Muhammed Rasheed Irshad ¹, Krishnakumar Archana ¹, Radhakumari Maya ² and Maria Longobardi ^{3,*} 

¹ Cochin University of Science and Technology, Kochi, India;

² Government College for Women, Thiruvananthapuram, India;

³ University of Napoli Federico II, Italy

* Correspondence: maria.longobardi@unina.it

Abstract: In the literature, estimation of weighted extropy are least considered. In this paper some non parametric estimators of weighted extropy are given. The validation and comparison of the estimators are implemented with the help of simulation study and data illustration. The usefulness of the estimators are demonstrated using real data sets.

Keywords: Weighted extropy; Log kernel estimation; Kernel estimation; Empirical estimation; simulation

1. Introduction

The concept of extropy and its use have been increased rapidly in the recent years. It measures the uncertainty contained in the probability distributions. For a non negative random variable (rv) X , the entropy introduced by [29] is given by

$$H(X) = - \int_0^{+\infty} f_X(x) \log f_X(x) dx, \quad (1)$$

where $f_X(x)$ is the probability density function (pdf) of a X . This measure is shift independent, that is, it is same for both X and $X + b$ and it cannot be applied in some fields such as in neurology. Thus [4] introduced the notion of weighted entropy measure as

$$H^w(X) = - \int_0^{+\infty} x f_X(x) \log f_X(x) dx. \quad (2)$$

They pointed out that occurrence of an event has an impact on uncertainty in two ways. Being quantitative and qualitative, it first shows the probability of occurrence of an event and later it shows utility in achieving qualitative characteristics of a goal. The variable x in the integral emphasize the weight related to the occurrence of the event $X = x$. Here it assigns more significance to large values of X . It is important to note that the information obtained when a device fails to operate or a neuron fails to release spikes in a specific time interval, differs significantly from the information obtained when such events occur in other equally wide intervals. This is why there is a need, in some cases, to employ a shift-dependent information measure that assigns varying measures to these distributions. The importance of the presence of weighted measures of uncertainty has been exhibited by [9].

The concept of extropy for a continuous rv X has been presented and discussed across numerous works in the literature. The differential extropy defined by [15] is:

$$J(X) = - \frac{1}{2} \int_0^{+\infty} f_X^2(x) dx. \quad (3)$$

One can refer [23] for the extropy properties of order statistics and record values. The applications of extropy in automatic speech recognition can be found in [3]. Various literature sources have presented

a range of extropy measures and their extensions. Analogous to weighted entropy, [1] introduced the concept of weighted extropy (WE) in the literature. It is given as

$$\mathfrak{J}^w(X) = -\frac{1}{2} \int_0^{+\infty} x f_X^2(x) dx, \quad (4)$$

which again can be alternatively expressed as

$$\mathfrak{J}^w(X) = -\frac{1}{2} \int_0^{+\infty} dy \int_y^{+\infty} f_X^2(x) dx. \quad (5)$$

They also illustrated that there exist some distributions with same extropy but different WE and vice versa. In the literature extropy, its different versions and their applications, have been studied by several authors (see, for instance, [6], [1], [14], [7]). In particular, a unified version of extropy in classical theory and in Dempster-Shafer theory has been studied by [5].

Let us define a *rv* X with unknown *pdf* $f_X(x)$. We assume that X is defined on R and $f_X(x)$ is continuously differentiable. Suppose $\{X_i; 1 \leq i \leq n\}$ be a sequence of independent and identically distributed (*iid*) *rvs*. The most commonly used estimator of $f_X(x)$ is the kernel density estimator (*KDE*), given by [22] and [27] as

$$\hat{f}_X(x) = \frac{1}{n\mathfrak{h}} \sum_{i=1}^n \mathfrak{K}\left(\frac{x - X_i}{\mathfrak{h}}\right), \quad (6)$$

where $\mathfrak{K}(x)$ is the kernel function which satisfies the following conditions.

- $\int_R \mathfrak{K}(x) dx = 1$
- $\int_R x \mathfrak{K}(x) dx = 0$
- $\int_R x^2 \mathfrak{K}(x) dx = 1$
- $\int_R \mathfrak{K}(x) dx < +\infty$

Here, \mathfrak{h}_n is a sequence of positive bandwidths such that $\mathfrak{h}_n \rightarrow 0$ and $n\mathfrak{h}_n \rightarrow +\infty$ as $n \rightarrow +\infty$.

When probability density functions are estimated in a non parametric way, standard kernel density estimators are frequently used. However, when we deal with data that fits distributions with heavy tails, multiple modes, or skewness, particularly those with positive values these estimators may lose their effectiveness. In all of these scenarios, applying a transformation, we can yield more consistent results. Such a transformation involves a logarithmic transformation to create a non parametric kernel estimator. An important aspect of the logarithmic transformation is its ability to compress the right tail of the distribution. The obtained *KDE* are called logarithmic *KDE* (denoted as $L - KDE$), (check, [8]). Let us define $Y = \log(X)$, $Y_i = \log(X_i)$; $i = 1, 2, \dots, n$ and let $f_Y(y)$ be the *pdf* of Y . The $L - KDE$ is defined as,

$$\begin{aligned} \hat{f}_{\log}(x) &= \frac{1}{n\mathfrak{h}} \sum_{i=1}^n \frac{1}{x} \mathfrak{K}\left(\frac{\log x - \log X_i}{\mathfrak{h}}\right) \\ &= \frac{1}{n} \sum_{i=1}^n L(x, X_i, \mathfrak{h}), \end{aligned} \quad (7)$$

where $L(x, z, \mathfrak{h}) = \frac{1}{x\mathfrak{h}} \mathfrak{K}\left(\log\left(\frac{x}{z}\right)\frac{1}{\mathfrak{h}}\right)$ is the log-kernel function with bandwidth $\mathfrak{h} > 0$, at location parameter z . For any $z, \mathfrak{h} \in (0, +\infty)$, $L(x, z, \mathfrak{h})$ satisfies the conditions: $L(x, z, \mathfrak{h}) \geq 0$ for all $x \in (0, +\infty)$ and $\int_0^{+\infty} L(x, z, \mathfrak{h}) dx = 1$.

For any $X \in (0, +\infty)$,

$$\text{Bias}(\hat{f}_{\log}(x)) = \frac{\mathfrak{h}^2}{2} \left[f_X(x) + 3x f_X^{(1)}(x) + x^2 f_X^{(2)}(x) \right] + O(\mathfrak{h}^2). \quad (8)$$

$$\text{Var}(\hat{f}_{\log}(x)) = \frac{\mathfrak{C}_k}{n\mathfrak{h}} \frac{f_X(x)}{x} dx + O\left(\frac{1}{n\mathfrak{h}}\right), \quad (9)$$

where $\mathfrak{C}_k = \int_{\mathbb{R}} \mathfrak{K}^2(z) dz$.

There are several papers available in the literature that delve into the estimation of entropy and its various versions. [24] proposed estimators of entropy and also worked on its application by testing uniformity. [26] approached the concept of length biased sampling in estimating entropy. Research on non parametric estimation using dependent data are also well explored in the literature. These covered the estimation of residual and past entropy both under α -mixing dependence condition and can be found in works such as [18] and [11]. Additionally, a work by [19] explained about the recursive and non recursive kernel estimation of negative cumulative entropy under α -mixing dependence condition. Recently, [20] discussed about the kernel estimation of entropy function using α -mixing dependent data. Moreover, [12] introduced the log kernel estimation of entropy.

Even if there are several works available on the literature related to the estimation of entropy, just a little has been done on *WE* and its estimation until now. There are situations in which we are forced to use *WE* instead of entropy. Unlike entropy, the qualitative characteristics of information are also represented here. [28] demonstrated the significance of employing *WE* as opposed to regular entropy in certain scenarios. There are instances where certain distributions possess identical entropy values but exhibit distinct *WE* values. In such situations, it becomes necessary to opt for weighted entropy. The estimators of *WE* can also be used in the selection of models in the reliability analysis also. Here we tried to find some estimators for *WE* and validated it using simulation study and data analysis.

The paper is organized as follows: In Section 2, we have introduced the log kernel estimation of *WE*. In Section 3 the estimation of *WE* using empirical kernel smoothed estimator is given. A simulation study is conducted to evaluate the estimators and we also made a comparison between log kernel and kernel estimation of *WE* in Section 4. Section 5 devotes to the real data analysis to examine the proposed estimators. Finally, we conclude the study in Section 6

2. Log Kernel Estimation Of Weighted Entropy

In this section we introduce the concept of log kernel based estimation of *WE*. Let (X_1, X_2, \dots, X_n) be a sample of *iid* observations. We obtain the log kernel estimators for *WE* by using the estimator defined in equation 7.

The $L - \text{KDE}$ for *WE* function is

$$\begin{aligned} \hat{\mathfrak{J}}_n^w(X) &= -\frac{1}{2} \int_0^{+\infty} x \hat{f}_{\log}^2(x) dx \\ &= -\frac{1}{2} \int_0^{+\infty} dy \int_y^{+\infty} \hat{f}_{\log}^2(x) dx. \end{aligned} \quad (10)$$

Theorem 1. The bias and variance of $\hat{\mathfrak{J}}_n^w(X)$ are given, respectively as

$$\text{Bias}(\hat{\mathfrak{J}}_n^w(X)) = -\int_0^{+\infty} dy \int_y^{+\infty} \frac{\mathfrak{h}^2}{2} \left[f_X(x) + 3x f_X^{(1)}(x) + x^2 f_X^{(2)}(x) \right] f_X(x) dx + O(\mathfrak{h}^2) \quad (11)$$

$$\text{Var}(\hat{\mathfrak{J}}_n^w(X)) = -\frac{\mathfrak{C}_k}{n\mathfrak{h}} \int_0^{+\infty} dy \int_y^{+\infty} \frac{f_X^{(3)}(x)}{x} dx + O\left(\frac{1}{n\mathfrak{h}}\right), \quad (12)$$

where $\mathfrak{C}_k = \int_{\mathbb{R}} K^2(z) dz$

Proof. By using Taylor's series expansion,

$$\begin{aligned}\hat{f}_{log}^2(x) &\simeq f_X^2(x) + (\hat{f}_{log}(x) - f_X(x)) f_X(x) \\ -\frac{1}{2} \int_0^{+\infty} dy \int_y^{+\infty} \hat{f}_{log}^2(x) dx &\simeq -\frac{1}{2} \int_0^{+\infty} dy \int_y^{+\infty} f_X^2(x) dx \\ &\quad + -\frac{2}{2} \int_0^{+\infty} dy \int_y^{+\infty} (\hat{f}_{log}(x) - f_X(x)) f_X(x) dx \\ \hat{\mathfrak{J}}_n^w(X) &\simeq \mathfrak{J}^w(X) - \int_0^{+\infty} dy \int_y^{+\infty} (\hat{f}_{log}(x) - f_X(x)) f_X(x) dx.\end{aligned}$$

Then,

$$\begin{aligned}Bias(\hat{\mathfrak{J}}_n^w(X)) &\simeq - \int_0^{+\infty} dy \int_y^{+\infty} Bias(\hat{f}_{log}(x)) f_X(x) dx \\ &= - \int_0^{+\infty} dy \int_y^{+\infty} \frac{\mathfrak{h}^2}{2} [f_X(x) + 3xf_X^{(1)}(x) + x^2f_X^{(2)}(x)] f_X(x) dx + O(\mathfrak{h}^2)\end{aligned}$$

and

$$\begin{aligned}Var(\hat{\mathfrak{J}}_n^w(X)) &\simeq \int_0^{+\infty} dy \int_y^{+\infty} Var(\hat{f}_{log}(x)) f_X^2(x) dx \\ &= \int_0^{+\infty} dy \int_y^{+\infty} \frac{1}{n\mathfrak{h}x} f_X(x) \int_R K^2(z) dz + O\left(\frac{1}{n\mathfrak{h}}\right) f_X^{(2)}(x) dx \\ &= -\frac{\mathfrak{C}_k}{n\mathfrak{h}} \int_0^{+\infty} dy \int_y^{+\infty} \frac{f_X^{(3)}(x)}{x} dx + O\left(\frac{1}{n\mathfrak{h}}\right).\end{aligned}$$

Hence the proof. \square

Theorem 2. $\hat{\mathfrak{J}}_n^w(X)$ is a consistent estimator of $\mathfrak{J}^w(X)$, where $\hat{\mathfrak{J}}_n^w(X)$ and $\mathfrak{J}^w(X)$ are defined in equations (10) and (4). Then, we can say that, as n tends to $+\infty$,

$$\hat{\mathfrak{J}}_n^w(X) = -\frac{1}{2} \int_0^{+\infty} dy \int_y^{+\infty} \hat{f}_{log}^2(x) \xrightarrow{p} -\frac{1}{2} \int_0^{+\infty} dy \int_y^{+\infty} f_X^{(2)}(x) dx = \mathfrak{J}^w(X). \quad (13)$$

Proof. From equations (11) and (12),

$$\begin{aligned}MSE(\hat{\mathfrak{J}}_n^w(X)) &= \left[-\frac{\mathfrak{h}^2}{2} \int_0^{+\infty} dy \int_y^{+\infty} [f_X(x) + 3xf_X^{(1)}(x) + x^2f_X^{(2)}(x)] f_X(x) dx \right]^2 \\ &\quad + \frac{\mathfrak{C}_k}{n\mathfrak{h}} \int_0^{+\infty} dy \int_y^{+\infty} \frac{f_X^{(3)}(x)}{x} dx + O(\mathfrak{h}^4) + O\left(\frac{1}{n\mathfrak{h}}\right).\end{aligned} \quad (14)$$

As n approaches to $+\infty$,

$$MSE(\hat{\mathfrak{J}}_n^w(X)) \rightarrow 0. \quad (15)$$

Therefore, we can say that $\hat{\mathfrak{J}}_n^w(X)$ is a consistent estimator of $\mathfrak{J}^w(X)$. \square

Theorem 3. If $\hat{\mathfrak{J}}_n^w(X)$ is the log kernel estimator according to equation (10), then $\hat{\mathfrak{J}}_n^w(X)$ behaves consistently and uniformly in terms of MSE as an estimator of $\mathfrak{J}^w(X)$.

Proof. Let us denote $MISE(\hat{\mathfrak{J}}_n^w(X))$ as the Mean Integrated Square Estimator of $\hat{\mathfrak{J}}_n^w(X)$. Here,

$$\begin{aligned} MISE(\hat{\mathfrak{J}}_n^w(X)) &= E \left[\int_{-\infty}^{+\infty} (\hat{\mathfrak{J}}_n^w(X) - \mathfrak{J}^w(X))^2 dx \right] \\ &= \int_{-\infty}^{+\infty} [\text{Var}(\hat{\mathfrak{J}}_n^w(X)) + \text{Bias}(\hat{\mathfrak{J}}_n^w(X))^2] dx \\ &= \int_{-\infty}^{+\infty} MSE(\hat{\mathfrak{J}}_n^w(X)) dx. \end{aligned}$$

Thus,

$$\begin{aligned} MISE(\hat{\mathfrak{J}}_n^w(X)) &= \int_0^{+\infty} \left[\left(-\frac{\mathfrak{h}^2}{2} \int_0^{+\infty} dy \int_y^{+\infty} [f_X(x) + 3xf_X^{(1)}(x) + x^2f_X^{(2)}(x)] f_X(x) dx \right)^2 \right. \\ &\quad \left. + \frac{\mathfrak{C}_k}{n\mathfrak{h}} \int_0^{+\infty} dy \int_y^{+\infty} \frac{f_X^{(3)}(x)}{x} dx + O(\mathfrak{h}^4) + O\left(\frac{1}{n\mathfrak{h}}\right) \right] dx. \end{aligned} \quad (16)$$

When n tends to $+\infty$, $MISE$ tends to 0. Hence we conclude that as n increases, $\hat{\mathfrak{J}}_n^w(X)$ is consistent integratedly in quadratic mean estimator of $\mathfrak{J}^w(X)$. (refer,[30]) \square

2.1. Optimal Bandwidth

Here, we give the expression for optimal bandwidth using equation (16). The asymptotic $MISE$ (denoted as $AMISE$) can be obtained by ignoring the higher order terms. The optimal bandwidth is attained after minimizing $AMISE$ with respect to \mathfrak{h} .

$$\begin{aligned} AMISE &= \frac{\mathfrak{h}^4}{4} \int_0^{+\infty} \left[\left(\int_0^{+\infty} dy \int_y^{+\infty} [f_X(x) + 3xf_X^{(1)}(x) + x^2f_X^{(2)}(x)] f_X(x) dx \right)^2 \right] dx \\ &\quad + \frac{1}{n\mathfrak{h}} \int_0^{+\infty} \left[\mathfrak{C}_k \int_0^{+\infty} dy \int_y^{+\infty} \frac{f_X^{(3)}(x)}{x} dx \right] dx. \end{aligned} \quad (17)$$

$$\begin{aligned} \frac{\partial AMISE}{\partial \mathfrak{h}} &= \mathfrak{h}^3 \int_0^{+\infty} \left[\left(\int_0^{+\infty} dy \int_y^{+\infty} [f_X(x) + 3xf_X^{(1)}(x) + x^2f_X^{(2)}(x)] f_X(x) dx \right)^2 \right] dx \\ &\quad - \frac{1}{n\mathfrak{h}^2} \int_0^{+\infty} \left[\mathfrak{C}_k \int_0^{+\infty} dy \int_y^{+\infty} \frac{f_X^{(3)}(x)}{x} dx \right] dx. \end{aligned}$$

Now,

$$\begin{aligned} \frac{\partial AMISE}{\partial \mathfrak{h}} &= 0 \implies \\ \mathfrak{h}^5 &= \frac{\frac{1}{n} \int_0^{+\infty} \left[\mathfrak{C}_k \int_0^{+\infty} dy \int_y^{+\infty} \frac{f_X^{(3)}(x)}{x} dx \right] dx}{\int_0^{+\infty} \left[\left(\int_0^{+\infty} dy \int_y^{+\infty} [f_X(x) + 3xf_X^{(1)}(x) + x^2f_X^{(2)}(x)] f_X(x) dx \right)^2 \right] dx}. \end{aligned}$$

Therefore,

$$\mathfrak{h} = \left(\frac{\frac{1}{n} \int_0^{+\infty} \left[\mathfrak{C}_k \int_0^{+\infty} dy \int_y^{+\infty} \frac{f_X^{(3)}(x)}{x} dx \right] dx}{\int_0^{+\infty} \left[\left(\int_0^{+\infty} dy \int_y^{+\infty} \left[f_X(x) + 3x f_X^{(1)}(x) + x^2 f_X^{(2)}(x) \right] f_X(x) dx \right)^2 \right] dx} \right)^{\frac{1}{5}} + n^{-\frac{1}{5}}$$

$$= O(n^{-\frac{1}{5}}).$$

3. Empirical Estimation Of Weighted Extropy

Non parametric estimation is a widely employed technique in various research papers for estimating extropy and its associated measures. One common approach within non parametric estimation is the use of kernel density estimation, which is a popular method in the literature in order to obtain smoothed estimates.

In this section, we introduce the empirical method for estimating pdf to assess WE . This estimation is achieved through the utilization of non parametric kernel estimator. The empirical kernel smoothed estimator for WE is,

$$\begin{aligned} \hat{\mathfrak{J}}_{n1}^w(X) &= -\frac{1}{2} \int_0^{+\infty} x \hat{f}_n^2(x) dx \\ &= -\frac{1}{2} \sum_{i=1}^{n-1} \int_{X_{i:n}}^{X_{i+1:n}} x \hat{f}_n^2(x) dx \\ &= -\frac{1}{2} \sum_{i=1}^{n-1} \frac{X_{i:n}^2 - X_{i+1:n}^2}{2} \hat{f}_n^2(X_{i:n}) \\ &= -\frac{1}{4} \sum_{i=1}^{n-1} (X_{i:n}^2 - X_{i+1:n}^2) \hat{f}_n^2(X_{i:n}), \end{aligned}$$

where $f_n(\cdot)$ is the kernel estimator given by [22] and $X_{i:n}$ is the i^{th} order statistic of the random sample.

Example 1. Let the samples be from the distribution with $pdf = 2x, 0 < x < 1$. Then X^2 follows standard uniform distribution. Moreover, Z_{i+1} is a beta distribution with mean and variance, respectively, $\frac{1}{2(n+1)}$ and $\frac{n}{4(n+1)^2(n+2)}$. Then, the mean and variance of $\hat{\mathfrak{J}}_{n1}^w(X)$ is given by

$$E(\hat{\mathfrak{J}}_{n1}^w(X)) = -\frac{1}{4(n+1)} \sum_{i=1}^{n-1} \hat{f}_n^2(X_{i:n}) \quad (18)$$

and,

$$V(\hat{\mathfrak{J}}_{n1}^w(X)) = \frac{n}{16(n+1)^2(n+2)} \sum_{i=1}^{n-1} \hat{f}_n^4(X_{i:n}),$$

where $\hat{f}_n(\cdot)$ is defined in equation (6).

Table 1. Mean and Variance of $\hat{\mathfrak{J}}_{n1}^w(X)$ for the distribution with $pdf = 2x, 0 < x < 1$.

n	Mean	Variance
10	-0.58602	0.03039
20	-0.57820	0.02182
30	-0.55622	0.01888
40	-0.56181	0.01083
50	-0.52369	0.00777
100	-0.52239	0.00597

Table 1 shows the values of mean and variance of the samples of Example 1. Hence it is clear that the mean is approaching towards the theoretical mean -0.5 and the variance is tending to zero when the sample size increases.

Example 2. Suppose X is a Rayleigh distribution with parameter 1. Then X^2 follows exponential distribution and $Z_{i+1} = \frac{X_{i:n}^2 - X_{i+1:n}^2}{2}$ is distributed as exponential distribution with mean $= \frac{1}{2(n-j)}$, for $j = 1, 2, \dots, n$. The mean and variance of $\hat{\mathfrak{J}}_{n1}^w(X)$ are

$$E(\hat{\mathfrak{J}}_{n1}^w(X)) = -\frac{1}{4} \sum_{i=1}^{n-1} \frac{\hat{f}_n^2(X_{i:n})}{n-j}. \quad (19)$$

$$V(\hat{\mathfrak{J}}_{n1}^w(X)) = \frac{1}{16} \sum_{i=1}^{n-1} \frac{\hat{f}_n^4(X_{i:n})}{(n-j)^2}. \quad (20)$$

Table 2. Mean and Variance of $\hat{\mathfrak{J}}_{n1}^w(X)$ for Rayleigh distribution with parameter=1.

n	Mean	Variance
10	-0.46220	0.02453
20	-0.33956	0.01626
30	-0.29213	0.00173
40	-0.28919	0.00121
50	-0.26677	0.00094
100	-0.23975	0.00024

From the table 2 it is clear that the variance is decreasing to zero and the mean is advanced towards the theoretical value -0.25 in the case of Rayleigh distribution with parameter one.

Remark 1. Based on the examples 1 and 2, it is clear that the proposed estimator is consistent (see, table 1 and 2), since the mean of the estimator is approaching towards the theoretical value and the variance tends to zero when the sample size increases.

4. Simulation Study

We manage a simulation study to evaluate the performance of the presented estimators. Random samples are generated corresponding to different sample sizes from some standard distributions, then both bias and MSE are calculated for 10000 samples.

It is obvious that the standard kernel estimation method performs better in many situations, but sometimes log kernel estimation methods outperforms it. So here to enable a comparison between log kernel and kernel estimators of weighted extropy we again proposed a kernel estimator for WE using equation (6). The estimator is given by

$$\begin{aligned} \hat{\mathfrak{J}}_{nk}^w(X) &= -\frac{1}{2} \int_0^{+\infty} x \hat{f}_n^2(x) dx \\ &= -\frac{1}{2} \int_0^{+\infty} dy \int_y^{+\infty} \hat{f}_n^2(x) dx, \end{aligned} \quad (21)$$

where $\hat{f}_n(x)$ is the kernel estimator given by [22]. Using the consistency property of kernel estimator it is clear that the proposed kernel estimator for WE is also consistent. To lay the ground work for comparison, we had generated samples from exponential distribution, log normal distribution-a heavy tailed distribution and uniform distribution. The Gaussian log transformed kernel and Gaussian kernel are the kernel functions used for simulation.

Table 3. Estimated value, $|bias|$ and MSE of $\hat{\mathfrak{J}}_n^w(X)$ and $\hat{\mathfrak{J}}_{nk}^w(X)$ from standard exponential distribution with $\mathfrak{J}^w(X) = -0.125$

$\hat{\mathfrak{J}}_n^w(X)$				$\hat{\mathfrak{J}}_{nk}^w(X)$		
n	H	$ bias $	MSE	H	$ bias $	MSE
50	-0.11909	0.00591	0.00031	-0.13364	0.00864	0.00030
100	-0.11830	0.00670	0.00017	-0.12979	0.00479	0.00013
150	-0.11886	0.00614	0.00011	-0.12872	0.00372	0.00009
200	-0.11904	0.00596	0.00009	-0.12806	0.00306	0.00006
250	-0.11978	0.00522	0.00007	-0.12741	0.00241	0.00005
300	-0.11952	0.00548	0.00007	-0.12720	0.00220	0.00004
350	-0.11975	0.00525	0.00006	-0.12728	0.00228	0.00004
400	-0.12025	0.00475	0.00005	-0.12679	0.00179	0.00003
450	-0.12028	0.00472	0.00005	-0.12646	0.00146	0.00003
500	-0.12061	0.00439	0.00004	-0.12648	0.00148	0.00002

Table 4. Estimated value, $|bias|$ and MSE of $\hat{\mathfrak{J}}_n^w(X)$ and $\hat{\mathfrak{J}}_{nk}^w(X)$ from lognormal distribution with $\mathfrak{J}^w(X) = -0.14105$.

$\hat{\mathfrak{J}}_n^w(X)$				$\hat{\mathfrak{J}}_{nk}^w(X)$		
n	H	$ bias $	MSE	H	$ bias $	MSE
50	-0.14370	0.00265	0.00025	-0.14621	0.00517	0.00025
100	-0.14243	0.00139	0.00012	-0.14375	0.00271	0.00013
150	-0.14189	0.00084	0.00008	-0.14241	0.00136	0.00008
200	-0.14199	0.00095	0.00006	-0.14207	0.00103	0.00006
250	-0.14175	0.00070	0.00005	-0.14204	0.00099	0.00006
300	-0.14171	0.00067	0.00004	-0.14180	0.00075	0.00005
350	-0.14155	0.00051	0.00003	-0.14169	0.00064	0.00004
400	-0.14140	0.00036	0.00003	-0.14158	0.00053	0.00004
450	-0.14127	0.00022	0.00003	-0.14153	0.00048	0.00004
500	-0.14121	0.00016	0.00001	-0.14137	0.00032	0.00004

Table 5. Estimated value, $|bias|$ and MSE of $\hat{\mathfrak{J}}_n^w(X)$ and $\hat{\mathfrak{J}}_{nk}^w(X)$ from standard uniform distribution with $\mathfrak{J}^w(X) = -0.25$.

$\hat{\mathfrak{J}}_n^w(X)$				$\hat{\mathfrak{J}}_{nk}^w(X)$		
n	H	$ bias $	MSE	H	$ bias $	MSE
50	-0.2097	0.0403	0.00286	-0.22576	0.02424	0.0017
100	-0.21562	0.03438	0.00184	-0.22786	0.02214	0.00100
150	-0.21826	0.03174	0.00143	-0.22829	0.02171	0.00080
200	-0.22059	0.02941	0.00120	-0.23056	0.01955	0.00067
250	-0.22285	0.02715	0.00099	-0.23045	0.01944	0.00066
300	-0.22277	0.02723	0.00097	-0.23201	0.01799	0.00053
350	-0.22426	0.02574	0.00088	-0.23276	0.01724	0.00047
400	-0.22511	0.02489	0.00079	-0.23325	0.01675	0.00044
450	-0.22575	0.02425	0.00076	-0.23399	0.01621	0.00039
500	-0.22668	0.02332	0.00068	-0.23379	0.01601	0.00039

Table 6. Estimated value, $|bias|$ and MSE of $\hat{\mathfrak{J}}_{n1}^w(X)$ from standard exponential distribution with $\mathfrak{J}^w(X) = -0.125$

$\hat{\mathfrak{J}}_{n1}^w(X)$			
n	H	$ bias $	MSE
50	-0.16379	0.03879	0.00256
100	-0.14418	0.01918	0.00052
150	-0.13857	0.01357	0.00027
200	-0.13496	0.00996	0.00015
250	-0.13285	0.00785	0.00010
300	-0.13234	0.00734	0.00009
350	-0.13111	0.00611	0.00007
400	-0.13043	0.00543	0.00006
450	-0.13005	0.00505	0.00005
500	-0.12976	0.00476	0.00005

Table 7. Estimated value, $|bias|$ and MSE of $\hat{\mathfrak{J}}_{n1}^w(X)$ from lognormal distribution with $\mathfrak{J}^w(X) = -0.14105$.

$\hat{\mathfrak{J}}_{n1}^w(X)$			
n	H	$ bias $	MSE
50	-0.22300	0.08195	0.06302
100	-0.17574	0.03469	0.00255
150	-0.16491	0.02386	0.00108
200	-0.15942	0.01837	0.00075
250	-0.15744	0.01639	0.00070
300	-0.15401	0.01296	0.00040
350	-0.15218	0.01113	0.00025
400	-0.15072	0.00967	0.00020
450	-0.15015	0.00911	0.00021
500	-0.14888	0.00783	0.00014

Table 8. Estimated value, $|bias|$ and MSE of $\hat{\mathfrak{J}}_{n1}^w(X)$ from standard uniform distribution with $\mathfrak{J}^w(X) = -0.25$.

$\hat{\mathfrak{J}}_{n1}^w(X)$			
n	H	$ bias $	MSE
50	-0.22669	0.02331	0.00172
100	-0.22713	0.02287	0.00107
150	-0.22892	0.02108	0.00085
200	-0.23084	0.01916	0.00066
250	-0.23194	0.01806	0.00056
300	-0.23168	0.01832	0.00055
350	-0.23195	0.01805	0.00049
400	-0.23317	0.01683	0.00045
450	-0.23335	0.01665	0.00041
500	-0.23361	0.01639	0.00039

From the above tables 3, 4, 5, 7 and 8, it is clear that the MSE and bias of both the estimators are decreasing with sample size. The decreasing MSE indicates that the estimator's predictions are getting closer to the true values with larger sample sizes, demonstrating enhanced accuracy and efficiency in estimation. The decreasing bias also shows the accuracy of the estimators.

The comparison of bias and MSE between kernel and logkernel estimators in simulation for WE reveals that the logkernel estimator outperforms the kernel estimator in certain scenarios, particularly when dealing with heavy-tailed distributions.

5. Data Analysis

5.1. Data 1

The comparison between log kernel and kernel estimators of WE has been demonstrated using the data given in [16]. The data demonstrates the quantity of thousands of cycles to failure for electrical appliances in a life test. The kurtosis of the data is found to be -0.918. We have fitted exponential distribution with parameter 0.640 to the data. The p-value obtained for Kolmogorov-Smirnov test (0.124) is 0.390, which reveals that exponential distribution is a good fit to the data. The estimate obtained using maximum likelihood estimation is -0.141.

The estimate of WE earned using log kernel and kernel estimation are $\hat{J}_n^w(X)=-0.127$, $\hat{J}_{nk}^w(X)=-0.144$ and $\hat{J}_{n1}^w(X)=-0.148$. Hence from closeness of estimates to maximum likelihood estimate of WE it is clear that the kernel estimators of corresponding function outperforms the log kernel estimator in this situation.

5.2. Data 2 (Heavy Tailed Data)

To illustrate the comparison of the proposed log kernel and kernel estimators using real data, we consider the data from [17]. The data represents the remission times(Months) of 137 cancer patients. A kurtosis value of 15.195 is obtained: it is exceptionally high and suggests a very heavy-tailed or leptokurtic distribution. Hence log normal distribution is fitted to the data and the parameters obtained are

$$\hat{\mu} = 1.756 \qquad \hat{\sigma} = 1.066.$$

Using Kolmogorov-Smirnov test with statistic as 0.06 and p value as 0.591, it supports that log normal distribution is a best fit to the data.

The estimates of WE using the proposed estimators and by maximum likelihood estimation are calculated for this data. We have obtained $\hat{J}_n^w(X)=-0.1346$, $\hat{J}_{nk}^w(X)=-0.141$. The estimate of WE using maximum likelihood estimation is secured as -0.1323, which signifies that log kernel estimator of WE performs better than the weighted extropy estimated with standard kernel estimation methods when dealing with heavy tailed data.

5.3. Data 3 (The Time Until Failure Of Three Systems)

The data is obtained from [25]. The observations represents the working time of 2000 hours, which starts from 0 for 3 systems. Table 9 shows the value of suggested estimators of WE for these systems.

Table 9. Values of $\hat{J}_n^w(X)$, $\hat{J}_{nk}^w(X)$ and $\hat{J}_{n1}^w(X)$ for the three systems

	System 1	System 2	System 3
$\hat{J}_n^w(X)$	-0.10526	-0.08201	-0.11902
$\hat{J}_{nk}^w(X)$	-0.13655	-0.15442	-0.17181
$\hat{J}_{n1}^w(X)$	-0.75986	-0.22782	-0.16940

According to [10] the system or component which is said to have less uncertainty is more reliable. In accordance with this concept we can infer that system 2 is more reliable than system 1 and system 3 in terms of $\hat{J}_n^w(X)$, while system 1 is more reliable with regard to $\hat{J}_{nk}^w(X)$ and system 3 is more reliable when using $\hat{J}_{n1}^w(X)$. This example vividly demonstrates how the estimation of WE is useful in choosing a reliable system among the several available competing models.

6. Conclusions

In this article, we have considered non parametric estimation of WE . A log kernel estimator and empirical kernel smoothed estimator for WE have been depicted. The bias, variance, optimal bandwidth and some properties of the log kernel estimator of the function have been also established

here. A kernel estimator has been also proposed to enable a comparison with log kernel estimator. From the simulation study and data analysis we have concluded that log kernel estimator of WE is performed better when using data sets suitable for heavy-tailed distributions. The real data analysis have also involved an assessment of the estimators performance and its utility in reliability modeling. Since log kernel estimator of WE performs better in the case of heavy tailed, skewed, multimodal distributions, the same can also be used in the case of income distributions to provide new features to the income data.

Author Contributions: All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Acknowledgments: Maria Longobardi is partially supported by the GNAMPA research group of INDAM (Istituto Nazionale di Alta Matematica) and MIUR-PRIN 2022 PNRR Statistical Mechanics of Learning Machines: from algorithmic and information-theoretical limits to new biologically inspired paradigms

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Balakrishnan, N., Buono, F., Longobardi, M. (2020). On Tsallis extropy with an application to pattern recognition. *Statistics and Probability Letters*, 180.
2. Balakrishnan, N., Buono, F., Longobardi, M. (2022). On weighted extropies. *Communication in Statistics-Theory and Methods*,
3. Becerra, A., de la Rosa, J.I., Gonzalez, E., Pedroza, A.D., Escalante, N.I. (2018). Training deep neural networks with non-uniform frame-level cost function for automatic speech recognition. *Multimed. Tools Appl.*, 77, 27231-27267.
4. Belis, M. Guiasu, S.(1968). A quantitative-qualitative measure of information in cybernatic systems. *IEEE Transactions on Information Theory*, 14(4), 593-594.
5. Buono, F., Deng, Y., Longobardi, M.(2023a). The unified extropy and its versions in classical and Dempster-Shafer theories. *Journal of Applied Probability*. DOI 10.1017/jpr.2023.68.
6. Buono, F., Kamari, O., Longobardi, M.(2023b). Interval extropy and weighted interval extropy. *Ricerche di Matematica*, 72 (1), 283-298.
7. Buono, F., Longobardi, M.(2020). A dual measure of uncertainty: The Deng extropy. *Entropy*, 22 (5).
8. Charpentier, A. and Flachaire, E. (2015). Log-transform kernel density estimation of income distribution. *L'Actualite economique, Revue d' Analyse Economique*, 91, 141-159.
9. Di Crescenzo, A., Longobardi, M. (2006). On weighted residual and past entropies, *Sci. Math. Jpn.*, 64, 255–266.
10. Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. *Sankhya: The Indian Journal of Statistics, Series A* 58 (1), 48–56.
11. Irshad, M.R., Maya, R., (2022). Non-parametric estimation of the past extropy under α -mixing dependence condition. *Ric. Mat.*, 71, 723-734.
12. Irshad, M.R., Maya, R., (2022). Non-parametric log kernel estimation of extropy function. *Chil. J. Stat.*, 13, 155–163.
13. Kazemi, R., Hashempour, M., Longobardi, M., (2022). Weighted Cumulative Past Extropy and Its Inference. *Entropy*, 24 (10).
14. Kazemi, R., Tahmasebi, S., Calì, C., Longobardi, M., (2021) Cumulative residual extropy of minimum ranked set sampling with unequal samples. *Results in Applied Mathematics*, 10.
15. Lad, F., Sanfilippo, G., Agro, G. (2015). Extropy: Complementary dual of entropy. *Statistical Science*, 30(1), 40-58.
16. Lawless, J.F. (2011). *Statistical Models and Methods for Lifetime Data* (Vol. 362), Hoboken: Wiley.
17. Lee, E. T, Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis*, Wiley and Sons, Third edition.
18. Maya, R., Irshad, M.R. (2019). Kernel estimation of the residual extropy under α -mixing dependence condition. *S. Afr. Stat. J.*, 53, 65-72.

19. Maya, R., Irshad, M.R., Archana, K. (2021). Recursive and non-recursive kernel estimation of negative cumulative residual extropy under α -mixing dependence condition. *Ric. Mat.*, 55, 1-21.
20. Maya, R., Irshad, M.R., Bakouch, H., Krishnakumar, A., Qarmalah, N. (2023). Kernel Estimation of the Extropy Function under α -Mixing Dependent Data. *Symmetry*, 15(4), 796. <https://doi.org/10.3390/sym15040796>.
21. Nguyen, H.D., Jones, A.T., and Mclachlan, G.J. (2018). Positive data kernel density estimation via the LogKDE Package for R. *16th Australasian Conference, AusDM 2018, Bahrurst, NSW, Australia*.
22. Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33, 1065-1076.
23. Qiu, G. (2017). Extropy of order statistics and record values. *Statistics and Probability Letters*, 120, 52-60.
24. Qiu, G., Jia, K. (2018). Extropy estimators with applications in testing uniformity. *J. Nonparametric Stat.*, 30, 182-196.
25. Rai, R. N., Chaturvedi, S. K., Bolia, N. (2020). Repairable systems reliability analysis: A comprehensive framework. *New Jersey: John Wiley and Sons*.
26. Rajesh, R., Rajesh, G., Sunoj, S. (2022). Kernel estimation of extropy function under length-biased sampling. *Stat. Probab. Lett.*, 181, 109290.
27. Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 27, 832-837.
28. Sathar, E.A., Nair, R.D. (2021). On dynamic weighted extropy. *Journal of Computational and Applied Mathematics*, 393, 113507.
29. Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423.
30. Wegman, E.J. (1972). Nonparametric probability density estimation: I. A summary of available methods. *Technometrics*, 14, 533-546.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.