Probability in the Engineering and Informational Sciences (2025), **39:**4 461–485
doi:10.1017/S0269964824000299

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Extropy-based dynamic cumulative residual inaccuracy measure: properties and applications

M Hashempour[1], M Mohammadi[2], S M A Jahanshahi[3] and A H Khammar[4]

[1]Department of Statistics, University of Hormozgan, Bandar Abbas, Hormozgan, Iran
[2]Department of Statistics, University of Zabol, Zabol, Sistan and Baluchestan, Iran
[3]Department of Statistics, University of Sistan and Baluchestan, Zahedan, Sistan and Baluchestan, Iran
[4]Department of Statistics, University of Birjand, Birjand, South Khorasan, Iran
Corresponding author: Morteza Mohammadi; Email: mo.mohammadi@uoz.ac.ir

Keywords: Cumulative residual extropy, Inaccuracy, Proportional hazard rate model, Characterization, Residual lifetime

MSC: 62N05; 62F10; 62E10; 94A17

Abstract

The cumulative residual extropy has been proposed recently as an alternative measure of extropy to the cumulative distribution function of a random variable. In this paper, the concept of cumulative residual extropy has been extended to cumulative residual extropy inaccuracy (CREI) and dynamic cumulative residual extropy inaccuracy (DCREI). Some lower and upper bounds for these measures are provided. A characterization problem for the DCREI measure under the proportional hazard rate model is studied. Nonparametric estimators for CREI and DCREI measures based on kernel and empirical methods are suggested. Also, a simulation study is presented to evaluate the performance of the suggested measures. Simulation results show that the kernel-based estimator performs better than the empirical-based estimator. Finally, applications of the DCREI measure for model selection are provided using two real data sets.

1. Introduction

Physicist Boltzmann [21] initially proposed the concept of entropy to characterize the level of disorder in a physical system. Shannon [36] further developed this concept, which has wide-ranging applications in statistics, information theory, probability, and other fields. Entropy measures the uncertainty associated with a random variable, as detailed by Cover and Thomas [6].

Cumulative residual entropy (CRE), introduced by Rao *et al.* [34], uses the survival function (SF) instead of the probability density function (PDF) in Shannon entropy, which provides a more organized measure of uncertainty over the remaining lifetime of a system:

$$\xi(F) = -\int_0^{+\infty} \bar{F}(x) \log \bar{F}(x) dx. \tag{1}$$

The SF is the negative derivative of the PDF, making it a more stable measure. Various aspects of CRE have been studied by Rao [33] and Navarro *et al.* [29].

In this regard, Kerridge [14] introduced a measure of inaccuracy, which quantifies differences between statistical models. The Kerridge inaccuracy measure differs from Shannon entropy by providing a way to assess the accuracy of statistical models. Properties of inaccuracy measures have been studied in coding theory (Nath [28] and Bhatia [4]), and further developments can be found in the studies by Kundu and Nanda [18], Kundu *et al.*[17], Kayal *et al.*[13], and Psarrakos and Di Crescenzo [32].

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Kumar and Taneja [16] defined the cumulative residual inaccuracy measure of two random variables X and Y as

$$\xi(F,G) = -\int_0^{+\infty} \bar{F}(x)\log\bar{G}(x)dx. \tag{2}$$

When $\bar{G}(x) = \bar{F}(x)$, this simplifies to equation (1), suggesting that the cumulative residual inaccuracy measure is a broader version of the cumulative residual measure.

In dynamic contexts, such as continuously changing information measures, equations (1) and (2) may not be suitable. Kumar and Taneja [16] extended the cumulative residual inaccuracy to a dynamic form based on Asadi and Zohrevand's [3] dynamic measure of cumulative residual entropy:

$$\xi(F,G;t) = -\int_t^{+\infty} \frac{\bar{F}(x)}{\bar{F}(t)}\log\frac{\bar{G}(x)}{\bar{G}(t)}dx, \qquad t \geq 0.$$

Moreover, $\xi(F,G;t)$ tends to $\xi(F,G)$ as $t \to 0$.

Lad *et al.* [19] proposed the extropy measure as an alternative to quantify uncertainty. Extropy is simpler to compute and has been explored in goodness-of-fit tests and inference methods. Yang *et al.* [37] studied the relationship between extropy and variational distance, identifying distributions with minimum or maximum extropy within a specified variational distance.

Jahanshahi *et al.* [12] introduced "cumulative residual extropy" (CRE) to measure the uncertainty of a non-negative continuous random variable X:

$$\xi J(F) = -\frac{1}{2}\int_0^{+\infty} \bar{F}^2(x)dx. \tag{3}$$

Abdul Sathar and Nair [1] suggested a dynamic version of CRE named dynamic survival extropy:

$$\xi J(F;t) = -\frac{1}{2\bar{F}^2(t)}\int_t^{+\infty} \bar{F}^2(x)dx, \qquad t \geq 0. \tag{4}$$

If the DCRE of variable X is less than that of Y, X is more uncertain than Y.

Hashempour and Mohammadi [9] introduced cumulative past extropy inaccuracy (CPEI) and dynamic cumulative past extropy inaccuracy (DCPEI):

$$\bar{\xi}J(F;G) = -\frac{1}{2}\int_0^{+\infty} F(x)G(x)dx,$$
$$\bar{\xi}J(F,G;t) = -\frac{1}{2}\int_0^{t} \frac{F(x)G(x)}{F(t)G(t)}dx.$$

They discussed topics such as characterization and stochastic ordering for these measures. Recent generalizations of these measures have been studied in several works [9, 10, 26]. For further reading on the relation between extropy and inaccuracy, see [11, 23–25].

In this paper, we introduce two novel metrics for assessing uncertainty, referred to as cumulative residual extropy inaccuracy (CREI) and dynamic cumulative residual extropy inaccuracy (DCREI). The fundamental concept involves substituting the PDF with the SF in extropy calculations. The SF is deemed to be more consistent than the PDF since the PDF is derived from the SF. CREI and DCREI are extensions of CRE measures. While CRE focuses on the uncertainty and information content over the remaining lifetime of a system, CREI and DCREI provide a refined view by capturing the accuracy and dynamics of these measures over time. This distinction is crucial in applications where the evolution of inaccuracy is as important as its static measure. CRE measures the uncertainty remaining in the lifetime of a system, whereas CREI and DCREI extend this concept to address the inaccuracies and dynamic

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

aspects over time. This temporal aspect is particularly valuable in reliability modeling, where the conditions and performance of systems can evolve. For instance, in industrial maintenance, equipment may undergo changes due to wear and repairs and a static measure might not fully capture these dynamics.

CREI and DCREI have solid theoretical foundations, validated through simulations and real data applications, demonstrating superior performance. By leveraging kernel-based and empirical methods for nonparametric estimation, these measures offer flexibility across distributions, which is essential for practical applications. Nonparametric methods, such as kernel-based estimators, do not rely on strict parametric assumptions, making them suitable for a wide range of data types. Additionally, DCREI's dynamic evaluation of inaccuracy at specific time points provides valuable insights into model performance. This is crucial for applications like industrial maintenance, where understanding how a model's accuracy changes over time can lead to better predictions of failure times, maintenance schedules, and risk assessments. Unlike static measures like Akaike information criterion (AIC) or Bayesian information criterion (BIC), which provide a single value, DCREI captures the variations in fit across different intervals, allowing for a more detailed assessment of a model's performance.

Additionally, by incorporating nonparametric estimators, CREI and DCREI allow for flexible modeling of complex data without strict assumptions. This flexibility is particularly important in real-world scenarios where data distributions may not follow conventional parametric forms. For example, in supply chain management, selecting the appropriate distribution impacts inventory policies, lead times, and service levels. CREI and DCREI can provide a more accurate and dynamic understanding of these distributions, leading to better decision-making and resource allocation. So, the introduction of CREI and DCREI addresses the need for dynamic and detailed measures of inaccuracy in reliability modeling and information theory. By providing a more comprehensive view of inaccuracy over time, these measures enable better model selection, improved predictions, and more informed decision-making.

This paper is organized into several sections to thoroughly address the proposed measures and their applications. In Section 2, the definitions of CREI and DCREI are provided. Additionally, alternative formulations of DCREI are proposed in this section, and various bounds based on DCREI are derived. In Section 3, we prove that the DCREI uniquely characterizes three specific lifetime distributions. Nonparametric estimators for the CREI and DCREI measures based on kernel and empirical methods are proposed in Section 4. In Section 5, a simulation study is presented to measure their accuracy. Finally, in Section 6, we present a real-life application of DCREI to find the best-fitted distribution from data.

2. DCREI on Results

In this section, we will discuss a novel measure of inaccuracy called dynamic cumulative residual extropy inaccuracy (DCREI). This measure applies to two continuous random variables with nonnegative values and the same range. We will also examine various properties of DCREI.

First, let us define the CREI measure by analogy to the CPEI as follows.

Definition 2.1. *Assume that $\bar{F}(x)$ and $\bar{G}(x)$ are SFs of non-negative continuous random variables X and Y, respectively. The CREI between X and Y is defined as*

$$\xi J(F,G) = -\frac{1}{2}\int_0^{+\infty} \bar{F}(x)\bar{G}(x)dx. \tag{5}$$

Remark 2.2. From (5), it holds $\xi J(F,G) \leq 0$ and $\xi J(F,G) = \xi J(G,F)$. Moreover, when two random variables X and Y have same SFs, the CREI reduces to the CRE given in (3). Indeed, the closer the value of CREI is to CRE, the better Y is an approximation of X (see Example 2.4).

Proposition 2.3. *Let X and Y be two non-negative continuous random variables with SFs $\bar{F}(x)$ and $\bar{G}(x)$, respectively. Then*

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

$\xi J(F,F) > (<)\ \xi J(F,G)$ *if* $\bar{F}(x) < (>)\ \bar{G}(x)$.

In the following remark, we illustrate an application of CREI for comparing statistical models.

Remark 2.4. Let the SF $\bar{F}(x) = (1-x)$, $0 < x < 1$, be the true statistical model for a random variable X that generated some data. Also, suppose $\bar{G}(x) = (1-x^2)$ and $\bar{I}(x) = (1-x^3)$, $0 < x < 1$, are two power SFs determined through nonparametric statistical tests to approximation X. From equation (5), we obtain $\xi J(F,F) = \xi J(F) = -0.167$, $\xi J(F,G) = -0.208$, and $\xi J(F,I) = -0.225$. Thus, the CREI between X and a random variable Y that follows the SF $\bar{G}(x)$ is closer to the CRE of X than the CREI between X and a random variable Z that follows the SF $\bar{I}(x)$. Therefore, Y provides a better approximation to X than Z, that is, the statistical model $\bar{G}(x)$ is closer than $\bar{I}(x)$ to the statistical model $\bar{F}(x)$ that generates data.

In life testing experiments, it is common for the experimenter to have knowledge about the current age of the system being studied. However, the existing CREI measure presented in (5) is not appropriate for such situations. Therefore, it needs to be revised in order to incorporate the current age of the system as well. In what follows, if X is the lifetime of a component that has already survived upto time t, then the random variable $X_t = [X - t|X > t]$ called the residual lifetime random variable has the SF

$$\bar{F}_t(x) = \begin{cases} \frac{\bar{F}(x)}{\bar{F}(t)} & \text{if } x > t, \\ 1 & \text{if } x \le t, \end{cases}$$

and similarly for $Y_t = [Y - t|Y > t]$. Thus, by analogy to the DCPEI, we define a dynamic version of the CREI between two random variables X and Y as follows.

Definition 2.5. *Let* $\bar{F}(x)$ *and* $\bar{G}(x)$ *be SFs of lifetime random variables X and Y, respectively. The DCREI between X and Y is defined as*

$$\xi J(F,G;t) = -\frac{1}{2}\int_0^{+\infty} \bar{F}_t(x)\bar{G}_t(x)dx = -\frac{1}{2}\int_t^{+\infty} \frac{\bar{F}(x)\bar{G}(x)}{\bar{F}(t)\bar{G}(t)}dx,\ x \ge 0,\ t \ge 0, \tag{6}$$

where $\bar{F}_t(x) = P(X - t > x|X > t) = \frac{\bar{F}(t+x)}{\bar{F}(t)}$ *with* $\bar{F}(t) > 0$ *and* $\bar{G}(t) > 0$.

Example 2.6. Suppose that X and Y have exponential distributions with SFs as follows:

$$\bar{F}(t) = e^{-\theta t}, \qquad t \ge 0,\ \theta > 0,$$
$$\bar{G}(t) = e^{-\lambda t}, \qquad t \ge 0,\ \lambda > 0.$$

From equation (5), we have

$$\xi J(F,G) = -\frac{1}{2(\theta+\lambda)}.$$

We observe that $\xi J(F,G) = \xi J(F,G;t)$ and so $\xi J(F,G;t)$ remains constant over time, which is expected due to the memoryless property of the exponential distribution. This constancy implies that the inaccuracy measure does not change with t, reinforcing the idea that the characteristics of the exponential distribution are time-invariant.

Functions $\xi J(F,G;t)$ and $\xi J(F,G)$ are shown in Figure 1 for some selected values of θ and λ. From the left panel, we observe that $\xi J(F,G)$ is a nondecreasing function with respect to θ. As θ increases,

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

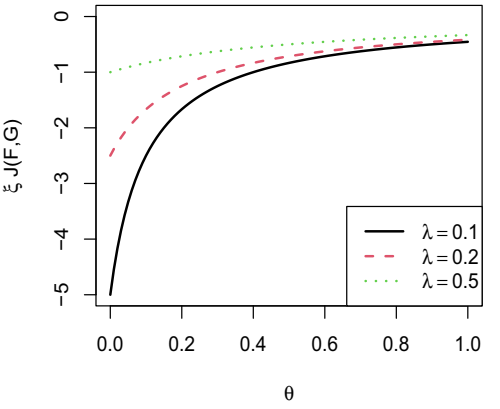


Figure 1. *Graphs of $\xi J(F,G)$ for Example 2.6.*

the value of $\xi J(F,G)$ becomes less negative, indicating that the inaccuracy between $F(t)$ and $G(t)$ decreases with larger θ. Similarly, higher values of λ also result in smaller (less negative) values of $\xi J(F,G)$, reflecting a decrease in inaccuracy as λ increases. This behavior aligns with the intuition that as the rates of decay (represented by θ and λ) increase, the survival functions become more aligned, reducing inaccuracy.

Example 2.7. Suppose that X and Y are two non-negative continuous random variables with SFs $\bar{F}(t) = e^{-\theta t^2};\ t \geq 0,\ \theta > 0$, and $\bar{G}(t) = e^{-\lambda t^2};\ t \geq 0,\ \lambda > 0$, respectively. From equation (5), we obtain

$$\xi J(F,G) = -\frac{\sqrt{\pi}}{4\sqrt{\theta+\lambda}}.$$

Also, from equation (6), we have

$$\xi J(F,G;t) = -\frac{\sqrt{\pi}}{4\sqrt{\theta+\lambda}} e^{(\theta+\lambda)t^2} erfc(\sqrt{\theta+\lambda}t),$$

where $erfc(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-x^2} dx$.

The behavior of functions $\xi J(F,G)$ and $\xi J(F,G;t)$ is shown in Figure 2. The left panel of Figure 2 shows the function $\xi J(F,G)$, which represents a static measure of inaccuracy between the SFs $\bar{F}(t)$ and $\bar{G}(t)$ as a function of θ with different values of λ. From the left panel, we observe that $\xi J(F,G)$ is a nondecreasing function of θ. This indicates that as θ increases, the inaccuracy measure $\xi J(F,G)$ becomes less negative, which means that the inaccuracy between the SFs $\bar{F}(t)$ and $\bar{G}(t)$ decreases. The right panel of Figure 2 illustrates the function $\xi J(F,G;t)$, which is a dynamic measure of inaccuracy evaluated at specific time points t. In this panel, $\xi J(F,G;t)$ increases over time, demonstrating that it is a nondecreasing function with respect to t, θ, and λ. This increase reflects the accumulating effect of differences between $F(t)$ and $G(t)$ over time.

Introducing the DCREI as a novel measure of inaccuracy for two continuous random variables with non-negative values and the same range is a significant contribution to the field. This measure provides a new perspective on assessing data accuracy and can offer valuable insights into the relationship between these variables. By examining various properties of DCREI, researchers can gain a deeper understanding of the dynamics and patterns within the data, leading to potential advancements in statistical analysis

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

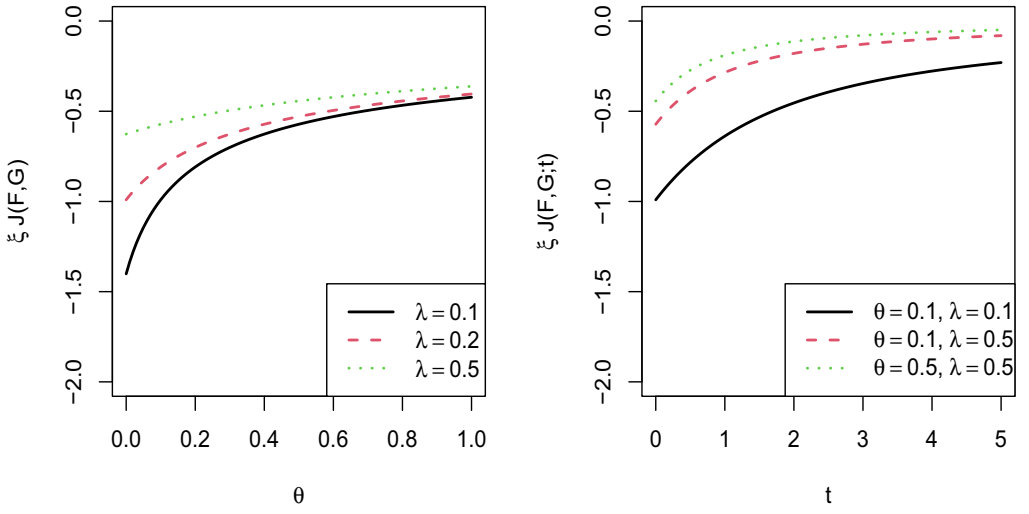


Figure 2. *Graphs of $\xi J(F,G)$ (left panel) and $\xi J(F,G;t)$ (right panel) for Example 2.7.*

and decision-making processes. This research opens up new possibilities for evaluating and interpreting data more nuanced and comprehensively.

Remark 2.8. From equation (6), it holds $\xi J(F,G;t) \leq 0$ and $\xi J(F,G;t) = \xi J(G,F;t)$. Furthermore, on taking limit as $t \to 0$ in equation (6), the DCREI becomes CREI (5). Additionally, when two random variables X and Y have same SFs, the DCREI reduces to the DCRE given in equation (4).

An alternative statement to (6) of the DCREI of a non-negative random variable X is presented herein. Let us consider the following function:

$$\begin{aligned}\psi_F(t,x) &= \int_t^x \bar{F}(z)dz \\ &= \bar{G}(t)\tilde{M}_{F,G}(t) - W_F(x) \\ &= \bar{F}(t)M_F(t) - \bar{F}(x)M_F(x),\end{aligned} \tag{7}$$

where $\tilde{M}_{F,G}(t) = \int_t^{+\infty} \frac{\bar{F}(x)}{\bar{G}(t)}dx$ is an auxiliary function (see the study by Psarrakos and Di Crescenzo [32]), $W_F(x) = \int_x^{+\infty} \bar{F}(z)dz$ is stop-loss transforms and $M_F(t) = \int_t^{+\infty} \frac{\bar{F}(x)}{\bar{F}(t)}dx$ is the mean residual life (MRL) function, which is a worthy tool for modeling and analyzing the data in reliability and survival analysis. It is worth noting that the partial derivative of $\psi_F(t,x)$ with respect to t is closely related to the SF of random variable X.

Theorem 2.9 *Let X and Y be two non-negative continuous random variables with SFs $\bar{F}(x)$ and $\bar{G}(x)$, respectively. Then*

$$\xi J(F,G;t) = -\frac{E[\psi_F(t,X)|Y > t]}{2\bar{F}(t)}. \tag{8}$$

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

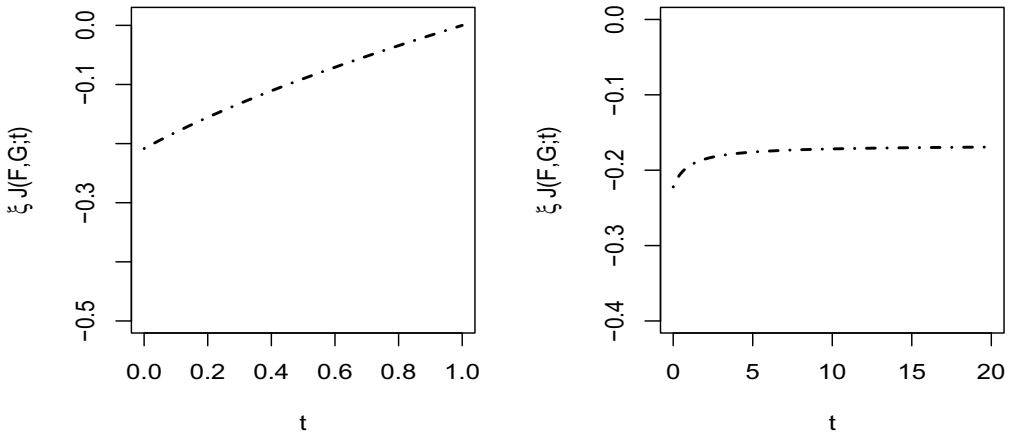


Figure 3. *Graphs of $\xi J(F,G;t)$ for Example 2.10 (left panel) and Example 2.11 (right panel).*

Proof. Using equation (7) and applying Fubini's theorem, we obtain

$$\begin{aligned} E[\psi_F(t,X)|Y>t] &= \int_t^{+\infty} \frac{g(x)}{\bar{G}(t)} \left(\int_t^x \bar{F}(z)dz \right) dx \\ &= \int_t^{+\infty} \bar{F}(z) \left(\int_z^{+\infty} \frac{g(x)}{\bar{G}(t)} dx \right) dz \\ &= \int_t^{+\infty} \bar{F}(t) \left(\frac{\bar{F}(z)}{\bar{F}(t)} \frac{\bar{G}(z)}{\bar{G}(t)} \right) dz \\ &= -2\bar{F}(t)\xi J(F,G;t), \end{aligned}$$

which leads to the outcome mentioned. □

Example 2.10. Let X be a non-negative random variable with SF $\bar{F}(x) = (1-x^2),\ 0<x<1$, and let the random variable Y be uniformly distributed over $(0,1)$ with SF given by $\bar{G}_Y(x) = (1-x),\ 0<x<1$. From both relations (6) and (8), we obtain

$$\xi J(F,G;t) = \frac{(t-1)\,(3t+5)}{24\,(t+1)}, \qquad 0 \le t < 1.$$

Example 2.11. Let X and Y be two non-negative random variables with SFs $\bar{F}(x) = (x+1)e^{-x}$ and $\bar{G}(x) = e^{-2x},\ x>0$, respectively. Substituting these functions in (6) or (8), we obtain

$$\xi J(F,G;t) = -\frac{3t+4}{18(1+t)}.$$

Figure 3 provides the graphs of $\xi J(F,G;t)$ for various values of t in the case where X and Y are random variables with SFs given by Examples 2.10 and 2.11. Note that $\xi J(F,G;t)$ is nondecreasing in Examples 2.10 and 2.11 in terms of t.

In what follows, we provide two lower bounds for the DCREI of non-negative random variables in terms of CREI and MRL functions.

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Proposition 2.12. *Assume that X and Y are two non-negative continuous random variables with SFs* $\bar{F}(x)$ *and* $\bar{G}(x)$*, respectively. Then*

$$\xi J(F,G;t) = \frac{1}{\bar{F}(t)\bar{G}(t)}\left(\frac{1}{2}H_{F,G}(t) + \xi J(F,G)\right)$$

and

$$\xi J(F,G;t) \geq \frac{\xi J(F,G)}{\bar{F}(t)\bar{G}(t)},$$

where $H_{F,G}(t) = \int_0^t \bar{F}(x)\bar{G}(x)dx$.

Proof. From relations (6) and (5), we have

$$\begin{aligned}
\xi J(F,G;t) &= -\frac{1}{2}\int_t^{+\infty} \frac{\bar{F}(x)}{\bar{F}(t)}\frac{\bar{G}(x)}{\bar{G}(t)}dx \\
&= -\frac{1}{2\bar{F}(t)\bar{G}(t)}\int_t^{+\infty} \bar{F}(x)\bar{G}(x)dx \\
&= -\frac{1}{2\bar{F}(t)\bar{G}(t)}\left(\int_0^{+\infty} \bar{F}(x)\bar{G}(x)dx - H_{F,G}(t)\right) \\
&= \frac{1}{\bar{F}(t)\bar{G}(t)}\left(\xi J(F,G) + \frac{1}{2}H_{F,G}(t)\right).
\end{aligned} \tag{9}$$

Since $H_{F,G}(t) \geq 0$, relation (9) gives $\xi J(F,G;t) \geq \frac{\xi J(F,G)}{\bar{F}(t)\bar{G}(t)}$. Thus, the proof is completed. □

Theorem 2.13 *Let X and Y be two non-negative continuous random variables with SFs* $\bar{F}(x)$ *and* $\bar{G}(x)$*, respectively. Then, we have*

$$\xi J(F,G;t) \geq \max\left\{-\frac{M_F(t)}{2}, -\frac{M_G(t)}{2}\right\}.$$

Proof. It is obvious that $\bar{G}(x)$ is nonincreasing in x. Therefore, we have $\frac{\bar{G}(x)}{\bar{G}(t)} \leq 1$, for $x \geq t$. This implies that $\xi J(F,G;t) \geq -\frac{M_F(t)}{2}$. Similar result too holds for Y, that is, $\xi J(F,G;t) \geq -\frac{M_G(t)}{2}$. The proof is completed. □

Example 2.14. Let $X \sim F$ and $Y \sim G$ have Pareto distribution with the same scale parameter σ, so that they have SFs and MRL functions as follows:

$$\bar{F}(t) = \left(\frac{\sigma}{t+\sigma}\right)^{\alpha_1},\ M_F(t) = \frac{t+\sigma}{\alpha_1 - 1}, \qquad t \geq 0, \sigma > 0, \alpha_1 > 1,$$

$$\bar{G}(t) = \left(\frac{\sigma}{t+\sigma}\right)^{\alpha_2},\ M_G(t) = \frac{t+\sigma}{\alpha_2 - 1}, \qquad t \geq 0, \sigma > 0, \alpha_2 > 1.$$

In this case, by using Proposition 2.12, we have

$$\xi J(F,G;t) = -\frac{t+\sigma}{2(\alpha_1+\alpha_2-1)} \geq -\frac{(t+\sigma)^{\alpha_1+\alpha_2}}{2(\alpha_1+\alpha_2-1)} \geq \frac{\xi J(F,G)}{\bar{F}(t)\bar{G}(t)},\ \alpha_1+\alpha_2 > 1, \qquad t \geq 0.$$

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

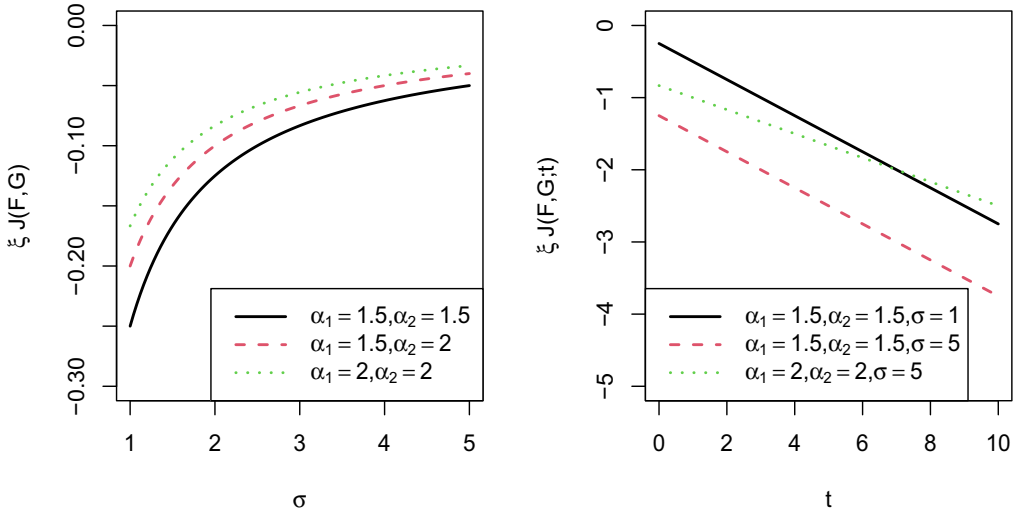


Figure 4. *Graphs of $\xi J(F,G;t)$ for Example 2.14 and various values of α_1, α_2, and σ.*

Moreover, by using Proposition 2.13, we find

$$\begin{aligned}\xi J(F,G;t) &\geq \max\left\{-\frac{t+\sigma}{2(\alpha_1-1)}, -\frac{t+\sigma}{2(\alpha_2-1)}\right\}\\ &= \max\left\{-\frac{M_F(t)}{2}, -\frac{M_G(t)}{2}\right\}, \qquad \alpha_1,\alpha_2 > 1,\ t \geq 0.\end{aligned}$$

The functions $\xi J(F,G)$ and $\xi J(F,G;t)$ are shown in Figure 4 for some selected values of α_1, α_2, and σ. The left panel of Figure 4 shows the function $\xi J(F,G)$ as a function of the scale parameter σ for various values of α_1 and α_2. From the left panel, we observe that $\xi J(F,G)$ is a nondecreasing function of σ. This indicates that as σ increases, the value of $\xi J(F,G)$ becomes less negative, suggesting that the inaccuracy between the SFs $\bar{F}(t)$ and $\bar{G}(t)$ decreases. Additionally, as α_1 and α_2 increase, the function $\xi J(F,G)$ reflects a lower level of inaccuracy. The right panel of Figure 4 illustrates the function $\xi J(F,G;t)$ as a dynamic measure of inaccuracy evaluated over time t for different combinations of α_1, α_2, and σ. In this panel, it is evident that $\xi J(F,G;t)$ does not increase with respect to time t. Instead, $\xi J(F,G;t)$ remains constant or decreases over time, illustrating that the inaccuracy measure is not strictly increasing with time t. This behavior is consistent with the properties of the Pareto distribution and the cumulative residual inaccuracy measure, reflecting the relationship between α_1, α_2, and σ in the context of the dynamic measure.

Example 2.15. Suppose that X and Y follow exponential distribution and Lindley distribution with SFs and MRL functions given by

$$\begin{aligned}&\bar{F}(t) = e^{-\theta t}, \quad M_F(t) = \frac{1}{\theta}, \qquad \theta > 0,\ t \geq 0,\\ &\bar{G}(t) = \frac{(\lambda+1+\lambda t)e^{-\lambda t}}{\lambda+1}, \quad M_G(t) = \frac{\lambda+2+\lambda t}{\lambda(\lambda+1+\lambda t)}, \qquad \lambda > 0,\ t \geq 0.\end{aligned}$$

In this case, due to (6) or (8) and by using Proposition 2.12, we have

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

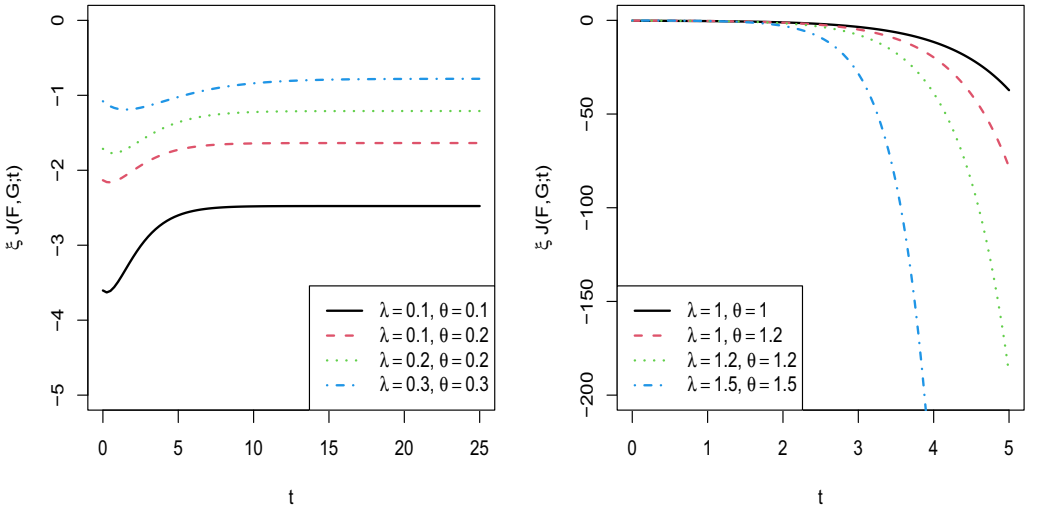


Figure 5. *Graphs of $\xi J(F,G;t)$ for Example 2.15 and various values of λ and θ.*

$$\xi J(F,G;t) = -\frac{\lambda+1+\lambda(\lambda+\theta)e^{t(\lambda+\theta)}\overline{\Gamma}_t(2,\lambda+\theta)}{2(\lambda+\theta)(\lambda+1+\lambda t)}$$
$$\geq -\frac{\lambda+1+\lambda(\lambda+\theta)e^{t(\lambda+\theta)}\Gamma(2,\lambda+\theta)}{2(\lambda+\theta)(\lambda+1+\lambda t)} = \frac{\xi J(F,G)}{\bar{F}(t)\bar{G}(t)}, \qquad \lambda,\theta>0,\ t\geq 0,$$

where $\bar{\Gamma}_x(\alpha,\beta)=\Gamma_x(\alpha)/\beta^\alpha$ and $\Gamma_x(\alpha)$ is the classical incomplete (upper) gamma function. Moreover, by using the Proposition 2.13, we find

$$\xi J(F,G;t) \geq \max\left\{-\frac{1}{2\theta}, -\frac{\lambda+2+\lambda t}{2\lambda(\lambda+1+\lambda t)}\right\} = \max\left\{-\frac{M_F(t)}{2}, -\frac{M_G(t)}{2}\right\}, \qquad \theta,\lambda>0,\ t\geq 0.$$

The function $\xi J(F,G;t)$ is shown in Figure 5 for selected values of θ and λ with respect to t. From the left panel, we observe that $\xi J(F,G;t)$ behaves nonmonotonically for smaller values of λ and θ. In the right panel, we see that for values of λ and θ greater than 1, $\xi J(F,G;t)$ consistently exhibits a decreasing trend over time. This decreasing behavior indicates that the measure of inaccuracy between the exponential and Lindley distributions diminishes as time progresses, without the nonmonotonic fluctuations observed at smaller parameter values.

In the following, we provide lower and upper bounds for the DCREI of non-negative random variables in terms of dynamic survival extropy based on the usual stochastic order.

Definition 2.16. *Let X and Y be two non-negative continuous random variables with SFs $\bar{F}(x)$ and $\bar{G}(x)$, respectively. Then X is said to be less than Y in the usual stochastic order (written as $X \overset{st}{\leqslant} Y$) if $\bar{F}(x) \leq \bar{G}(x)$.*

Corollary 2.17. *Let X and Y be two non-negative continuous random variables with SFs $\bar{F}(x)$ and $\bar{G}(x)$, respectively. If $X_t \overset{st}{\leqslant} Y_t$, $t>0$, then $\xi J(G;t) \leq \xi J(F,G;t) \leq \xi J(F;t)$.*

Proof. Suppose that $X_t \overset{st}{\leqslant} Y_t$, $t>0$. Then

$$\bar{F}_t(x) \leq \bar{G}_t(x), \qquad x\geq t. \tag{10}$$

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Multiplying both sides of (10) by $-\bar{F}_t(x)/2$, it follows that $\xi J(F,G;t) \leq \xi J(F;t)$. Again, multiplying both sides of (10) by $-\bar{G}_t(x)/2$, we have $\xi J(F,G;t) \geq \xi J(G;t)$. This completes the proof. □

Proposition 2.18. *For non-negative random variable X, Abdul Sathar and Nair [1] proposed lower and upper bounds for dynamic survival extropy in terms of MRL function as*

$$-\frac{M_F(t)}{2} \leq \xi J(F;t) \leq -\frac{e^{-E[M_F(X)|X\geq t]}}{2}. \tag{11}$$

Based on (11), Proposition 2.17 gives the following lower and upper bounds for the DCREI in terms of MRL functions:

$$-\frac{M_G(t)}{2} \leq \xi J(F,G;t) \leq -\frac{e^{-E[M_F(X)|X\geq t]}}{2}. \tag{12}$$

3. Characterization results

The main challenge is to identify the conditions under which the given inaccuracy measure referenced as (6) provides a unique characterization of the distribution function. The proportional hazard model (PHM) is described by two non-negative continuous random variables X and Y as (Cox [7] and Efron [8])

$$\bar{G}(x) = [\bar{F}(x)]^{\delta}, \qquad \delta > 0. \tag{13}$$

In the following, we propose some characterization results for the DCREI under the PHM (13). In Theorem 3, we propose that if DCREI remains finite and increases for all $t \geq 0$, then it uniquely determines the SF $\bar{F}(x)$ of the variable X. In Theorem 4, we further provide the characterization of three specific lifetime distributions based on the DCREI under the PHM. For instance, X follows a finite range distribution, Pareto distribution, and exponential distribution under certain conditions for the constant c. These theorems establish the theoretical basis for the unique characterization of these distributions by DCREI.

Theorem 3.1 *Let X and Y be two non-negative random variables with SFs $\bar{F}(x)$ and $\bar{G}(x)$ satisfying the PHM. If the function $\xi J(F;G,t)$ remains finite and increases for all $t \geq 0$, then it uniquely determines the SF $\bar{F}(x)$ of the variable X.*

Proof. Since X and Y satisfy the PHM, the DCREI (6) can be rewritten as

$$\xi J(F,G;t) = -\frac{1}{2}\int_t^{+\infty} \left[\frac{\bar{F}(x)}{\bar{F}(t)}\right]^{\delta+1} dx. \tag{14}$$

Differentiating (14) with respect to t on both sides, we obtain

$$\frac{\partial}{\partial t}\xi J(F,G;t) = \frac{1}{2} + (\delta+1)\,\lambda_F(t)\xi J(F,G;t), \tag{15}$$

where $\lambda_F(x) = \frac{f_X(x)}{\bar{F}_X(x)}$ is the hazard rate function. Assume that $F_1(x)$, $G_1(x)$ and $F_2(x)$, $G_2(x)$ are two sets of the PDFs satisfying the PHM, that is, $\bar{G}_1(x) = \bar{F}_1(x)$ and $\bar{G}_2(x) = \bar{F}_2(x)$. Suppose that

$$\xi J(F_1,G_1;t) \quad = \quad \xi J(F_2,G_2;t), \;\; t \geq 0. \tag{16}$$

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

By taking the derivative of equation (16) with respect to t on both sides, equation (15) can be used to determine the value of

$$\lambda_{F_1}(t)\,(\delta+1)\,\xi J(F_1,G_1;t)=\lambda_{F_2}(t)\,(\delta+1)\,\xi J(F_2,G_2;t). \tag{17}$$

If $\forall\, t\geq 0$, $\lambda_{F_1}(t)=\lambda_{F2}(t)$, then $\bar{F}_1(x)=\bar{F}_2(x)$ and the evidence will be finished.

Otherwise, assume that $A=\left\{t: t\geq 0,\ \textit{and}\ \lambda_{F_1}(t)\neq\lambda_{F_2}(t)\right\}$, and let set A to be nonempty. Thus for some $t_0\in A$, $\lambda_{F_1}(t_0)\neq\lambda_{F_2}(t_0)$. Without loss of generality, let $\lambda_{F_2}(t_0)>\lambda_{F_1}(t_0)$. Using this, equation (17) gives $\lambda_{F_1}(t)\,(\delta+1)\,\xi J(F_1,G_1;t)\geq\lambda_{F_2}(t)\,(\delta+1)\,\xi J(F_2,G_2;t)$, so for $t=t_0$, $\lambda_{F_2}(t_0)<\lambda_{F_1}(t_0)$. Therefore, set A is empty, which completes the proof. □

In what follows, we give the characterization of three distributions based on the DCREI.

Theorem 3.2 *Let X and Y be two non-negative continuous random variables satisfying the PHM. Then, for $t\geq 0$,*

$$\xi J(F,G;t)=c\,M_F(t), \tag{18}$$

if and only if

(I) the random variable X follows the finite range distribution for $c=-\frac{a+1}{2(a\delta+(a+1))}$, $a,\delta>0$,
(II) the random variable X follows the Pareto distribution for $c=-\frac{a-1}{2(a(\delta+1)-1)}$, $a,\delta>0$,
(III) X follows the exponential distribution for $c=-\frac{1}{2(\delta+1)}$, $\delta>0$.

Proof. First we prove the if part.

(I) Let X follow a finite range distribution with SF and MRL as follows:

$$\bar{F}(t)=(1-t)^a,\qquad M_F(t)=\frac{1-t}{1+a}.$$

Then, DCREI (6) under PHM is given by

$$\xi J(F,G;t)=-\frac{1-t}{2\,(1+(\delta+1)a)}=c\,M_F(t),$$

where $c=-\frac{a+1}{2(a\delta+(a+1))}$.

(II) Suppose that X follows the Pareto distribution with SF and MRL as follows:

$$\bar{F}(t)=\left(1+\frac{t}{b}\right)^{-a},\qquad M_F(t)=\frac{t+b}{a-1},\qquad t\geq 0.$$

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Then, DCREI (6) under PHM is obtained as

$$\xi J(F,G;t) = -\frac{t}{2\,(a(\delta+1)-1)} = c\,M_F(t),$$

where $c = -\frac{a-1}{2(a(\delta+1)-1)}$.

(III) Let X have the exponential distribution with SF $\bar{F}(x) = e^{-\theta x}$, $\theta > 0$, and MRL function $M_F(t) = \frac{1}{\theta}$. Then, under PHM and for $c = -\frac{1}{2(\delta+1)}$, we have

$$\xi J(F,G;t) = -\frac{1}{2\theta(\delta+1)} = c\,M_F(t).$$

To prove the only if part, let relation (18) hold. Then, under the PHM assumptions, we have

$$\int_t^{+\infty} \bar{F}^{\delta+1}(x)dx = -2c\,M_F(t)\bar{F}^{\delta+1}(t).$$

Differentiating both sides with respect to t and using the relation $\lambda_F(x) = \frac{M_F'(x)+1}{M_F(x)}$, we get

$$M_F'(t) = -\frac{1+2c(\delta+1)}{2c\delta},$$

which by integration gives

$$M_F(t) = -\frac{1+2c(\delta+1)}{2c\delta}t + W,$$

where W is a constant of integration. Now, $c = -\frac{a+1}{2(a\delta+(a+1))}$ and $W = \frac{1}{a+1}$ imply that $M_F(t) = \frac{1-t}{1+a}$. Thus, we have the finite range distribution with parameter $a > 0$. The proof of the other cases is similar and hence omitted. □

Since the MRL function for a random variable X having the exponential distribution is constant, part (*III*) of Theorem 3.2 gives the following result.

Remark 3.3. Let X be an exponentially distributed random variable. Then $\xi J(F,G;t)$ is independent of t if and only if two random variables X and Y satisfy the PHM (13).

Example 3.4. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed (iid) non-negative random variables having SFs $\bar{F}(x)$. If $X_{i:n}$ denotes the ith order statistics in this sample of size n, then the lifetime of series system is determined by $X_{1:n}$ with SF $\bar{F}_{1:n}(x) = [\bar{F}(x)]^n$. Hence, the two random variables X and $X_{1:n}$ satisfy the PHM. Therefore, Remark 3.3 gives

$$\xi J(F,F_{1:n};t) = -\frac{M_F^{n+1}(t)}{2} = -\frac{1}{2(n+1)\theta},$$

which is independent of t.

4. Nonparametric estimation

In this section, we design nonparametric estimators for the CREI and DCREI measures. Suppose that $X_1, \ldots, X_n$ is a random sample from $F(\cdot)$ and $G(\cdot)$, which is a known Cumulative Distribution Function

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

(CDF). From (5) and (6), we define nonparametric estimators of CREI and DCREI, respectively, as

$$\xi J(\widehat{F}, G) = -\frac{1}{2}\int_0^\infty \widehat{\bar{F}}(x)\bar{G}(x)dx, \tag{19}$$

$$\xi J(\widehat{F}, G; t) = -\frac{1}{2}\int_t^\infty \frac{\widehat{\bar{F}}(x)}{\widehat{\bar{F}}(t)}\frac{\bar{G}(x)}{\bar{G}(t)}dx, \tag{20}$$

where $\bar{G}(\cdot) = 1 - G(\cdot)$ in which $\widehat{\bar{F}}(\cdot) = 1 - \widehat{F}(\cdot)$, and $\widehat{F}(\cdot)$ denotes the estimator of $F(\cdot)$. We provide empirical and kernel methods to estimate the CDF $F(x)$ used in (19) and (20).

In the empirical method, the CDF is estimated by $\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^n I(X_i \le x)$. So, denoting by $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ the order statistics of random samples, we define the empirical CREI as

$$\begin{aligned}\xi J(\widehat{F}_n, G) &= -\frac{1}{2}\int_0^\infty (1 - \widehat{F}_n(x))\bar{G}(x)dx \\ &= -\frac{1}{2}\sum_{j=1}^{n-1}\int_{X_{(j)}}^{X_{(j+1)}} (1 - \frac{j}{n})\bar{G}(x)dx \\ &= -\frac{1}{2}\sum_{j=1}^{n-1}(1 - \frac{j}{n})U_{(j+1)},\end{aligned}$$

where $U_{(j+1)} = V(X_{(j+1)}) - V(X_{(j)}), j = 1, 2, \cdots, n-1$ are based on the sample spacings and $\frac{\partial V(x)}{\partial x} = \bar{G}(x)$. Also, by considering the sample values greater than t as $X_{(j)} \le X_{(j+1)} \le \cdots \le X_{(n)}$, we define the empirical DCREI as

$$\begin{aligned}\xi J(\widehat{F}_n, G; t) &= -\frac{1}{2\bar{G}(t)}\int_t^\infty (1 - \widehat{F}_n(x))\bar{G}(x)dx \\ &= -\frac{n}{2\bar{G}(t)}\sum_{j=i}^{n-1}\int_{X_{(j)}}^{X_{(j+1)}} (1 - \frac{j}{n})\bar{G}(x)dx \\ &= -\frac{n}{2\bar{G}(t)}\sum_{j=i}^{n-1}(1 - \frac{j}{n})U_{(j+1)}.\end{aligned}$$

Now, we provide another method for the estimation of CREI and DCREI measures by replacing the kernel estimator of CDF in (19) and (20). The empirical estimation is the most commonly used nonparametric estimation of CDF, but this estimation of CDF is a step function even in the case that CDF is continuous. So, Nadaraya [27] proposed the kernel distribution estimator as

$$\widehat{F}_h(x) = \frac{1}{n}\sum_{i=1}^n W(\frac{x - X_i}{h}),$$

where h is a bandwidth or smoothness parameter and $W(x) = \int_{-\infty}^x K(t)dt$ is a CDF of a kernel function $K(\cdot)$ defined by Parzen [30]. When applying $\widehat{F}_h$, one needs to choose the kernel and the bandwidth.

The kernel function is a fundamental component in kernel density estimation and other nonparametric statistical methods. It is a symmetric, non-negative function that integrates to one over its entire domain, ensuring that the estimated density is a valid probability distribution. The kernel function smoothens data points by placing a weighted influence around each point, with the bandwidth parameter

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

controlling the width of this influence. Common kernel functions include the Gaussian, Epanechnikov, and uniform kernels, each with unique properties but all sharing the essential characteristics of symmetry, non-negativity, and normalization. These properties ensure that the kernel function effectively balances bias and variance in the density estimation process, providing a smooth and accurate representation of the underlying data distribution. It is shown by Lejeune and Sarda [22] that the choice of the kernel is less important than the choice of the bandwidth for the performance of the estimation of CDF.

The bandwidth parameter in the kernel estimator of the cumulative distribution function (CDF) plays a crucial role in determining the smoothness of the estimated distribution. Essentially, the bandwidth parameter controls the width of the kernel function, which is used to smooth the data points. A larger bandwidth parameter results in a smoother CDF estimate, reducing the variance but potentially introducing bias by over-smoothing the data. Conversely, a smaller bandwidth parameter captures more details and variations in the data, reducing bias but increasing variance and the risk of overfitting to noise. Selecting an appropriate bandwidth is vital for balancing this trade-off between bias and variance, ensuring an accurate and reliable estimation of the underlying distribution.

In general, the idea underlying bandwidth selection is the minimization of the mean integrated squared error (MISE), defined as

$$MISE(h) = E\Big[\int_{-\infty}^{+\infty} \big(\widehat{F}_h(x) - F(x)\big)^2 dx\Big], \tag{21}$$

where E denotes the expected value with respect to that sample. The MISE is a key metric in CDF estimation that measures the average squared difference between the estimated CDF and the true CDF over the entire data range. It combines both bias and variance of the estimator, providing a comprehensive assessment of its accuracy. In the context of bandwidth selection, MISE is crucial because the bandwidth controls the smoothness of the density estimate. A well-chosen bandwidth minimizes the MISE, balancing the trade-off between bias (under-smoothing) and variance (over-smoothing), leading to a more accurate and reliable CDF estimate. This makes MISE a fundamental concept in nonparametric statistics, guiding the optimal bandwidth selection for CDF estimation.

For bandwidth selection in CDF estimation, the plug-in (PI) method and cross-validation (CV) method are commonly used. The PI method involves estimating the optimal bandwidth by substituting estimates of unknown quantities into an asymptotic formula for the mean integrated squared error (MISE). This approach leverages theoretical properties of the data distribution to derive a bandwidth that minimizes the MISE. On the other hand, the CV method involves dividing the data into subsets, using one subset to estimate the CDF and another to evaluate the error. The bandwidth that minimizes the CV error is selected. This method is more data-driven and does not rely on theoretical assumptions, making it versatile but computationally intensive. Both methods aim to balance the trade-off between bias and variance to achieve an accurate CDF estimate.

Choosing between the PI method and the CV method for bandwidth selection often depends on the specific context and requirements of your analysis. The PI method is generally preferred when you have a good theoretical understanding of the data distribution, as it leverages this knowledge to provide a more computationally efficient bandwidth estimate. However, if you lack strong assumptions about the data or prefer a more data-driven approach, the CV method is advantageous. It directly evaluates the performance of different bandwidths on the data itself although it can be more computationally intensive. Ultimately, the choice hinges on the balance between computational efficiency and the need for flexibility in handling diverse data characteristics.

In this paper, we use the PI approach provided by Polansky and Baker [31] that developed the previous ideas by Altman and Leger [2]. They showed that $h^{PI} = \widehat{C}n^{-1/3}$, where $\widehat{C}$ is estimated through the data sample. A well-known method on the bandwidth selection for the CDF estimation is the CV method that is initially proposed by Sarda [35]. Here, we use the CV approach proposed by

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Bowman *et al.* [5]. They considered a CV bandwidth selection as

$$h^{CV} = \arg\min_{h} \frac{1}{n} \sum_{i=1}^{n} \int_{-\infty}^{\infty} \left(I(x - x_i \geq 0) - \widehat{F}_{h,-i}(x)\right)^2 dx,$$

where $\widehat{F}_{h,-i}(X_i)$ denotes the kernel estimator constructed from the data with observation x_i omitted. The simulation study by Bowman *et al.* [5] showed that the CV approach has better performance than the PI approach.

5. Simulation study

Now, we provide a simulation study to evaluate the performance of the estimators proposed in the previous section. The primary objective of this simulation study is to assess the accuracy and reliability of the estimators under various conditions. Specifically, we aim to compare the performance of these estimators using two key metrics: Bias and root mean squared error (RMSE). Bias measures the difference between the expected value of the estimator and the true value of the parameter being estimated. It provides insight into the systematic error present in the estimators. A lower bias indicates that the estimator is, on average, closer to the true parameter value. RMSE combines both the variance and the bias of the estimator, offering a comprehensive measure of its accuracy. RMSE is particularly useful as it penalizes larger errors more heavily, thus providing a more stringent assessment of the estimator's performance. By doing so, we gain valuable insights into how well these estimators perform under various conditions and identify any potential strengths or limitations.

Considering various sample sizes, parameters, and times in the simulation study enhances our understanding of estimator behaviour, robustness, and applicability. Researchers can make informed decisions based on these insights. By comparing estimators across different sample sizes, we assess their statistical power. Larger sample sizes tend to yield more precise estimates and reduce bias. Understanding how estimators behave under varying sample sizes informs our confidence in their performance. Investigating performance across different sample sizes and parameters helps us understand how estimators behave in practical scenarios. Smaller samples may represent limited data availability, while larger samples provide more robust estimates. Also, investigating estimators across various parameters helps us identify optimal settings. Properly tuned parameters lead to better estimation accuracy. Sensitivity analysis reveals how estimators respond to parameter changes. Furthermore, estimators can exhibit different behaviour at different time points. Studying estimators at multiple time points allows us to capture dynamic processes. Some estimators may adapt better to changing distributions over time, while others may be more stable.

For simulations, we consider exponential, Weibull, and Pareto distributions used in Examples 2.6, 2.7, and 2.14. Also, we calculate the bias and RMSE of DCREI estimators for various sample sizes ($n = 20, 100$) and times ($t = 0, 0.2, 0.5, 1$). It is obvious that in the case when $t = 0$, the CREI measure will be estimated. The Bias and RMSE of the proposed estimators are computed as follows: for the exponential distribution with scale parameters ($\lambda = 0.1$ and 0.5), see Table 1; for the Weibull distribution with scale parameters ($\lambda = 0.1$ and 0.5) and shape parameter ($\sigma = 2$), see Table 2; for the Pareto distribution with scale parameters ($\sigma = 1$ and 5) and shape parameter ($\alpha = 1.5$), see Table 3; and for the Pareto distribution with scale parameters ($\sigma = 1$ and 5) and shape parameter ($\alpha = 2$), see Table 4.

By evaluating the performance of the estimators, the following results were obtained:

- Kernel-based estimators generally outperform empirical-based estimators in terms of root mean squared error (RMSE) across various scenarios. These estimators leverage kernel functions to smooth the data and provide more accurate estimates.
- The choice of kernel and bandwidth parameters significantly impacts the performance of kernel-based estimators. Properly tuning these parameters is crucial for achieving optimal results.

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Table 1. *Bias and RMSE estimation of DCREI estimators based on exponential distribution with parameter θ for the $F(\cdot)$ and parameter λ for the $G(\cdot)$*

n	(θ,λ)	t	$\xi J(\widehat{F}_n, G; t)$		$\xi J(\widehat{F}_{h^{PI}}, G; t)$		$\xi J(\widehat{F}_{h^{CV}}, G; t)$	
			Bias	RMSE	Bias	RMSE	Bias	RMSE
20	(0.1,0.1)	0	–0.31606	0.45735	–0.28171	0.43560	–0.00775	0.34992
		0.2	–0.30175	0.44911	–0.26783	0.42947	–0.00642	0.33007
		0.5	–0.27227	0.43148	–0.23637	0.40992	–0.00715	0.32947
		1	–0.22014	0.41215	–0.18453	0.39308	0.00364	0.31353
	(0.1,0.5)	0	–0.05775	0.06531	–0.05356	0.06289	–0.00136	0.05984
		0.2	–0.05549	0.06412	–0.05068	0.06180	–0.00213	0.05809
		0.5	–0.05013	0.06078	–0.04453	0.05844	0.00022	0.05778
		1	–0.04350	0.05688	–0.03820	0.05505	–0.00397	0.05477
	(0.5,0.5)	0	–0.06399	0.09276	–0.05773	0.08906	–0.00142	0.07919
		0.2	–0.04566	0.08283	–0.03880	0.07928	–0.00077	0.07637
		0.5	–0.02110	0.07427	–0.01544	0.07379	0.00296	0.07219
		1	–0.01064	0.07287	0.01037	0.07092	0.00272	0.06634
100	(0.1,0.1)	0	–0.25968	0.29937	–0.19204	0.24408	–0.00489	0.16817
		0.2	–0.23606	0.28141	–0.16613	0.22556	0.00167	0.15605
		0.5	–0.21611	0.26547	–0.14316	0.20936	–0.00239	0.15295
		1	–0.16847	0.22638	–0.10027	0.17996	0.00058	0.14829
	(0.1,0.5)	0	–0.05124	0.05328	–0.04163	0.04519	0.00014	0.02761
		0.2	–0.04730	0.04975	–0.03671	0.04108	0.00122	0.02607
		0.5	–0.04364	0.04645	–0.03247	0.03749	0.00011	0.02545
		1	–0.03530	0.03896	–0.02282	0.03016	0.00122	0.02393
	(0.5,0.5)	0	–0.05130	0.05879	–0.03781	0.04773	0.00000	0.03881
		0.2	–0.03380	0.05115	–0.02016	0.04658	0.00013	0.03397
		0.5	–0.01631	0.04952	–0.00627	0.03904	–0.00046	0.03264
		1	–0.00307	0.03907	–0.00111	0.03158	–0.00109	0.03004

- Among the estimators considered, the one denoted as $\xi J(\widehat{F}_{h^{CV}}, G; t)$ exhibits the best performance. This estimator likely benefits from CV bandwidth selection, which helps adapt to the underlying data distribution.
- The estimator $\xi J(\widehat{F}_{h^{PI}}, G; t)$ also performs well, surpassing the performance of the simple empirical estimator $\xi J(\widehat{F}_n, G; t)$. The use of PI bandwidth selection likely contributes to its improved accuracy.
- Increasing the sample size has a positive impact on both bias and RMSE values for the proposed estimators. As the sample size grows, estimators tend to converge to the true distribution, resulting in reduced bias and more precise estimates.

The implications of the Bias and RMSE values in Tables 1–4 for each estimator and distribution are as follows:

1. **Exponential and Weibull Distributions**:
   - RMSE Decreases with Time (t): As time increases, the RMSE values decrease for both exponential and Weibull distributions. This suggests that estimators become more accurate as more data become available over time.

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Table 2. *Bias and RMSE estimation of DCREI estimators based on Weibull distribution with parameters* $(\theta, \sigma = 2)$ *for the* $F(\cdot)$ *and parameters* $(\lambda, \sigma = 2)$ *for the* $G(\cdot)$

n	(θ, λ)	t	$\xi J(\widehat{F}_n, G; t)$		$\xi J(\widehat{F}_{h^{PI}}, G; t)$		$\xi J(\widehat{F}_{h^{CV}}, G; t)$	
			Bias	RMSE	Bias	RMSE	Bias	RMSE
20	(0.1,0.1)	0	0.00119	0.07685	−0.01176	0.06715	−0.02992	0.06401
		0.2	0.00229	0.07505	−0.00944	0.06518	−0.01297	0.06289
		0.5	0.00316	0.06878	−0.01423	0.06089	−0.01909	0.05893
		1	−0.00303	0.06491	−0.01877	0.05729	−0.02492	0.05661
	(0.1,0.5)	0	0.00046	0.02327	0.00551	0.01935	0.00678	0.01872
		0.2	−0.00123	0.02215	−0.00457	0.01553	0.00520	0.01448
		0.5	−0.00064	0.01971	−0.00249	0.01355	−0.00272	0.01224
		1	0.00020	0.01712	−0.00278	0.01089	−0.00364	0.00996
	(0.5,0.5)	0	−0.00082	0.03302	−0.00951	0.02960	−0.01731	0.02834
		0.2	−0.00074	0.03168	−0.00779	0.02816	−0.00970	0.02723
		0.5	0.00035	0.02985	−0.00687	0.02571	−0.00989	0.02484
		1	0.00062	0.02695	−0.00460	0.02144	−0.00708	0.02083
100	(0.1,0.1)	0	0.00014	0.03347	0.00645	0.03093	0.00608	0.03030
		0.2	−0.00092	0.03216	−0.00675	0.03010	−0.00802	0.02961
		0.5	0.00092	0.03150	−0.00894	0.02919	−0.01109	0.02849
		1	0.00026	0.02944	−0.01063	0.02793	−0.01350	0.02792
	(0.1,0.5)	0	0.00002	0.01005	0.00432	0.00955	0.00495	0.00947
		0.2	−0.00023	0.00980	0.00071	0.00747	0.00098	0.00722
		0.5	0.00005	0.00936	−0.00129	0.00666	−0.00151	0.00629
		1	0.00013	0.00772	−0.00147	0.00538	−0.00196	0.00507
	(0.5,0.5)	0	0.00033	0.01406	0.00063	0.01403	0.00040	0.01373
		0.2	−0.00034	0.01364	−0.00448	0.01308	−0.00538	0.01299
		0.5	0.00032	0.01321	−0.00468	0.01230	−0.00603	0.01131
		1	0.00009	0.01267	−0.00403	0.01071	−0.00494	0.00977

- Scale Parameter Influence: By decreasing the scale parameter (θ) in $F(\cdot)$ or increasing the scale parameter (λ) in $G(\cdot)$, the RMSE values decrease. Properly tuning these parameters improves estimator performance.

2. **Pareto Distribution**:
   - RMSE Increase with Time (t): Unlike exponential and Weibull distributions, the RMSE values increase with time for the Pareto distribution. This indicates that estimators struggle to handle the tail-heavy nature of the Pareto distribution.
   - Reducing Scale Parameter Improves Performance: By reducing the scale parameter in the Pareto distribution, both Bias and RMSE values decrease. This adjustment enhances estimator accuracy.
   - Scale Parameter Impact on Bias and RMSE: The comparison between Tables 3 and 4 highlights that reducing the scale parameter in the Pareto distribution leads to improved estimator performance. Both Bias and RMSE values benefit from this adjustment.

In summary, understanding how different distributions and parameters affect Bias and RMSE values helps us choose appropriate estimators for specific scenarios. Researchers can use this information to make informed decisions when analyzing real-world data.

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Table 3. *Bias and RMSE estimation of DCREI estimators based on Pareto distribution with parameters* $(\sigma_1 = 1, \alpha_1)$ *for the* $F(\cdot)$ *and parameters* $(\sigma_2 = 1, \alpha_2)$ *for the* $G(\cdot)$

n	(α_1, α_2)	t	$\xi J(\widehat{F}_n, G; t)$		$\xi J(\widehat{F}_{h^{PI}}, G; t)$		$\xi J(\widehat{F}_{h^{CV}}, G; t)$	
			Bias	RMSE	Bias	RMSE	Bias	RMSE
20	(1.5,1.5)	0	–0.05075	0.07689	–0.04065	0.06715	0.00137	0.04264
		0.2	–0.02613	0.07805	–0.01638	0.07023	–0.00101	0.06284
		0.5	0.01061	0.08521	0.01241	0.08166	0.00205	0.08152
		1	0.03085	0.15601	0.02299	0.14099	0.00323	0.13274
	(1.5,2)	0	–0.03982	0.04553	–0.03235	0.04280	–0.00091	0.03060
		0.2	–0.01970	0.05030	–0.01276	0.04504	–0.00085	0.04030
		0.5	0.00263	0.05488	0.00452	0.05200	–0.00129	0.05150
		1	0.01874	0.12321	0.01363	0.08575	–0.00218	0.07132
	(2,2)	0	–0.03095	0.03615	–0.02532	0.03123	0.00034	0.02703
		0.2	–0.00863	0.04023	–0.00450	0.03690	–0.00012	0.03646
		0.5	0.01114	0.05315	0.00942	0.05174	0.00271	0.04396
		1	0.00695	0.10283	0.00377	0.06800	–0.00533	0.05255
100	(1.5,1.5)	0	–0.04245	0.02504	–0.02767	0.02279	–0.00009	0.01981
		0.2	–0.01755	0.03372	–0.00716	0.03168	–0.00054	0.02757
		0.5	0.00609	0.03968	0.00185	0.03721	0.00152	0.03187
		1	0.00738	0.06721	–0.00279	0.06677	–0.00105	0.06147
	(1.5,2)	0	–0.03196	0.01580	–0.02088	0.01561	–0.00007	0.01304
		0.2	–0.01317	0.02186	–0.00489	0.01881	0.00049	0.01748
		0.5	0.00258	0.02427	0.00034	0.02362	–0.00019	0.02256
		1	0.00582	0.03730	–0.00114	0.03641	–0.00179	0.03429
	(2,2)	0	–0.02447	0.01282	–0.01590	0.01058	0.00025	0.00911
		0.2	–0.00456	0.01801	–0.00052	0.01611	–0.00010	0.01573
		0.5	0.00479	0.02259	0.00084	0.02141	–0.00074	0.02032
		1	0.00159	0.03725	–0.00144	0.03560	–0.00016	0.03291

6. Application

The objective of the application section is to demonstrate the usefulness of the dynamic cumulative residual extropy inaccuracy (DCREI) measure in real-life problems. Specifically, the authors apply the DCREI measure to model selection. By doing so, they showcase how this novel inaccuracy measure can be practically employed in statistical analysis and decision-making processes. Overall, the application section aims to illustrate the practical implications and benefits of the DCREI measure in statistical modeling and inference.

The DCREI measure plays a crucial role in identifying the best-fitted distribution from data. DCREI allows us to compare different lifetime distribution models based on their inaccuracy. By calculating DCREI for each candidate distribution, we can assess how well they fit the observed data and lower DCREI values indicate better model fit. So, This ensures that the selected model minimizes the discrepancy between the theoretical and observed cumulative residual extropy. In reliability analysis, DCREI helps choose an appropriate distribution for modeling failure times. It guides practitioners toward the most suitable distribution, improving predictions and decision-making. In summary, DCREI serves as a practical tool for assessing model fit and selecting the best-fitted distribution, enhancing statistical inference and reliability modeling.

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Table 4. *Bias and RMSE estimation of DCREI estimators based on Pareto distribution with parameters* $(\sigma_1 = 5, \alpha_1)$ *for the* $F(\cdot)$ *and parameters* $(\sigma_2 = 5, \alpha_2)$ *for the* $G(\cdot)$

n	(α_1, α_2)	t	$\xi J(\widehat{F}_n, G; t)$		$\xi J(\widehat{F}_{h^{PI}}, G; t)$		$\xi J(\widehat{F}_{h^{CV}}, G; t)$	
			Bias	RMSE	Bias	RMSE	Bias	RMSE
20	(1.5,1.5)	0	−0.25376	0.38446	−0.20322	0.33567	0.00681	0.21321
		0.2	−0.24039	0.39853	−0.18655	0.35083	−0.00476	0.24352
		0.5	−0.17665	0.41368	−0.12621	0.37587	0.00918	0.26063
		1	−0.12628	0.43736	−0.07773	0.39543	−0.00853	0.30487
	(1.5,2)	0	−0.19912	0.25381	−0.16176	0.22546	−0.00153	0.15304
		0.2	−0.18008	0.26070	−0.14401	0.22795	−0.00659	0.15752
		0.5	−0.15132	0.26588	−0.11188	0.23467	−0.00458	0.17372
		1	−0.10189	0.27634	−0.06541	0.24398	−0.00810	0.20502
	(2,2)	0	−0.15473	0.19802	−0.12659	0.18204	0.00699	0.13511
		0.2	−0.13370	0.20540	−0.10563	0.18408	−0.00408	0.14234
		0.5	−0.08505	0.21594	−0.05830	0.19257	0.00525	0.15806
		1	−0.03280	0.23073	−0.01259	0.20613	0.00639	0.18477
100	(1.5,1.5)	0	−0.21224	0.18435	−0.13834	0.15792	−0.00049	0.09904
		0.2	−0.18617	0.19625	−0.11355	0.14756	−0.00355	0.10588
		0.5	−0.13577	0.23574	−0.06840	0.17621	0.00481	0.11621
		1	−0.08926	0.25187	−0.03873	0.18945	−0.00148	0.13884
	(1.5,2)	0	−0.15979	0.12406	−0.10437	0.09803	−0.00035	0.06520
		0.2	−0.13832	0.14545	−0.08265	0.10319	−0.00079	0.06880
		0.5	−0.11254	0.16262	−0.05952	0.11371	−0.00327	0.07461
		1	−0.06569	0.17899	−0.02407	0.12806	0.00803	0.09343
	(2,2)	0	−0.12235	0.09197	−0.07950	0.08271	0.00124	0.05696
		0.2	−0.09953	0.10331	−0.05638	0.07791	−0.00098	0.06407
		0.5	−0.06527	0.12559	−0.02732	0.09053	0.00088	0.07126
		1	−0.02320	0.14111	−0.00280	0.10291	−0.00023	0.08664

Now, we show how the DCREI measure can be a proper measure to find the best-fitted distribution from data in two real data sets.

1. Gas oil production data

We consider the data set that consists of gas oil production (in cubic meters per day) in the country's refineries in Iran from 1,992 to 2,014 as follows:

64,731, 69,545, 69,945, 70,879, 71,923, 73,154, 77,037, 79,215, 80,473, 81,549, 84,957, 88,702, 90,951, 94,677, 93,595, 97,689, 96,016.

The data on gas oil (crude oil) production in Iran from 1,992 to 2,014 are significant for understanding Iran's energy landscape during that period. This data set is reported in Khorashadizadeh [15]. In generally, this type of data comes from various reputable sources, including the International Energy Statistics provided by organizations such as the Energy Information Administration and the Organization of the Petroleum Exporting Countries. These organizations collect and report data on energy production, consumption, and related metrics globally. The data represent the annual production of crude oil (gas oil) in Iran, measured in barrels per day (bbl/d). It provides insights into Iran's oil industry, its capacity, and fluctuations over time. Understanding Iran's crude oil production during this time frame is crucial for assessing its economic stability, energy security, and geopolitical influence. Researchers,

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Table 5. *Model selection criteria for Gas oil production data*

Distribution	Log-likelihood	AIC	BIC	K–S	*P*-Value	$\overline{\eta_G(t)}$
Exponential	−209.236	420.473	421.306	0.548	2.53E-05	246.931
Weibull	−181.487	366.973	368.639	0.141	0.841	54.818
Log-normal	−181.123	366.245	367.912	0.137	0.866	15.943
Logistic	−183.953	371.906	373.573	0.182	0.566	94.872
Pareto	−214.811	433.621	435.287	0.590	3.67E-06	256.572

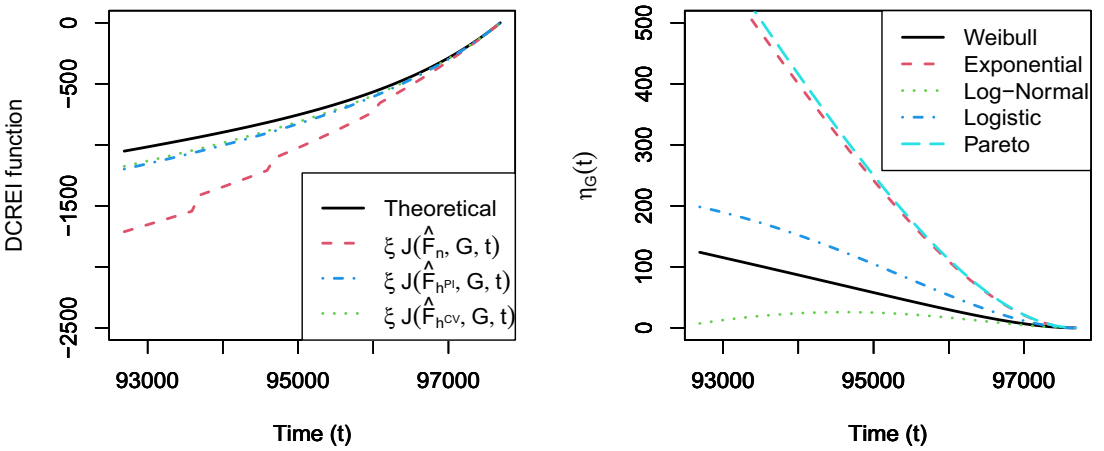


Figure 6. *Graphs of DCREI function (left panel) and $\eta_G(t)$ (right panel) for gas oil production data.*

policymakers, and analysts use these data to study trends, identify patterns, and evaluate the impact of global events (such as sanctions, conflicts, or technological advancements) on Iran's oil production. In summary, the gas oil production data from 1,992 to 2,014 provide valuable context for analyzing Iran's energy dynamics and its role in the global oil market.

In Table 5, the values of the log-likelihood, AIC, and BIC, as well as the Kolmogorov–Smirnov (K–S) goodness of fit test, are presented for choosing the best model among exponential, Weibull, log-normal, Logistic, and Pareto distributions. The results of this table show that among the considered distributions, the log-normal distribution is closer to the real distribution of Gas oil production data. In this distribution, the maximum likelihood estimation of the location and dispersion parameters are equal to 11.30 and 0.12, respectively.

In the right panel of Figure 6, the behavior of the proposed estimators compared to the theoretical values of $\xi J(F, G; t)$ based on the log-normal distribution is plotted. It can be seen that the behavior of estimator $\xi J(\widehat{F}_{h^{CV}}, G; t)$ is closer to the theoretical values than other estimators, which can be a reason for the better performance of this estimator compared to other estimators.

Now, we propose the following criterion for model selection among the proposed distributions based on DCREI measures, as

$$\eta_G(t) = |\xi J(\widehat{F}, G; t) - \xi J(G, G; t)|, \qquad t \in T,$$

where T is determined based on the change interval of real data. For gas oil production data, $T \in [92686, 97689]$ is considered. In the right panel of Figure 6, the behavior of $\eta_G(t)$ with respect to time (t) is plotted. It can be seen that the values of $\eta_G(t)$ on different times for the log-normal distribution have smaller values than other distributions.

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

The criterion $\eta_G(t)$ is defined as the absolute difference between two terms $\xi J(\widehat{F}, G; t)$ and $\xi J(G, G; t)$. Essentially, $\eta_G(t)$ quantifies the deviation of the estimated cumulative residual extropy from the theoretical one. $\eta_G(t)$ is indeed defined as the absolute difference between $\xi J(\widehat{F}, G; t)$ and $\xi J(G, G; t)$. The term $\xi J(G, G; t)$ represents the CRE of the model G, which serves as our theoretical benchmark. This expression does not directly depend on the survival functions of both X and Y but rather captures the CRE of the fitted distribution G under consideration. We know that $\xi J(\widehat{F}, G; t)$ represents the estimated cumulative residual extropy using the empirical data $\widehat{F}$ and the model G and $\xi J(G, G; t)$ is the cumulative residual extropy based purely on the model G, without influence from the empirical data. This serves as the theoretical value for comparison. In essence, $\eta_G(t)$ quantifies the deviation of the empirical CRE from the theoretical CRE derived from the model G. By minimizing $\eta_G(t)$ over the specified time interval T, we aim to select the distribution that best aligns with the observed data.

To choose the best-fitting distribution, one should consider the average value of $\eta_G(t)$ over the specified time interval T. The preferred model is the one that minimizes $\overline{\eta_G}(t)$, where $\overline{\eta_G}(t)$ is defined as the average value of $\eta_G(t)$ across the interval T. In simpler terms, opt for the distribution that, on average, minimizes the difference between the estimated and theoretical cumulative residual extropy. Researchers can confidently utilize this model selection criterion to make informed decisions. By minimizing DCREI, they ensure that the selected distribution aligns well with the observed data. The chosen distribution, with consistently smaller $\eta_G(t)$ values, emerges as a strong candidate for practical applications. Therefore, the proposed criterion guides us toward selecting a distribution that optimally fits the data, highlighting the significance of DCREI in statistical modeling and inference.

The log-normal distribution effectively models right-skewed data, which is common in gas oil production due to various factors. It has a uni-modal shape, aligning with the central tendency of production values. Also, taking the natural logarithm of data transforms it into a normal distribution, aiding analysis and the log-normal distribution ensures non-negative production rates. Thus, the log-normal distribution suits gas oil production modeling and provides valuable insights for decision-making in the energy industry.

In the last column of Table 5, values of $\overline{\eta_G}(t)$ are calculated for considered distributions. The log-normal distribution has the minimum value of $\overline{\eta_G}(t)$ and therefore has the best fits for Gas oil production data. After that, the Weibull distribution has the minimum value of $\overline{\eta_G}(t)$, and so this distribution is the second best fitted distribution. These results confirm the results obtained based on log-likelihood, AIC, and BIC. Since the difference between the value of $\overline{\eta_G}(t)$ for the considered distributions is higher, it can show the difference between the distributions more clearly and so the best distribution can be chosen more accurately.

2. Failure time data

The following data are from Lawless [20], and they involve failure times for 36 appliances that underwent an automatic life test. The failure times refer to the time until an appliance fails or stops functioning. The study explores survival analysis techniques and their applications in reliability modeling.

111, 351, 491, 1,701, 3,291, 3,811, 7,081, 9,581, 10,621, 11,671, 15,941, 19,251, 19,901, 22,231, 23,271, 24,001, 24,511, 24,711, 25,511, 25,651, 25,681, 26,941, 27,021, 27,611, 28,311, 30,341, 30,591, 31,121, 32,141, 34,781, 35,041, 43,291, 63,671, 69,761, 78,461, 13,403.

Table 6 provides various statistical measures for evaluating different probability distributions (exponential, Weibull, log-normal, logistic, and Pareto) based on their fit to the observed failure time data. The results indicate that the exponential distribution is the best-fitting model among those considered. Specifically, the maximum likelihood estimate for the scale parameter in this distribution is approximately 4.13e−05.

The column in Table 6 contains the calculated values of $\overline{\eta_G}(t)$ for each considered distribution. Recall that $\overline{\eta_G}(t)$ represents the weighted cumulative residual extropy, which is a measure of inaccuracy or deviation from the ideal distribution. Lower values of $\overline{\eta_G}(t)$ indicate better fits to the observed failure times data. The exponential distribution exhibits the minimum $\overline{\eta_G}(t)$ value among all the distributions considered. This finding suggests that the exponential distribution provides the best overall fit to the

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

Table 6. *Model selection criteria for failure times data*

Distribution	Log-likelihood	AIC	BIC	K–S	*P*-Value	$\overline{\eta_G}(t)$
Exponential	−399.415	800.830	802.414	0.215	0.061	121.364
Weibull	−399.259	802.518	805.685	0.170	0.222	130.730
Log-normal	−408.624	821.249	824.416	0.249	0.019	194.934
Logistic	−402.201	808.401	811.568	0.145	0.398	148.967
Pareto	−401.675	807.350	810.517	0.251	0.017	136.632

data. Following the exponential distribution, the Weibull distribution has the second-lowest $\overline{\eta_G}(t)$ value. While not as optimal as the exponential distribution, it still reasonably fits the failure times data.

The results align with those obtained based on the Bayesian Information Criterion (BIC). BIC is a model selection criterion that penalizes complexity, favoring simpler models. The fact that the exponential distribution performs well both in terms of $\overline{\eta_G}(t)$ and BIC reinforces its suitability for modeling the data. In summary, the exponential distribution emerges as the preferred choice for modeling the failure times data, with the Weibull distribution as a close second. These insights can guide further statistical analysis and decision-making.

Overall, the proposed criterion $\eta_G(t)$ for model selection, which compares DCREI, offers several justifications and enhancements compared to existing methods like log-likelihood, AIC, and BIC. Unlike likelihood-based criteria, $\eta_G(t)$ explicitly considers cumulative residual extropy. Extropy captures information beyond likelihood, especially in reliability modeling, where understanding the inaccuracy between model predictions and observed data is crucial. On the other hand, the DCREI accounts for changes over time by comparing extropy at different intervals. It provides insights into how well a distribution fits the data across various time points, addressing temporal variations. Also, $\eta_G(t)$ accommodates nonparametric estimators (e.g., kernel-based methods), allowing flexibility beyond parametric assumptions. Researchers can apply it even when specific distributional assumptions are challenging or unknown. While AIC and BIC balanced model complexity and goodness of fit, $\eta_G(t)$ offers an alternative approach. $\eta_G(t)$ enriches the model selection toolkit by emphasizing extropy, adaptability, and dynamic assessment, making it a valuable tool for choosing the best-fitted distribution.

The dynamic evaluation of inaccuracy provided by CREI and DCREI can be beneficial in numerous applications beyond model selection. For instance, in healthcare, these measures can be used to improve the accuracy of predictive models for patient outcomes, enabling more effective treatment planning and resource allocation. In finance, CREI and DCREI can enhance risk assessment models by capturing the time-varying nature of financial risk, leading to better investment decisions and portfolio management. In environmental studies, these measures can aid in modeling and predicting the impacts of climate change by providing a more dynamic understanding of environmental data. Moreover, the flexibility of nonparametric methods and kernel-based estimators in CREI and DCREI allows for their application in various domains where data distributions may not follow conventional parametric forms. This adaptability is crucial for accurately modeling complex, real-world scenarios, such as in supply chain management, industrial maintenance, and beyond.

7. Conclusion

In this paper, by considering the concept of the CREI measure, its dynamic version was proposed. Characterization of certain specific lifetime distribution functions such as Pareto, exponential, and finite range distributions was derived, which play a vital role in reliability modeling. Moreover, upper and lower bounds and some inequalities concerning DCREI were determined. The results reported for the new inaccuracy measure are a generalization of the results obtained for CREI. Instead of a single overall measure, DCREI evaluates inaccuracy at specific time points and it captures variations in fit across different intervals. When we calculate DCREI for a specific time point ($t = 0$), it reduces to CREI. DCREI

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

provides a more detailed view of inaccuracy, allowing us to track deviations over time. By analyzing DCREI across multiple time points, we gain richer insights than CREI alone. Thus, DCREI generalizes CREI by considering temporal dynamics, providing a more nuanced understanding of distribution fit.

The nonparametric estimators of the DCREI measure were provided. Results of the simulation study showed that kernel-based estimators outperform empirical-based estimators in terms of Root Mean Squared Error (RMSE) across various sample sizes, parameters, and time points. Among the kernel-based estimators, the estimator $\xi J(\widehat{F}_{h^{CV}}, G; t)$ exhibits the best performance, followed by $\xi J(\widehat{F}_{h^{PI}}, G; t)$, which also surpasses the empirical estimator $\xi J(\widehat{F}_n, G; t)$. Additionally, both bias and RMSE values of the proposed estimators decrease as the sample size increases, highlighting the effectiveness of larger sample sizes in improving estimation accuracy. These findings underscore the importance of selecting appropriate bandwidths and leveraging kernel-based methods for more reliable nonparametric estimation. However, it is also important to acknowledge potential limitations. For instance, while kernel-based estimators generally perform well, they may require careful bandwidth selection to avoid issues like over-smoothing or under-smoothing. Additionally, in situations with very small sample sizes, nonparametric methods might not perform as robustly as parametric alternatives due to higher variance.

By minimizing DCREI, we ensure that the selected distribution closely aligns with observed data, and researchers and practitioners can confidently use DCREI as a criterion for model comparison. Therefore, improved reliability models lead to better predictions of failure times, maintenance schedules, and risk assessments. Unlike static criteria (e.g., AIC or BIC), DCREI evaluates inaccuracy dynamically over time. It captures variations in fit, especially when data patterns change. So, decision-makers gain insights into how well a model adapts to evolving conditions. For instance, in industrial equipment maintenance, DCREI helps adjust reliability estimates as machinery ages or undergoes repairs. Also, DCREI accommodates nonparametric estimators (kernel-based methods), where these estimators handle complex data distributions without strict assumptions. Furthermore, DCREI informs choices beyond statistical modeling. In supply chain management, for instance, selecting a distribution impacts inventory policies, lead times, and service levels. Decision-makers can optimize resource allocation based on reliable distribution models supported by DCREI analysis. In summary, DCREI bridges theory and practice by enhancing distribution selection, adapting to changing contexts, and aiding informed decisions. Its dynamic nature and flexibility make it a valuable tool in reliability modeling and beyond.

Acknowledgements. The authors thank the editor-in-chief, the associate editor, and the anonymous reviewers for their useful comments on the earlier version of this paper.

Competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Abdul Sathar, E.I. and Nair, R. & D. (2021). On dynamic survival extropy. *Communications in Statistics-Theory and Methods* **50**(6): 1295–1313.
[2] Altman, N. & Leger, C. (1995). Bandwidth selection for kernel distribution function estimation. *Journal of Statistical Planning and Inference* **46**(2): 195–214.
[3] Asadi, M. & Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. *Journal of Statistical Planning and Inference* **137**(6): 1931–1941.
[4] Bhatia, P.K. (1995). Useful inaccuracy of order α and 1.1 coding. *Soochow Journal of Mathematics* **21**(1): 81–87.
[5] Bowman, A., Hall, P. & Prvan, T. (1998). Bandwidth selection for the smoothing of distribution functions. *Biometrika* **85**(4): 799–808.
[6] Cover, T.M. & Thomas, J.A. (2006). *Elements of Information theory.* 2nd ed. Hoboken, New York, NY: John Wiley & Sons. Inc.
[7] Cox, D.R. (1959). The analysis of exponentially distributed lifetimes with two type of failure. *Journal of the Royal Statistical Society: Series B (Methodological)* **21**(2): 411–421.
[8] Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association* **76**(374): 312–319.
[9] Hashempour, M. & Mohammadi, M. (2024a). On dynamic cumulative past inaccuracy measure based on extropy. *Communications in Statistics-Theory and Methods* **53**(4): 1294–1311.

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press

[10] Hashempour, M. & Mohammadi, M. (2024b). A new measure of inaccuracy for record statistics based on extropy. *Probability in the Engineering and Informational Sciences* **38**(1): 207–225.
[11] Hashempour, M., Mohammadi, M., Jahanshahi, S.M.A. & Khammar, A.H. (2024). Modified cumulative extropies of doubly truncated random variables. *Sankhya B* **86**(2): 558–585.
[12] Jahanshahi, S.M.A., Zarei, H. & Khammar, A. (2020). On cumulative residual extropy. *Probability in the Engineering and Informational Sciences* **34**(4): 605–625.
[13] Kayal, S., Madhavan, S.S. & Ganapathy, R. (2017). On dynamic generalized measures of inaccuracy. *Statistica* **77**(2): 133–148.
[14] Kerridge, D.F. (1961). Inaccuracy and inference. *Journal of the Royal Statistical Society. Series B (Methodological)* **23**(1): 184–194.
[15] Khorashadizadeh, M. (2019). A quantile approach to the interval shannon entropy. *Journal of Statistical Research of Iran JSRI* **15**(2): 317–333.
[16] Kumar, V. & Taneja, H.C. (2015). Dynamic cumulative residual and past inaccuracy measures. *Journal of Statistical Theory and Applications* **14**(4): 399–412.
[17] Kundu, C., Di Crescenzo, A. & Longobardi, M. (2016). On cumulative residual (past) inaccuracy for truncated random variables. *Metrika* **79**(3): 335–356.
[18] Kundu, C. & Nanda, A.K. (2015). Characterizations based on measure of inaccuracy for truncated random variables. *Statistical Papers* **56**(3): 619–637.
[19] Lad, F., Sanfilippo, G. & Agro, G. (2015). Extropy: Complementary dual of entropy. *Statistical Science* **30**(1): 40–58.
[20] Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime data.* New York, NY: Wiley.
[21] Lebowitz, L.J. (1993). Boltzmann's entropy and time's arrow. *Physics Today* **46**(9): 8–32 10.1063/1.881363.
[22] Lejeune, M. & Sarda, P. (1992). Smooth estimators of distribution and density functions. *Computational Statistics and Data Analysis* **14**(4): 457–471.
[23] Mohammadi, M. & Hashempour, M. (2022). On interval weighted cumulative residual and past extropies. *Statistics* **56**(5): 1029–1047.
[24] Mohammadi, M. & Hashempour, M. (2023). Extropy based inaccuracy measure in order statistics. *Statistics* **57**(6): 1490–1510.
[25] Mohammadi, M. & Hashempour, M. (2024). On weighted version of dynamic cumulative residual inaccuracy measure based on extropy. *Statistical Papers* **65**(7): 4599–4629.
[26] Mohammadi, M., Hashempour, M. & Kamari, O. (2024). On the dynamic residual measure of inaccuracy based on extropy in order statistics. *Probability in the Engineering and Informational Sciences* **38**(3): 481–502.
[27] Nadaraya, E.A. (1964). On estimating regression. *Theory of Probability and Its Applications* **9**(1): 141–142.
[28] Nath, P. (1968). Inaccuracy and coding theory. *Metrika* **13**(1): 123–135.
[29] Navarro, J., Aguila, Y. & Asadi, M. (2010). Some new results on the cumulative residual entropy. *Journal of Statistical Planning and Inference* **140**(1): 310–322.
[30] Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical statistics* **33**(3): 1065–1076.
[31] Polansky, A.M. & Baker, E.R. (2000). Multistage plug-in bandwidth selection for kernel distribution function estimates. *Journal of Statistical Computation and Simulation* **65**(1): 63–80.
[32] Psarrakos, G. & Di Crescenzo, A. (2018). A residual inaccuracy measure based on the relevation transform. *Metrika* **81**(1): 37–59.
[33] Rao, M. (2005). More on a new concept of entropy and information. *Journal Theoretical Probability* **18**(4): 967–981.
[34] Rao, M., Chen, Y., Vemuri, B.C. & Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE Transactions on Information Theory* **6**: 1220–1228.
[35] Sarda, P. (1993). Smoothing parameter selection for smooth distribution functions. *Journal of Statistical Planning and Inference* **35**(1): 65–75.
[36] Shannon, C.E. (1948). A mathematical theory of communication. *The Bell System Technical journal* **27**(3): 379–423.
[37] Yang, J., Xia, W. & Hu, T. (2019). Bounds on extropy with variational distance constraint. *Probability in the Engineering and Informational Sciences* **33**(2): 186–204.

Cite this article: Hashempour M., Mohammadi M., Jahanshahi SMA. and Khammar AH. (2025). Extropy-based dynamic cumulative residual inaccuracy measure: properties and applications. *Probability in the Engineering and Informational Sciences* 39(4): 461–485. https://doi.org/10.1017/S0269964824000299

https://doi.org/10.1017/S0269964824000299 Published online by Cambridge University Press