Probability in the Engineering and Informational Sciences (2026), **40:**2 272–287
doi:10.1017/S0269964825100132

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Non-parametric estimation of the generalized past entropy function under α-mixing sample

Radhakumari Maya[1], Muhammed Rasheed Irshad[1], Francesco Buono[2] and Maria Longobardi[3]

[1]Department of Statistics, Cochin University of Science and Technology, Cochin, Kerala, India
[2]Institute of Statistics, RWTH Aachen University, Aachen, Germany. Now at Dipartimento di Matematica e Applicazioni, Università degli Studi di Napoli Federico II, Naples, Italy
[3]Dipartimento di Matematica e Applicazioni, Università degli Studi di Napoli Federico II, Naples, Italy
Corresponding author: Maria Longobardi; Email: maria.longobardi@unina.it

Keywords: Generalized past entropy function, α-mixing, Kernel estimator, Past entropy function, Residual entropy function

MSC: 62B10; 62G20; 94A17

Abstract

Measure of uncertainty in past lifetime distribution plays an important role in the context of information theory, forensic science and other related fields. In the present work, we propose non-parametric kernel type estimator for generalized past entropy function, which was introduced by Gupta and Nanda [9], under α-mixing sample. The resulting estimator is shown to be weak and strong consistent and asymptotically normally distributed under certain regularity conditions. The performance of the estimator is validated through simulation study and a real data set.

1. Introduction

The concept of entropy was first introduced and thoroughly explored by Shannon [19], an electrical engineer at Bell Telephone Laboratories, in his research on communication networks. Around the same time, Wiener [22] independently investigated the concept in his work on Cybernetics, albeit with a different motivation. Let X be a non-negative rv with an absolutely continuous cumulative distribution function (cdf) $F(x) = P[X \leq x]$ and with probability density function (pdf) $f(x)$. Then the Shannon entropy associated with X is defined as

$$H(X) = -\int_0^{+\infty} f(x) \log f(x)\, dx. \tag{1.1}$$

Starting from the pioneering work of Shannon [19], different researchers have shown applications of entropy in different fields. Apart from thermodynamics and information theory, application of Shannon entropy varies over diverse fields such as statistics, economics, finance, psychology, wavelet analysis, image recognition, computer science, fuzzy sets and so on. In spite of its well known applications, it is necessary to modify this measure under different scenarios, specifically the concept of residual entropy proposed by Ebrahimi [7] substantiate this claim.

If the system has already survived for some units of time, Shannon entropy will not be applicable in such cases which led to the development of residual entropy. Residual entropy is the Shannon entropy associated with the rv $[X - t|X > t]$, $t \geq 0$, and is defined as (see, Ebrahimi [7])

© 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/S0269964825100132 Published online by Cambridge University Press

$$H(f;t) = -\int_t^{+\infty} \frac{f(x)}{\overline{F}(t)} \log\left(\frac{f(x)}{\overline{F}(t)}\right) dx, \tag{1.2}$$

where $\overline{F}(t)$ denotes the survival function. Given that an item has survived up to time t, $H(f;t)$ measures the uncertainty about its remaining life.

Ebrahimi [7] used this concept to define new ageing classes such as decreasing (increasing) uncertainty of residual life (DURL (IURL)). Belzunce et al. [3] proposed kernel-type estimation of the residual entropy function in the case of independent complete data sets. Belzunce et al. [4] showed that for a rv X with an increasing residual entropy $H(f;t)$, the function $H(f;t)$ uniquely determines the cdf. Rajesh et al. [15] thoroughly elucidated the necessity of developing inferential aspects under dependent sample and proposed non-parametric kernel type estimators for the residual entropy function based on α-mixing dependent sample.

Di Crescenzo and Longobardi [6] discussed the necessity of developing the uncertainty measure based on the reversed residual life (past lifetime) and they developed the concept of past entropy. If X denotes the lifetime of a component, a system or a living organism, then the past entropy of X at time t is defined as

$$\overline{H}(f;t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)}\, dx. \tag{1.3}$$

This measure has become significant in forensic science, particularly in determining the exact time of failure (or death, in the case of humans) when a unit is found in a failed state at some time t. It also has applications in actuarial science, as discussed by [18]. Maya [12] introduced non-parametric estimators for this measure, applicable to both complete and censored α-mixing dependent samples.

Numerous researchers have explored various definitions of entropy and their associated properties. Balakrishnan et al. [2] established relationships between certain cumulative entropies and the moments of order statistics. Furthermore, a general formulation of entropy was presented in Balakrishnan et al. [1]. Building on the successful applications of the past entropy function, Gupta and Nanda [9] extended (1.3) to define the past entropy function of order β. Now the past entropy function of order β is defined as

$$\overline{H}^{\beta}(f;t) = \frac{1}{(1-\beta)} \left\{ \log \int_0^t \left(\frac{f(x)}{F(t)}\right)^{\beta} dx \right\}, \text{ for } \beta \neq 1, \beta > 0. \tag{1.4}$$

Later on Nanda and Paul [13] elucidated certain ordering and ageing properties and some characterization results based on this generalized past entropy measure. As stated by Ebrahimi [7], systems or components with high uncertainty tend to exhibit lower reliability. Building on this concept, these measures can be effectively used to choose the most reliable system among competing models.

However, in terms of inferential aspects, no studies appear to have been conducted in the existing literature. This gap motivates the authors to explore the development of a non-parametric estimator for the generalized past entropy function using kernel-based estimation. The study focuses on cases where the observations exhibit dependence. More specifically, the proposed estimator is based on an α-mixing dependent sample (see Rosenblatt [17]), and its definition is provided below.

Definition 1.1. *Let* $\{X_i; i \geq 1\}$ *denote a sequence of rvs. Given a positive integer n, set*

$$\alpha(n) = \sup_{k \geq 1} \{|P(A \cap B) - P(A)P(B)|; A \epsilon \mathfrak{F}_1^k, B \epsilon \mathfrak{F}_{k+n}^{+\infty}\}, \tag{1.5}$$

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

where $\mathfrak{F}_i^k$ *denote the* σ*- field of events generated by* $\{X_j; i \le j \le k\}$*. The sequence is said to be* α*-mixing (strong mixing) if the mixing coefficient* $\alpha(n) \to 0$ *as* $n \to +\infty$*.*

Among various mixing conditions, α-mixing is reasonably weak and has many practical applications. In the same way, Irshad et al. [10] have proposed non-parametric kernel type estimators for two versions of cumulative residual Tsallis entropies.

The rest of the article is organized as follows. In Section 2, we propose a non-parametric kernel type estimator of $\overline{H}^{\beta}(f;t)$ given in (1.4) by α-mixing dependent sample. Asymptotic properties of the proposed estimator are discussed in Section 3. The performance of the estimator is validated using simulation study in Section 4. In Section 5, a real data set is used to illustrate the performance of the proposed estimator and then the asymptotic normality is analyzed by the use of histograms. Finally, the article is concluded in Section 6.

2. Non-parametric estimation of generalized past entropy function

In this section, we propose a non-parametric kernel type estimator for the generalized past entropy function.

Let $\{X_i; 1 \le i \le n\}$ be a sequence of identically distributed rvs. Note that $X_i's$ need not be mutually independent, that is, the observations are supposed to be α-mixing. Rosenblatt [16] and Parzen [14] proposed a non-parametric kernel type estimator of $f(x)$ that is given by

$$f_n(x) = \frac{1}{nb_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{b_n}\right), \tag{2.1}$$

where $K(x)$ is a kernel of order s satisfying the conditions:

a. $K(x) \ge 0$ for all x,
b. $\int_{-+\infty}^{+\infty} K(x)dx = 1$,
c. $K(\cdot)$ is symmetric about zero and satisfies the Lipschitz condition, that is, there exists a constant M such that $|K(x) - K(y)| \le M|x - y|$,
d. $K_n(x) = \frac{1}{b_n} K\left(\frac{x}{b_n}\right)$, where $\{b_n\}$ is a bandwidth sequence of positive numbers such that $b_n \to 0$ and $nb_n \to +\infty$ as $n \to +\infty$.

The expressions for the bias and variance of $f_n(x)$ under α-mixing dependence sample are respectively given by

$$Bias\,(f_n(x)) \backsimeq \frac{b_n^s \Delta_s}{s!} f^{(s)}(x) \tag{2.2}$$

and

$$Var\,(f_n(x)) \backsimeq \frac{f(x)}{nb_n} \Delta_K, \tag{2.3}$$

where $\Delta_s = \int_{-+\infty}^{+\infty} u^s K(u)\, du$, $\Delta_K = \int_{-+\infty}^{+\infty} K^2(u)\, du$ and $f^{(s)}(\cdot)$ denotes the s-th derivative of f.

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

Based on (2.1), we propose a non-parametric kernel type estimator for $\overline{H}^{\beta}(f;t)$ that is defined as

$$\overline{H}_n^{\beta}(f;t) = \frac{1}{(1-\beta)} \left\{ \log \int_0^t f_n^{\beta}(x)\, dx - \beta \log F_n(t) \right\}, \tag{2.4}$$

where $f_n(x)$ is a non-parametric estimator of $f(x)$ and is given in (2.1) and $F_n(t) = \int_0^t f_n(x)dx$ is a non-parametric estimator of cdf $F(t)$. The bias and variance of $F_n(t)$ are respectively given by (see, Maya [12])

$$Bias\,(F_n(t)) \backsimeq \frac{b_n^s \Delta_s}{s!} \int_0^t f^{(s)}(x)\, dx \tag{2.5}$$

and

$$Var\,(F_n(t)) \backsimeq \frac{1}{nb_n} \Delta_K F(t). \tag{2.6}$$

3. Asymptotic properties of generalized past entropy function

In this section, some asymptotic properties of generalized past entropy function are established.

For computational simplicity, define the following

$$a_n(t) = \log \int_0^t f_n^{\beta}(x)\, dx, a(t) = \log \int_0^t f^{\beta}(x)\, dx, \tag{3.1}$$

and

$$m_n(t) = \log F_n(t) \text{ and } m(t) = \log F(t). \tag{3.2}$$

Therefore,

$$\overline{H}^{\beta}(f;t) = \frac{1}{(1-\beta)} \left[a(t) - \beta\, m(t) \right] \tag{3.3}$$

and

$$\overline{H}_n^{\beta}(f;t) = \frac{1}{(1-\beta)} \left[a_n(t) - \beta\, m_n(t) \right]. \tag{3.4}$$

In the following theorems, the weak and strong consistency properties of $\overline{H}_n^{\beta}(f;t)$ are proved.

Theorem 3.1. *Suppose $\overline{H}_n^{\beta}(f;t)$ is a non-parametric estimator of $\overline{H}^{\beta}(f;t)$ defined in (2.4) satisfying the assumptions given in Section 2. Then $\overline{H}_n^{\beta}(f;t)$ is a consistent estimator of $\overline{H}^{\beta}(f;t)$.*

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

Proof. By using Taylor's series expansion, we have

$$\log \int_0^t f_n^{\beta}(x)\,dx \backsimeq \log \int_0^t f^{\beta}(x)\,dx + (f_n(t) - f(t)) \frac{\int_0^t \beta f^{\beta-1}(x)dx}{\int_0^t f^{\beta}(x)dx}.$$

Using (3.1), we get

$$a_n(t) \backsimeq a(t) + (f_n(t) - f(t))u(t), \tag{3.5}$$

where

$$u(t) = \frac{\int_0^t \beta f^{\beta-1}(x)dx}{\int_0^t f^{\beta}(x)dx}.$$

Also,

$$\log F_n(t) \backsimeq \log F(t) + \frac{(F_n(t) - F(t))}{F(t)}.$$

From (3.2), we write

$$m_n(t) \backsimeq m(t) + \frac{(F_n(t) - F(t))}{F(t)}, \tag{3.6}$$

Using (3.5) and (3.6), the bias and variance of $a_n(t)$ and $m_n(t)$ are given by

$$Bias\,(a_n(t)) \backsimeq \frac{b_n^s}{s!}\Delta_s f^{(s)}(t)u(t), \tag{3.7}$$

$$Var\,(a_n(t)) \backsimeq \frac{1}{nb_n}\Delta_K f(t)u^2(t), \tag{3.8}$$

$$Bias\,(m_n(t)) \backsimeq \frac{b_n^s \Delta_s}{s!F(t)} \int_0^t f^{(s)}(x)\,dx, \tag{3.9}$$

and

$$Var\,(m_n(t)) \backsimeq \frac{1}{nb_n}\frac{\Delta_K}{F(t)}. \tag{3.10}$$

From (3.7) and (3.8), as $n \to +\infty$

$$\text{MSE}\,(a_n(t)) \to 0.$$

Therefore, the estimator $a_n(t)$ is consistent (in the probability sense), i.e.,

$$a_n(t) \xrightarrow{p} a(t).$$

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

From (3.9) and (3.10), as $n \to +\infty$

$$\text{MSE}\,(m_n(t)) \to 0.$$

Therefore, the estimator $m_n(t)$ is consistent (in the probability sense), i.e.,

$$m_n(t) \xrightarrow{p} m(t).$$

Therefore

$$\overline{H}_n^{\beta}(f;t) = \frac{1}{(1-\beta)}\,[a_n(t) - \beta\, m_n(t)] \xrightarrow{p} \frac{1}{(1-\beta)}\,[a(t) - \beta\, m(t)] = \overline{H}^{\beta}(f;t).$$

That is, $\overline{H}_n^{\beta}(f;t)$ is a consistent (in the probability sense) estimator of $\overline{H}^{\beta}(f;t)$. □

Theorem 3.2. *Let $\overline{H}_n^{\beta}(f;t)$ be a non-parametric estimator of $\overline{H}^{\beta}(f;t)$ defined in (2.4) satisfying the assumptions given in Section 2. Suppose the kernel $K(\cdot)$ satisfies the requirements:*

a. $K(u) \to 0$ *as* $|u| \to +\infty$,
b. $\int_{-\infty}^{+\infty} |K'(u)|du < +\infty$,
c. $\int_{-\infty}^{+\infty} |u|K(u)\,du < +\infty$.

Let $\bar{t} = \sup\{t \in \mathbb{R}; F(t) < 1\}$ *and J be any compact subset of* $(0,\ \bar{t})$. *Then*

$$\lim_{n\to+\infty} \sup_{t\in J} \left|\overline{H}_n^{\beta}(f;t) - \overline{H}^{\beta}(f;t)\right| = 0\ a.s.$$

Proof. By direct calculations, we get

$$\overline{H}_n^{\beta}(f;t) - \overline{H}^{\beta}(f;t) = \frac{1}{(1-\beta)}\,(a_n(t) - a(t)) - \frac{\beta}{(1-\beta)}\,(m_n(t) - m(t))\,.$$

$$|\overline{H}_n^{\beta}(f;t) - \overline{H}^{\beta}(f;t)| \backsimeq \frac{u(t)}{(1-\beta)}\,|f_n(t) - f(t)| + \frac{\beta}{(1-\beta)}\left(\frac{|F_n(t) - F(t)|}{1 - F(t)}\right).$$

By using Theorem 3.3 and Theorem 4.1 of Cai and Roussas [5], the result is immediate. Thus, we conclude that $\overline{H}_n^{\beta}(f;t)$ is a strong consistent estimator of $\overline{H}^{\beta}(f;t)$. □

In order to prove that $\overline{H}_n^{\beta}(f;t)$ is an integratedly uniformly consistent in quadratic mean estimator of $\overline{H}^{\beta}(f;t)$, we consider the following definition from Wegman [21].

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

Definition 3.1. *The density estimator $f_n(x)$ is said to be integratedly uniformly consistent in quadratic mean if the mean integrated squared error (MISE) approaches zero, i.e.,*

$$\lim_{n\to+\infty} \mathbb{E}\left[\int (f_n(x) - f(x))^2\, dx\right] = 0$$

In the following theorem, we prove that $\overline{H}_n^{\beta}(f;t)$ is integratedly uniformly consistent in quadratic mean estimator of $\overline{H}^{\beta}(f;t)$.

Theorem 3.3. *Suppose $\overline{H}_n^{\beta}(f;t)$ is a kernel estimator of $\overline{H}^{\beta}(f;t)$ as defined in (2.4). Then, $\overline{H}_n^{\beta}(f;t)$ is integratedly uniformly consistent in quadratic mean estimator of $\overline{H}^{\beta}(f;t)$.*

Proof. MISE of $\overline{H}_n^{\beta}(f;t)$ is given by

$$\begin{aligned}
\mathrm{MISE}(\overline{H}_n^{\beta}(f;t)) &= \mathrm{E}\int_0^{+\infty}\left[\overline{H}_n^{\beta}(f;t) - \overline{H}^{\beta}(f;t)\right]^2 dt \\
&= \int_0^{+\infty}\left[[\mathrm{Bias}(\overline{H}_n^{\beta}(f;t))]^2 + \mathrm{Var}(\overline{H}_n^{\beta}(f;t))\right] dt \\
&= \int_0^{+\infty}\left(\frac{b_n^s \Delta_s}{(1-\beta)s!}\left(f^{(s)}(t)u(t) - \frac{\beta}{F(t)}\int_0^t f^{(s)}(x)dx\right)\right)^2 dt \\
&+ \int_0^{+\infty}\left(\frac{\Delta_K}{nb_n(1-\beta)^2}\left(f(t)u^2(t) + \frac{\beta^2}{F(t)}\right)\right) dt.
\end{aligned} \tag{3.11}$$

We have, as $n \to +\infty$,

$$\mathrm{MSE}\left(\overline{H}_n^{\beta}(f;t)\right) = Bias(\overline{H}_n^{\beta}(f;t))]^2 + \mathrm{Var}(\overline{H}_n^{\beta}(f;t)) \to 0.$$

Therefore, from (3.11), we have

$$\mathrm{MISE}\left(\overline{H}_n^{\beta}(f;t)\right) \to 0, \text{as } n \to +\infty. \tag{3.12}$$

From (3.12), we can say that $\overline{H}_n^{\beta}(f;t)$ is integratedly uniformly consistent in quadratic mean estimator of $\overline{H}^{\beta}(f;t)$.

Thus the theorem is proved. □

In the following theorem, we obtained the optimal bandwidth.

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

Theorem 3.4. *Suppose $\overline{H}_n^{\beta}(f;t)$ is a non-parametric estimator of $\overline{H}^{\beta}(f;t)$ defined in (2.4) satisfying the assumptions given in Section 2. Then the optimal bandwidth is given by*

$$b_n^{'} = \left[\frac{\frac{1}{n}\int_0^{+\infty} \Delta_K \left(f(t)u^2(t) + \frac{\beta^2}{F(t)} \right) dt}{2s \int_0^{+\infty} \left(\frac{\Delta_s}{s!} \left(f^{(s)}(t)u(t) - \frac{\beta}{F(t)} \int_0^t f^{(s)}(x)dx \right) \right)^2 dt} \right]^{\frac{1}{2s+1}}. \tag{3.13}$$

Proof. By using (3.11), the asymptotic-MISE (A-MISE) is given by

$$\text{A-MISE}(\overline{H}_n^{\beta}(f;t)) = b_n^{2s} \int_0^{+\infty} \left(\frac{\Delta_s}{(1-\beta)s!} \left(f^{(s)}(t) - \frac{\beta}{F(t)} \int_0^t f_{(x)}^{(s)} dx \right) \right)^2 dt + \frac{1}{nb_n} \int_0^{+\infty} \frac{\Delta_K}{(1-\beta)^2} \left(f(t)u^2(t) + \frac{\beta^2}{F(t)} \right) dt.$$

By minimizing $\text{MISE}(\overline{H}_n^{\beta}(f;t))$ with respect to the parameter b_n, we get the optimal bandwidth $b_n^{'}$.

$\frac{\partial \text{A-MISE}(\overline{H}_n^{\beta}(f;t))}{\partial b_n} = 0 \Rightarrow$

$$2sb_n^{2s-1} \int_0^{+\infty} \left\{ \frac{\Delta_s}{(1-\beta)s!} \left(f^{(s)}(t) - \frac{\beta}{F(t)} \int_0^t f_{(x)}^{(s)} dx \right) \right\}^2 dt = \frac{1}{nb_n^2} \int_0^{+\infty} \frac{\Delta_K}{(1-\beta)^2} \left(f(t)u^2(t) + \frac{\beta^2}{F(t)} \right) dt.$$

$\Rightarrow$

$$b_n^{2s+1} = \frac{\frac{1}{n}\int_0^{+\infty} \frac{\Delta_K}{(1-\beta)^2} \left(f(t)u^2(t) + \frac{\beta^2}{F(t)} \right) dt}{2s \int_0^{+\infty} \left\{ \frac{\Delta_s}{(1-\beta)s!} \left(f^{(s)}(t) - \frac{\beta}{F(t)} \int_0^t f_{(x)}^{(s)} dx \right) \right\}^2 dt}.$$

Therefore,

$$b_n^{'} = \left[\frac{\int_0^{+\infty} \Delta_K \left(f(t)u^2(t) + \frac{\beta^2}{F(t)} \right) \mathrm{d}t}{2s \int_0^{+\infty} \left\{ \frac{\Delta_s}{s!} \left(f^{(s)}(t) - \frac{\beta}{F(t)} \int_0^t f_{(x)}^{(s)} dx \right) \right\}^2 \mathrm{d}t} \right]^{\frac{1}{2s+1}} n^{-\frac{1}{2s+1}} = O\left(n^{-\frac{1}{2s+1}} \right), s = 1, 2, \ldots.$$

□

The following lemma is used to derive the asymptotic normality of $\overline{H}^{\beta}(f;t)$ and is established in Theorem 3.4.

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

Lemma 3.1. *Let $K(x)$ be a kernel of order s, let $\{\alpha(i)\}$ be the mixing coefficients and let $\{b_n\}$ be a sequence of numbers satisfying the assumptions given in Section 2. Then, for all fixed point x, with $f(x) > 0$,*

$$(nb_n)^{\frac{1}{2}} \left\{ \frac{(f_n(x) - f(x))}{\sigma_f} \right\} \tag{3.14}$$

has a standard normal distribution as $n \to +\infty$ with $\sigma_f^2 \backsimeq \frac{1}{n} f(x)\Delta_k$.

Proof. The proof follows directly by applying the same steps as in the proof of Theorem 8 in Masry [11]. □

Theorem 3.5. *Let $\overline{H}_n^{\beta}(f;t)$ be a non-parametric estimator of $\overline{H}^{\beta}(f;t)$, $K(x)$ be a kernel of order s satisfying the assumptions given in Section 2. Then for fixed t,*

$$\sqrt{nb_n} \left\{ \frac{(\overline{H}_n^{\beta}(f;t) - \overline{H}^{\beta}(f;t))}{\sigma_{\overline{H}}} \right\} \tag{3.15}$$

has a standard normal distribution as $n \to +\infty$ with

$$\sigma_{\overline{H}}^2 \backsimeq \frac{\Delta_K}{nb_n(1-\beta)^2} \left\{ f(t)u^2(t) + \frac{\beta^2}{F(t)} \right\}. \tag{3.16}$$

Proof.

$$\begin{aligned}
\sqrt{nb_n} \left(\overline{H}_n^{\beta}(f;t) - \overline{H}^{\beta}(f;t) \right) &= \frac{1}{(1-\beta)} (nb_n)^{\frac{1}{2}} \left\{ \log \int_0^t f_n^{\beta}(x)\, dx - \log \int_0^t f^{\beta}(x)\, dx \right\} \\
&\quad - \frac{\beta}{(1-\beta)} (nb_n)^{\frac{1}{2}} \{\log F_n(t) - \log F(t)\} \\
&\backsimeq \frac{1}{(1-\beta)} (nb_n)^{\frac{1}{2}} \{(f_n(t) - f(t))u(t)\} \\
&\quad - \frac{\beta}{(1-\beta)} (nb_n)^{\frac{1}{2}} \frac{\int_0^t (f_n(x) - f(x))\, dx}{F(t)}.
\end{aligned}$$

By using the asymptotic normality of $f_n(x)$ given in Lemma 1, the proof is immediate. □

4. Simulation

A simulation study is conducted to evaluate the performance of the proposed estimator $\overline{H}_n^{\beta}(f;t)$. Here the process $\{X_i\}$ is generated form exponential AR(1), with correlation coefficient $\phi = 0.2$ and parameter $\lambda = 1$. The Gaussian kernel is employed as the kernel function for estimation. The estimated value, bias, and MSE of the proposed estimator are calculated for various sample sizes. We vary the values of β and the bandwidth parameter is determined using the plug-in method as proposed by Sheather [20]. The result for the exponential AR(1) process are presented in Table 1.

From Table 1, it is evident that both the *MSE* and bias of the estimators decrease as the sample size increases. The reduction in *MSE* signifies that the estimators' predictions become closer to the true values with larger sample sizes, reflecting improved accuracy and efficiency in estimation. Similarly, the decreasing bias highlights the growing precision of the estimator.

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

Table 1. *Estimated value, bias ad MSE of $\overline{H}_n^{\beta}(f;t)$ for the Exponential AR(1) process along with the corresponding theoretical value $\overline{H}^{\beta}(f;t)$.*

	$\beta = 2$				$\beta = 4$			
n	$\overline{H}_n^{\beta}(f;t)$	bias	*MSE*	$\overline{H}^{\beta}(f;t)$	$\overline{H}_n^{\beta}(f;t)$	bias	*MSE*	$\overline{H}^{\beta}(f;t)$
100	−0.05474	0.02405	0.00239	−0.07879	−0.08847	0.05484	0.00693	−0.14331
200	−0.05302	0.02577	0.00152		−0.09441	0.04889	0.00478	
300	−0.05726	0.02153	0.00119		−0.09913	0.04418	0.00393	
400	−0.0579	0.02089	0.00096		−0.10095	0.04236	0.00328	
	$\beta = 6$				$\beta = 8$			
100	−0.11872	0.07284	0.01181	−0.19156	−0.1378	0.08929	0.01524	−0.22709
200	−0.12471	0.06686	0.00823		−0.14546	0.08163	0.01131	
300	−0.12888	0.06268	0.00677		−0.15353	0.07356	0.00928	
400	−0.13256	0.059	0.006		−0.15745	0.06963	0.00785	

Table 2. *Bootstrap bias and mean-squared error of $\overline{H}_n^{\beta}(f;t)$ for the real data set.*

t	Bias	MSE	t	Bias	MSE
2.0	−0.0713	0.1461	5.0	−0.0307	0.0133
2.5	0.0041	0.0299	5.5	−0.0873	0.0196
3.0	−0.0232	0.0668	6.0	−0.0785	0.0217
3.5	0.0112	0.0187	6.5	−0.0470	0.0178
4.0	−0.0786	0.0243	7.0	−0.1472	0.0383
4.5	−0.0428	0.0661			

5. Numerical examples

To demonstrate the practical applicability of the proposed estimator discussed in Section 2, we analyze real-world data consisting of the failure times (in months) of 20 electric carts used for internal transport and delivery in a large manufacturing facility (see Zimmer et al. [23]). The bootstrapping procedure is employed to determine the optimal value of b_n (see Efron [8]).

For estimation, we utilize the Gaussian kernel function:

$$K(z) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right).$$

At each value of t, we compute the biases and mean-squared errors of $\overline{H}_n^{\beta}(f;t)$ using 250 bootstrap samples of size 20. Table 2 presents the bootstrap biases and mean-squared errors for the estimator $\overline{H}_n^{\beta}(f;t)$. Figure 1 compares the theoretical value $\overline{H}^{\beta}(f;t)$ with the estimated value $\overline{H}_n^{\beta}(f;t)$. From Figure 1, it is evident that for the given dataset, the generalized past entropy function increases over time.

Next, we aim to examine the asymptotic normality of the estimator presented in Theorem 3.5. To achieve this, we conduct a numerical evaluation of $\overline{H}_n^{\beta}(f;t)$ and analyze it through Monte Carlo simulations. Let X follow an exponential distribution with parameter λ, where the mean is given by $\mu = 1/\lambda$. Then, the past entropy function of order β is expressed as

$$\overline{H}^{\beta}(f;t) = \frac{1}{1-\beta}\left[(\beta-1)\log\lambda - \log\beta - \beta\log(1-e^{-\lambda t}) + \log(1-e^{-\lambda\beta t})\right]. \tag{5.1}$$

Before obtaining our estimator, it is necessary to fix a function K and a sequence $\{b_n\}_{n\in\mathbb{N}}$ which satisfy the assumptions given in Section 2. Here, we consider

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x \in \mathbb{R} \tag{5.2}$$

$$b_n = \frac{1}{\sqrt{n}}, \quad n \in \mathbb{N}. \tag{5.3}$$

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

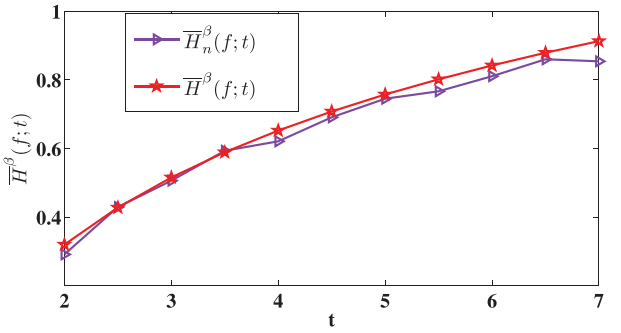
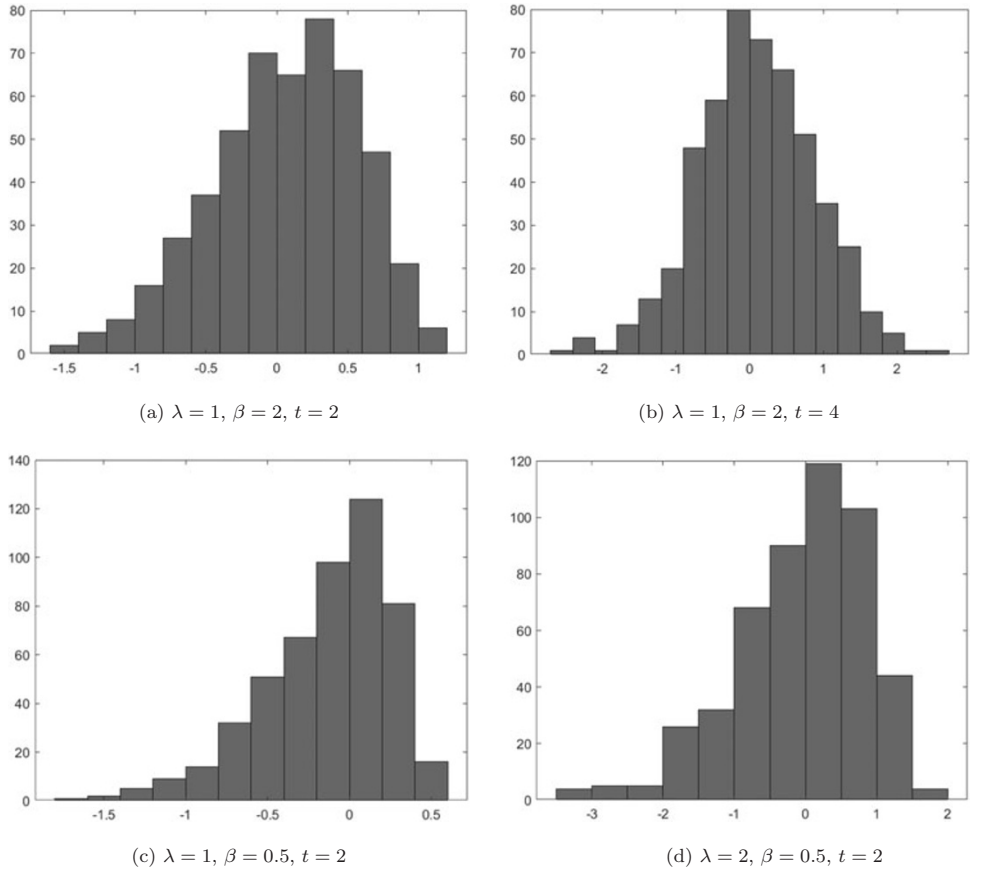


Figure 1. *Plots of estimates of generalized past entropy function for the first failure of 20 electric carts.*



(a) $\lambda = 1$, $\beta = 2$, $t = 2$ (b) $\lambda = 1$, $\beta = 2$, $t = 4$

(c) $\lambda = 1$, $\beta = 0.5$, $t = 2$ (d) $\lambda = 2$, $\beta = 0.5$, $t = 2$

Figure 2. *Histograms of (3.15) with parameters given in Section 5 and different choices of α, β and t. a) $\lambda = 1$, $\beta = 2$, $t = 2$ b) $\lambda = 1$, $\beta = 2$, $t = 4$ c) $\lambda = 1$, $\beta = 0.5$, $t = 2$ d) $\lambda = 2$, $\beta = 0.5$, $t = 2$.*

By using these assumptions, we obtain

$$\Delta_K = \frac{1}{2\sqrt{\pi}},$$

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

and, for the exponential distribution, we have

$$u(t) = \frac{\beta^2}{\lambda(\beta-1)} \frac{1-e^{-\lambda(\beta-1)t}}{1-e^{-\lambda\beta t}}.$$

We aim to verify whether the quantity defined in (3.15) follows a standard normal distribution. Using the `exprnd` function in *MATLAB*, we generate 500 samples, each of size $N = 50$, from an exponential distribution with parameter λ. For each simulated sample, we extract only the values that do not exceed a fixed threshold t, and the count of these selected values is used as n in our evaluations. Subsequently, using the fixed parameters, we compute the quantity given in (3.15) for each sample. To assess its asymptotic normality, we construct a histogram based on the 500 computed values. Specifically, we explore various choices for λ, β, and t, as illustrated in Figure 2.

6. Conclusion

This paper examines non-parametric estimators for the generalized past entropy function using kernel-based estimation, with α-mixing dependent observations. The asymptotic properties of the proposed estimator are analyzed and established. Additionally, a simulation study and real data analysis are carried out to evaluate the performance of the proposed estimator. We obtained that the estimator is performing good on both simulation study and data analysis owing to bias and *MSE*. As discussed earlier these estimate can be used in selecting a reliable system from the other available competing models.

Future research will extend beyond the α-mixing dependence condition to explore the inferential properties of generalized past entropy function under ϕ-mixing and ρ-mixing dependence conditions.

Declaration of interests. 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.

Acknowledgements Francesco Buono and Maria Longobardi are members of the research group GNAMPA of INdAM (Istituto Nazionale di Alta Matematica). ML has been partially supported by MIUR—PRIN 2022 PNRR, project "Stochastic Modeling of Compound Events", no. P2022KZJTZ, funded by European Union—Next Generation EU.

References

1. N. Balakrishnan, F. Buono and M. Longobardi. A unified formulation of entropy and its application. *Physica A: Statistical Mechanics and its Applications*. **596** (2022), 127214.
2. N. Balakrishnan, F. Buono and M. Longobardi. On Cumulative Entropies in Terms of Moments of Order Statistics. *Methodology and Computing in Applied Probability*. **24** (2022), 345–359.
3. F. Belzunce, A. Guillamón, J. Navarro and J. M. Ruiz. Kernel estimation of residual entropy. *Communications in Statistics—Theory and Methods*. **30** (2001), 1243–1255.
4. F. Belzunce, J. Navarro, J. M. Ruiz and Y. del Aguila. Some results on residual entropy function. *Metrika*. **59** (2004), 147–161.
5. Z. Cai and G. G. Roussas. Uniform strong estimation under α-mixing, with rates. *Statistics and Probability Letters*. **15** (1992), 47–55.
6. A. Di Crescenzo and M. Longobardi. Entropy-based measure of uncertainty in past lifetime distributions. *Journal of Applied Probability*. **39** (2002), 434–440.
7. N. Ebrahimi. How to measure uncertainty in the residual life time distribution. *Sankhya A*. **58** (1996), 48–56.
8. B. Efron. Censored data and the bootstrap. *Journal of American Statistical Association*. **76** (1981), 312–319.
9. R. D. Gupta and A. K. Nanda. α- and β-entropies and relative entropies of distributions. *Journal of Statistical Theory and Applications*. **1** (2002), 177–190.
10. M. R. Irshad, R. Maya, F. Buono and M. Longobardi. Kernel Estimation of Cumulative Residual Tsallis Entropy and Its Dynamic Version under ρ-Mixing Dependent Data. *Entropy*. **24** (2022), 9.
11. E. Masry. Recursive probability density estimation for weakly dependent stationary processes. *IEEE*. **32** (1986), 254–267.
12. R. Maya. Kernel estimation of the past entropy function with dependent data. *Journal of the Kerala Statistical Association*. **24** (2013), 12–36.
13. A. K. Nanda and P. Paul. Some results on generalized past entropy. *Journal of Statistical Planning and Inference*. **136** (2006), 3659–3674.
14. E. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*. **33** (1962), 1065–1076.

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

15. G. Rajesh, E.I Abdul-Sathar, R. Maya and K. R. M. Nair. Nonparametric estimation of the residual entropy function with censored dependent data. *Brazilian Journal of Probability and Statistics*. **29** (2015), 866–877.
16. M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*. **27** (1956), 832–837.
17. M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*. **42** (1956), 43–47.
18. A. Sachlas and T. Papaioannou. Residual and past entropy in actuarial science and survival models. *Methodology and Computing in Applied Probability*. **16** (2014), 79–99.
19. C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal* **27**(3) (1948), 379–423, 623–656.
20. S. J. Sheather and M. C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*. **53** (1991), 683–90.
21. E. J. Wegman. Nonparametric probability density estimation: I. A summary of available methods. *Technometrics*. **14** (1972), 533–546.
22. N. Wiener. Cybernetics: or control and communication in the animal and the machine. *John Wiley*. (1948).
23. W.I Zimmer, J. B. Keats and F. K. Wang. The Burr xii distribution in reliability analysis. *Journal of Quality Technology*. **30** (1998), 386–394.

Appendix

R-code to find $\overline{H}_n^{\beta}(f;t)$ by simulation for size n=100.

```
rm(list = ls())
library(MASS)
library(extremefit)
library(xlsx)
library(stats)

# Upper limit of integration
t <- 1

# PDF of the exponential distribution with rate = 1.5
f <- function(x) {
dexp(x, rate = 1)
}

# CDF of the exponential distribution at t
F_t <- pexp(t, rate = 1)

betas <- c( 2,4, 6, 8, 10) # Add more beta values as needed
results_list <- list()

for (beta in betas) {
r2 <- NULL # Reset results for each beta

n <- 100
r <- 1000
t <- 1

Hhat <- numeric(r)

for (s in 1:r) {

print(c(beta, k, s))
#set.seed(10)
```

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

```
# Generate data for the process
lambda <- 1
phi <- 0.2

X <- numeric(n)
a <- numeric(n)

for (i in 2:(n + 1)) {
a[i] <- sample(c(0, rexp(1, lambda)), prob = c(phi, 1 - phi), size = 1)
X[i] <- phi * X[i - 1] + a[i]
}

X <- X[-1]

# Bandwidth estimation
h <- bw.SJ(X)
# The integrand function: (f(x) / F_t)^beta
integrand <- function(x) {
(f(x) / F_t)^beta
}

# Integrate the function from 0 to t, passing beta
integral_result <- integrate(integrand, lower = 0, upper = t)

# Extract the integral value
integral_value <- integral_result$value

# PDF approximation using a kernel density estimator
int <- function(x) {
fn <- numeric(length(x))
for (i in 1:length(x)) {
fn[i] <- mean(dnorm((x[i] - X) / h) / h)
}
return(fn)
}

# Compute the non-parametric CDF Fn(t)
compute_Fn <- function(data, t) {
integrate(Vectorize(int), lower = 0, upper = t)$value
}

# Compute the integral of f_n(x)^beta from 0 to t
compute_integral <- function(data, t, beta){
integrand <- function(x) int(x)^beta
integrate(Vectorize(integrand), lower = 0, upper = t)$value
}

# Function to compute H^beta_n(f; t)
H_beta_n <- function(data, t, beta) {
integral_value <- compute_integral(data, t, beta)
```

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press

```
Fn_t <- compute_Fn(data, t)

result <- (log(integral_value) - beta * log(Fn_t))
return(result)
}

# Compute H^beta_n(f; t)
Hhat[s] <- (1 / (1 - beta)) * H_beta_n(X, t, beta)
}

# Compute the estimator and related metrics
est <- (1 / (1 - beta)) * log(integral_value)
H <- mean(Hhat)
Mse <- (est - Hhat)^2
bias <- H - est
MSE <- mean(Mse)
}
```

Cite this article: Maya R., Irshad MR., Buono F. and Longobardi M. (2026). Non-parametric estimation of the generalized past entropy function under α-mixing sample. *Probability in the Engineering and Informational Sciences* 40(2): 272–287. https://doi.org/10.1017/S0269964825100132

https://doi.org/10.1017/S0269964825100132 Published online by Cambridge University Press