Shahid Bahonar University of Kerman
Journal of Mahani Mathematical Research
Print ISSN: 2251-7952 Online ISSN: 2645-4505

# NONPARAMETRIC KERNEL ESTIMATION OF TSALLIS-TYPE CUMULATIVE RESIDUAL ENTROPY UNDER LENGTH-BIASED LIFETIME DATA

M. Ajami ✉ and R. Zamini

Article type: Research Article
(Received: 05 January 2026, Received in revised form 09 March 2026)
(Accepted: 05 May 2026, Published Online: 05 May 2026)

Abstract. We develop nonparametric inference methods for Tsallis-based cumulative residual entropy functionals when dealing with length-biased lifetime observations. The paper introduces kernel-type estimators for both the static measure and its time-dependent version, with explicit corrections for the sampling mechanism that systematically oversamples longer-lived units. We derive large-sample approximations for bias and variance under standard smoothness assumptions and appropriate bandwidth choices, establish weak and $L^2$-consistency, and prove central limit theorems. Numerical experiments using exponential and Weibull distributions examine finite-sample behavior through bias, variance, and mean squared error calculations, while normality diagnostics validate the asymptotic approximations. We also apply the methodology to automotive component durability data, where results confirm the stable performance and practical value in realistic length-biased scenarios where standard sampling assumptions break down.

*Keywords*: Asymptotic normality; Tsallis entropy functionals; Dynamic entropy measures; Weighted kernel estimation, Length-biased sampling, Reliability analysis.
*2020 MSC*: 62G05, 62G20.

## 1. Introduction

Consider a nonnegative lifetime random variable $X$ with distribution function $F(t)$, density $f(t)$, and survival function $S(t) = 1 - F(t)$. Entropy measures have been used for decades to quantify distributional uncertainty. More recently, researchers have turned their attention to alternative formulations based on survival characteristics, especially when analyzing lifetime phenomena.

One such measure is the cumulative residual entropy (CRE), which quantifies uncertainty in the residual lifetime distribution. For a continuous random

✉ m.ajami@vru.ac.ir, ORCID: 0000-0002-6045-128x
https://doi.org/10.22103/jmmr.2026.26632.1924
Publisher: Shahid Bahonar University of Kerman
How to cite: M. Ajami, R. Zamini, *Nonparametric kernel estimation of Tsallis-type cumulative residual entropy under length-biased lifetime data*, J. Mahani Math. Res. 2026; 15(2): 377-394.

variable $X$, the CRE is

$$\mathcal{C}(X) = -\int_0^\infty S(x) \log S(x)\, dx,$$

where $\log(\cdot)$ is the natural logarithm with standard convention $0 \log 0 = 0$ (see [11]).

This measure connects naturally with reliability concepts like mean residual life. Several extensions have appeared in the literature; see [12], [3], [9], and [18].

[13] introduced the cumulative residual Tsallis entropy (CRTE) of order $\alpha$ to allow more flexible control over tail behavior:

$$\Psi_\alpha(X) = \frac{1}{\alpha - 1}\left(1 - \int_0^\infty S^\alpha(x)\, dx\right), \qquad \alpha \neq 1,\ \alpha > 0. \tag{1}$$

To avoid ambiguity, we clarify that the parameter $\alpha$ appearing in $\Psi_\alpha(X)$ and $S^\alpha(x)$ denotes the same order parameter of the Tsallis entropy, controlling the tail behavior as introduced by [13]. It's straightforward to verify that $\Psi_\alpha(X) \to \mathcal{C}(X)$ as $\alpha \to 1$.

The static form in (1) doesn't account for elapsed operation time, which is often important in reliability applications. A dynamic version at time $t$ is defined as

$$\Psi_\alpha(X \mid t) = \frac{1}{\alpha - 1}\left(1 - \int_t^\infty \frac{S^\alpha(x)}{S^\alpha(t)}\, dx\right), \qquad \alpha \neq 1,\ \alpha > 0. \tag{2}$$

These measures have been applied in physics, engineering, information theory, and reliability studies; see [9], [7], [14], [10], and [15].

In real data collection, simple random sampling often fails. One common violation is length-biased (or size-biased) sampling, where selection probability is proportional to the unit's magnitude. Specifically, if $X$ has density $f$ and mean $\mu$, then under length-biased sampling, the observed random variable $Y$ follows the weighted density

$$g(y) = \frac{y\, f(y)}{\mu}, \quad y \geq 0. \tag{3}$$

If we ignore this sampling mechanism, standard estimation procedures will be biased.

Length-biased data show up naturally in renewal theory [17], reliability studies, survival analysis, and many other areas of applied statistics; see [19] and [2]. Since these schemes systematically oversample longer lifetimes, we need appropriate corrections for valid inference.

With these issues in mind, this paper develops nonparametric kernel estimation methods for $\Psi_\alpha(X)$ and $\Psi_\alpha(X \mid t)$ under length-biased sampling. We use weighted kernel techniques to build smooth estimators and study their large-sample properties, including bias, variance, consistency, integrated $L^2$-consistency, and asymptotic normality.

We organize the rest of the paper as follows. Section 2 introduces the weighted kernel estimators for CRTE and its dynamic version. Section 3 presents our main theoretical results: explicit bias-variance formulas, consistency properties, and central limit theorems. Section 4 reports Monte Carlo simulation results for exponential and Weibull models. Section 5 applies the methodology to automotive component lifetime data. Section 6 concludes and suggests future research directions for parallelism with the actual content.

The following Table 1 provides a summary of the key notations used throughout this paper. This table aims to enhance clarity and facilitate the reader's understanding of the mathematical expressions presented.

Table 1. Notation

| **Notation** | **Expression** |
|---|---|
| $\mathcal{C}(X)$ | The cumulative residual entropy (CRE) for a continuous random variable |
| $S(t)$ | The survival function |
| $\Psi_\alpha(X)$ | The cumulative residual Tsallis entropy (CRTE) of order $\alpha$ |
| $\Psi_\alpha(X \mid t)$ | The dynamic CRTE at time $t$ |
| $\tilde{f}_n(t)$ | Kernel density estimator under length-biased data |
| $\widetilde{S}_n(t)$ | The survival function estimator under length-biased data |
| $\widetilde{\Psi}_\alpha(X)$ | Kernel-based estimator for CRTE |
| $\widetilde{\Psi}_\alpha(X \mid t)$ | Kernel-based estimator for dynamic CRTE |
| $\widetilde{\mathcal{M}}_\alpha(t)$ | $\int_t^\infty \widetilde{S}_n^\alpha(x)\,dx$ |
| $\mathcal{M}_\alpha(t)$ | $\int_t^\infty S^\alpha(x)\,dx$ |
| $Z_\Psi$ | $\sqrt{nh_n}\,\dfrac{\widetilde{\Psi}_\alpha(X) - \Psi_\alpha(X)}{\widehat{\sigma}_\Psi}$ |
| $Z_{\Psi,t}$ | $\sqrt{nh_n}\,\dfrac{\widetilde{\Psi}_\alpha(X \mid t) - \Psi_\alpha(X \mid t)}{\widehat{\sigma}_{\Psi,t}}$ |

## 2. **Kernel-Based Estimators under Length-Biased Sampling**

Let $X$ be a positive lifetime variable with density $f(t)$ and finite expectation $\mu = \mathbb{E}(X)$. As mentioned above under length-biased sampling, the observed random variable $Y$ has weighted density $g(t)$ which is defined in (3). Suppose $Y_1, \ldots, Y_n$ are independent and identically distributed observations from $g(\cdot)$.

To estimate the underlying density $f(t)$, we need to correct for the sampling bias. Following [6], and based on (3), it can be easily observed that $\widehat{\mu} =$

$\frac{n}{\sum_{i=1}^{n} Y_i^{-1}}$, so we use the weighted kernel estimator

$$\tilde{f}_n(t) = \frac{1}{h_n \sum_{j=1}^{n} Y_j^{-1}} \sum_{i=1}^{n} Y_i^{-1} K\left(\frac{t - Y_i}{h_n}\right),$$

where $K(\cdot)$ is a symmetric kernel function and $\{h_n : n \geq 1\}$ is a bandwidth sequence with $h_n \to 0$ and $nh_n \to \infty$ as $n \to \infty$. The bandwidth $h_n$ controls the tradeoff between smoothness and variance.

To keep the analysis clean, we impose the following regularity conditions.

**Regularity Conditions**

- **(K1):** $K$ is a bounded symmetric kernel with $\int_{\mathbb{R}} K(u)\,du = 1$ and compact support $[-1, 1]$.
- **(K2):** $\int_{\mathbb{R}} u\,K(u)\,du = 0$.
- **(K3):** $\int_{\mathbb{R}} |K(u)|\,du < \infty$.
- **(K4):** $\kappa_2 := \int_{\mathbb{R}} u^2 K(u)\,du \in (0, \infty)$ and $\int |dK(u)| < \infty$.
- **(F1):** $\tau := \inf\{t : F(t) < 1\} < \infty$.
- **(F2):** $\int_0^{\infty} u^{-2} G^{1/r}(u)\,du < \infty$ for some $r > 2$, where $G$ is the distribution of $Y$.(This controls the tail behavior of the length-biased distribution and ensures finite variance in the kernel estimator.)
- **(F3):** The second derivative $f''$ exists and is bounded on $(0, \tau]$, and $\mu_{-1} := \mathbb{E}[X^{-1}] < \infty$.(the expected value of the inverse of X)

When condition (F3) holds, the leading-order bias and variance of $\tilde{f}_n(t)$ are (see [4])

$$\mathbb{E}[\tilde{f}_n(t)] - f(t) \simeq \frac{h_n^2}{2}\,\kappa_2\,f''(t), \tag{4}$$

$$\mathrm{Var}(\tilde{f}_n(t)) \simeq \frac{\mu}{nh_n}\,t^{-1} f(t)\,\|K\|_2^2, \tag{5}$$

where $\|K\|_2^2 := \int K^2(u)\,du$.

Using $\tilde{f}_n$, we construct a survival function estimator

$$\widetilde{S}_n(t) = \int_t^{\infty} \tilde{f}_n(x)\,dx.$$

We then define kernel-based estimators for CRTE and dynamic CRTE:

$$\widetilde{\Psi}_\alpha(X) = \frac{1}{\alpha - 1}\left\{1 - \int_0^{\infty} \widetilde{S}_n^{\alpha}(x)\,dx\right\},$$

$$\widetilde{\Psi}_\alpha(X \mid t) = \frac{1}{\alpha - 1}\left\{1 - \frac{\int_t^{\infty} \widetilde{S}_n^{\alpha}(x)\,dx}{\widetilde{S}_n^{\alpha}(t)}\right\}.$$

In the next section, we study the asymptotic properties of $\widetilde{\Psi}_\alpha(X)$ and $\widetilde{\Psi}_\alpha(X \mid t)$.

## 3. Asymptotic Theory

This section presents our main theoretical results for the proposed estimators. We start with bias-variance approximations for $\widetilde{S}_n(t)$, then derive corresponding results for the entropy functionals, and finally establish consistency and asymptotic normality.

### 3.1. Bias and Variance of the Survival Estimator.

**Theorem 3.1.** *Under conditions (K1)–(K3) and (F3), the asymptotic bias and variance of $\widetilde{S}_n(t)$ satisfy*

$$\mathbb{E}[\widetilde{S}_n(t)] - S(t) \simeq \frac{h_n^2}{2}\,\kappa_2 \int_t^\infty f''(x)\,dx,$$

$$Var(\widetilde{S}_n(t)) \simeq \frac{\mu}{nh_n}\,\|K\|_2^2 \int_t^\infty x^{-1} f(x)\,dx.$$

*Proof.* The results follow by integrating (4) and (5) over $(t,\infty)$ and applying standard integration-by-parts arguments. □

### 3.2. Consistency of the Entropy Estimators.

**Theorem 3.2.** *Assume conditions (K1)–(K3) and (F3) hold. For $\alpha > 1/2$ and $\alpha \neq 1$, we have:*

(i) $\widetilde{\Psi}_\alpha(X) \xrightarrow{P} \Psi_\alpha(X)$ *as* $n \to \infty$;
(ii) $\widetilde{\Psi}_\alpha(X \mid t) \xrightarrow{P} \Psi_\alpha(X \mid t)$ *as* $n \to \infty$.

*Proof. Part (i):* A first-order Taylor expansion gives

$$\int_0^\infty \widetilde{S}_n^\alpha(x)\,dx \simeq \int_0^\infty S^\alpha(x)\,dx + \alpha \int_0^\infty S^{\alpha-1}(x)[\widetilde{S}_n(x) - S(x)]\,dx.$$

Using Theorem 3.1, the asymptotic bias of the integral is

(6)
$$\mathbb{E}\left[\int_0^\infty \widetilde{S}_n^\alpha(x)\,dx\right] - \int_0^\infty S^\alpha(x)\,dx \simeq \frac{\alpha h_n^2 \kappa_2}{2} \int_0^\infty S^{\alpha-1}(x) \left(\int_x^\infty f''(y)\,dy\right) dx,$$

and the variance is

(7) $$\mathrm{Var}\left(\int_0^\infty \widetilde{S}_n^\alpha(x)\,dx\right) \simeq \frac{\alpha^2 \mu \|K\|_2^2}{nh_n} \int_0^\infty S^{2\alpha-2}(x) \left(\int_x^\infty y^{-1} f(y)\,dy\right) dx.$$

Thus, the Mean Squared Eerror (MSE) satisfies

$$\mathrm{MSE}\left(\int_0^\infty \widetilde{S}_n^\alpha(x)\,dx\right) \simeq \left(\frac{\alpha h_n^2 \kappa_2}{2} \int_0^\infty S^{\alpha-1}(x) \left(\int_x^\infty f''(y)\,dy\right) dx\right)^2 + \frac{\alpha^2 \mu \|K\|_2^2}{nh_n} \int_0^\infty S^{2\alpha-2}(x) \left(\int_x^\infty y^{-1} f(y)\,dy\right) dx \to 0,$$

as $n \to \infty$. Hence,

$$\widetilde{\Psi}_\alpha(X) = \frac{1}{\alpha - 1} \left\{ 1 - \int_0^\infty \widetilde{S}_n^\alpha(x)\, dx \right\} \xrightarrow{P} \frac{1}{\alpha - 1} \left\{ 1 - \int_0^\infty S^\alpha(x)\, dx \right\} = \Psi_\alpha(X).$$

*Part (ii):* A similar Taylor expansion gives

(8)
$$\mathbb{E}\left[ \int_t^\infty \widetilde{S}_n^\alpha(x)\, dx \right] - \int_t^\infty S^\alpha(x)\, dx \simeq \frac{\alpha h_n^2 \kappa_2}{2} \int_t^\infty S^{\alpha-1}(x) \left( \int_x^\infty f''(y)\, dy \right) dx,$$

$$\mathbb{E}[\widetilde{S}_n^\alpha(t)] - S^\alpha(t) \simeq \frac{\alpha h_n^2 \kappa_2}{2}\, S^{\alpha-1}(t) \int_t^\infty f''(x)\, dx, \tag{9}$$

with corresponding variance expressions. The MSE of both numerator and denominator vanish as $n \to \infty$, which proves the result. □

3.3. **Bias and Variance of the Entropy Functionals.** For the dynamic version, we need an auxiliary result.

**Proposition 3.3.** *Under conditions (K1)–(K3) and (F3), the estimation error for the dynamic CRTE satisfies*

$$\widetilde{\Psi}_\alpha(X \mid t) - \Psi_\alpha(X \mid t) \simeq \frac{-1}{(\alpha - 1)\, \mathcal{A}_\alpha(t)} \left[ \widetilde{\mathcal{M}}_\alpha(t) - \widetilde{\mathcal{A}}_\alpha(t)\, \frac{\mathcal{M}_\alpha(t)}{\mathcal{A}_\alpha(t)} \right], \tag{10}$$

*where* $\widetilde{\mathcal{M}}_\alpha(t) = \int_t^\infty \widetilde{S}_n^\alpha(x)\, dx$, $\mathcal{M}_\alpha(t) = \int_t^\infty S^\alpha(x)\, dx$, $\widetilde{\mathcal{A}}_\alpha(t) = \widetilde{S}_n^\alpha(t)$, *and* $\mathcal{A}_\alpha(t) = S^\alpha(t)$.

*Proof.* From Theorem 3.2(ii), $\widetilde{\mathcal{A}}_\alpha(t) \xrightarrow{P} \mathcal{A}_\alpha(t)$, so $(\mathcal{A}_\alpha(t)/\widetilde{\mathcal{A}}_\alpha(t) - 1) = O_P(1)$. The result follows by standard ratio expansion; see Proposition 1 in [5]. □

**Theorem 3.4.** *Under the conditions of Theorem 3.1, the asymptotic biases are*

$$\mathbb{E}[\widetilde{\Psi}_\alpha(X)] - \Psi_\alpha(X) \simeq \frac{-\alpha h_n^2 \kappa_2}{2(\alpha - 1)} \int_0^\infty S^{\alpha-1}(x) \left( \int_x^\infty f''(y)\, dy \right) dx, \tag{11}$$

*and*

$$\begin{aligned} \mathbb{E}[\widetilde{\Psi}_\alpha(X \mid t)] - \Psi_\alpha(X \mid t) &\simeq \frac{\alpha h_n^2 \kappa_2}{2(\alpha - 1)\, S^\alpha(t)} \\ &\quad \times \left\{ \frac{\int_t^\infty S^\alpha(x)\, dx}{S(t)} \int_t^\infty f''(x)\, dx \right. \\ &\quad \left. - \int_t^\infty S^{\alpha-1}(x) \left( \int_x^\infty f''(y)\, dy \right) dx \right\}. \end{aligned} \tag{12}$$

*For $\alpha > 1/2$, the asymptotic variances are*

$$Var(\widetilde{\Psi}_\alpha(X)) \simeq \frac{\alpha^2 \mu \|K\|_2^2}{(\alpha - 1)^2 n h_n} \int_0^\infty S^{2\alpha-2}(x) \left( \int_x^\infty y^{-1} f(y)\, dy \right) dx, \tag{13}$$

*and*

$$\text{Var}(\widetilde{\Psi}_\alpha(X \mid t)) \simeq \frac{\alpha^2 \mu \|K\|_2^2}{(\alpha-1)^2 n h_n S^{2\alpha}(t)} \times \left\{ \left( \frac{\int_t^\infty S^\alpha(x)\,dx}{S(t)} \right)^2 \int_t^\infty x^{-1} f(x)\,dx + \int_t^\infty S^{2\alpha-2}(x) \left( \int_x^\infty y^{-1} f(y)\,dy \right) dx \right\}. \tag{14}$$

*Proof.* Expressions (11) and (13) follow from (6) and (7) by the delta method. The results for the dynamic version combine Proposition 3.3 with equations (8)–(9). □

**Remark 1.** These results show that bias is of order $O(h_n^2)$ while variance is of order $O((nh_n)^{-1})$. Thus, for the bandwidth choice satisfying $nh_n^3 \to 0$, variance dominates the MSE in moderate samples, consistent with the simulation findings in Section 4.

### 3.4. **Integrated $L^2$-Consistency.**

**Theorem 3.5.** *Under the conditions of Theorem 3.1, for $\alpha > 1/2$ and $\alpha \neq 1$:*

(i) *$\widetilde{\Psi}_\alpha(X)$ is uniformly consistent in the integrated squared sense for $\Psi_\alpha(X)$;*
(ii) *$\widetilde{\Psi}_\alpha(X \mid t)$ is uniformly consistent in the integrated squared sense for $\Psi_\alpha(X \mid t)$.*

*Proof. Part (i):* The mean integrated squared error (MISE) of $\widetilde{\Psi}_\alpha(X)$ is

$$\text{MISE}(\widetilde{\Psi}_\alpha(X)) = \mathbb{E} \int_0^\infty [\widetilde{\Psi}_\alpha(X) - \Psi_\alpha(X)]^2\,dx = \int_0^\infty \text{MSE}(\widetilde{\Psi}_\alpha(X))\,dx.$$

From (11) and (13),

$$\text{MSE}(\widetilde{\Psi}_\alpha(X)) \simeq \left( \frac{-\alpha h_n^2 \kappa_2}{2(\alpha-1)} \int_0^\infty S^{\alpha-1}(x) \left( \int_x^\infty f''(y)\,dy \right) dx \right)^2 + \frac{\alpha^2 \mu \|K\|_2^2}{(\alpha-1)^2 n h_n} \int_0^\infty S^{2\alpha-2}(x) \left( \int_x^\infty y^{-1} f(y)\,dy \right) dx \to 0,$$

as $n \to \infty$. So $\text{MISE}(\widetilde{\Psi}_\alpha(X)) \to 0$, which gives uniform $L^2$-consistency; see [16].

*Part (ii):* The proof for the dynamic version is similar, using (12) and (14). □

### 3.5. **Asymptotic Normality.**

**Theorem 3.6.** *Suppose $f$ has a bounded derivative on $(0, \tau]$ and conditions (K1)–(K4), (F1)–(F3) are satisfied. Assume $\lim_{n\to\infty}(nh_n^3) = 0$ and $n^{2\lambda} h_n \to \infty$ for some $0 < \lambda < 1/2 - 1/r$ $(r > 2)$. Then for $\alpha > 1/2$ and $\alpha \neq 1$,*

$$\sqrt{nh_n}\, \frac{\widetilde{\Psi}_\alpha(X) - \Psi_\alpha(X)}{\sigma_\Psi} \xrightarrow{d} N(0,1) \quad \text{as } n \to \infty, \tag{15}$$

*where*

$$\sigma_{\Psi}^2 = \frac{\alpha^2 \mu \|K\|_2^2}{(\alpha-1)^2} \int_0^\infty S^{2\alpha-2}(x) \left( \int_x^\infty y^{-1} f(y)\, dy \right) dx. \tag{16}$$

*Proof.* The result follows from the asymptotic normality of $\tilde{f}_n(x)$ established in Corollary 1 of [1], combined with standard functional delta-method arguments similar to Theorem 5 in [5]. □

**Theorem 3.7.** *Under the conditions of Theorem 3.6, for $\alpha > 1/2$ and $\alpha \neq 1$,*

$$\sqrt{nh_n}\, \frac{\widetilde{\Psi}_\alpha(X \mid t) - \Psi_\alpha(X \mid t)}{\sigma_{\Psi,t}} \xrightarrow{d} N(0,1) \quad \text{as } n \to \infty, \tag{17}$$

*where*

$$\begin{aligned} \sigma_{\Psi,t}^2 = {} & \frac{\alpha^2 \mu \|K\|_2^2}{(\alpha-1)^2\, S^{2\alpha}(t)} \\ & \times \left[ \int_t^\infty S^{2\alpha-2}(x) \left( \int_x^\infty y^{-1} f(y)\, dy \right) dx \right. \\ & \left. + \frac{\left( \int_t^\infty S^\alpha(x)\, dx \right)^2}{S^2(t)} \int_t^\infty x^{-1} f(x)\, dx \right]. \end{aligned} \tag{18}$$

*Proof.* The proof combines the asymptotic normality of $\tilde{f}_n(x)$ with the ratio expansion in Proposition 3.3, using similar techniques as Theorem 6 of [5]. □

## 4. Simulation Study

We investigate the finite-sample behavior of $\widetilde{\Psi}_\alpha(X)$ and $\widetilde{\Psi}_\alpha(X \mid t)$ using Monte Carlo experiments. Throughout, we set bandwidth $h = n^{-\beta}$ with $\beta = 0.60$, which ensures $h \to 0$ and $nh \to \infty$. Our simulation study has two main goals: first, to check numerically how well the bias, variance, and MSE formulas from Theorems 3.4–3.5 perform in finite samples; second, to verify the asymptotic normality results stated in Theorems 3.6–3.7.

4.1. **Data Generation under Size-Biased Sampling.** Let $X$ be a nonnegative lifetime variable with density $f(\cdot)$ and survival $S(\cdot)$. Under length-biased sampling, observed values $Y_1, \ldots, Y_n$ are i.i.d. from $g(y) = yf(y)/\mu$. In each Monte Carlo replicate, we generate length-biased samples using two models:

- **Exponential model:** Here $X \sim \text{Exp}(\lambda)$, so $S(t) = e^{-\lambda t}$. The length-biased distribution turns out to be $\text{Gamma}(2, \lambda)$, which we can simulate directly.
- **Weibull model:** We take $X \sim \text{Weibull}(k, \theta)$ with survival function
$$S(t) = \exp\{-(t/\theta)^k\}.$$
To generate length-biased draws, we use
$$Y = \theta Z^{1/k}, \quad \text{where } Z \sim \text{Gamma}(1 + 1/k, 1).$$

We consider sample sizes $n \in \{50, 100, 200, 500\}$, Tsallis orders $\alpha \in \{1.5, 2, 3\}$ (all satisfying $\alpha > 1/2$ and $\alpha \neq 1$), and time points

$$t \in \{0.5, 1.0, 1.5\},$$

when examining the dynamic functional.

4.2. **Computational Details.** Given observations $\{Y_i\}_{i=1}^n$, we compute

$$\tilde{f}_n(t) = \frac{1}{h_n \sum_{j=1}^n Y_j^{-1}} \sum_{i=1}^n Y_i^{-1} K\left(\frac{t - Y_i}{h_n}\right), \quad \widetilde{S}_n(t) = \int_t^\infty \tilde{f}_n(x)\, dx,$$

and then build the entropy estimators

$$\widetilde{\Psi}_\alpha(X) = \frac{1}{\alpha - 1}\left\{1 - \int_0^\infty \widetilde{S}_n^\alpha(x)\, dx\right\},$$
$$\widetilde{\Psi}_\alpha(X \mid t) = \frac{1}{\alpha - 1}\left\{1 - \frac{\int_t^\infty \widetilde{S}_n^\alpha(x)\, dx}{\widetilde{S}_n^\alpha(t)}\right\}.$$

**Kernel:** We use the Gaussian kernel, $K(u) = \phi(u)$, where $\phi$ denotes the standard normal density.

**Bandwidth:** We set $h_n = n^{-2/5}$, which satisfies the CLT conditions in Theorem 3.6.

**Integration:** Integrals are approximated by Riemann sums on $[0, 8 \times \max\{Y_i\}]$ using 2500 grid points.

**Mean estimation:** We estimate the mean by $\widehat{\mu} = n / \sum_{i=1}^n Y_i^{-1}$, which is consistent under length-biased sampling.

4.3. **Performance Metrics.** For each scenario $(n, \alpha, t)$, we run $B = 1000$ Monte Carlo replicates and calculate

$$\text{Bias} = \frac{1}{B}\sum_{b=1}^B (\widehat{\theta}^{(b)} - \theta), \quad \text{Var} = \frac{1}{B-1}\sum_{b=1}^B (\widehat{\theta}^{(b)} - \bar{\widehat{\theta}})^2, \quad \text{MSE} = \text{Bias}^2 + \text{Var},$$

where $\theta$ is either $\Psi_\alpha(X)$ in (1) or $\Psi_\alpha(X \mid t)$ in (2), and $\widehat{\theta}$ is the corresponding estimator.

4.4. **Results: CRTE.** Table 2 shows bias, variance, and MSE results for $\widetilde{\Psi}_\alpha(X)$ across both the exponential and Weibull models.

Inspection of Table 2 reveals a number of systematic features that are consistent across both models. As the sample size increases, variance and mean squared error decrease steadily, in agreement with the theoretical rates derived in Theorems 3.2 and 3.5. In particular, the empirical variances closely follow the $O((nh_n)^{-1})$ rate predicted by (13), while the bias component decays faster at rate $O(h_n^2) = O(n^{-4/5})$, making variance the dominant contributor to MSE for moderate sample sizes.

The effect of the Tsallis order $\alpha$ is also clearly visible. Smaller values of $\alpha$, especially those closer to one, lead to increased bias and variability, which is

TABLE 2. Monte Carlo bias, variance, and MSE of $\widetilde{\Psi}_\alpha(X)$ under length-biased sampling with Gaussian kernel, $h_n = n^{-2/5}$ and $B = 1000$ replicates.

| Model | $\alpha$ | $n$ | Bias | Var | MSE |
|---|---|---|---|---|---|
| **Exponential** | | | | | |
| | 1.5 | 50 | 0.1638 | 0.1487 | 0.1756 |
| | 1.5 | 100 | 0.0997 | 0.0942 | 0.1042 |
| | 1.5 | 200 | 0.0749 | 0.0515 | 0.0571 |
| | 1.5 | 500 | 0.0508 | 0.0260 | 0.0286 |
| | 2 | 50 | 0.0731 | 0.0335 | 0.0388 |
| | 2 | 100 | 0.0484 | 0.0175 | 0.0198 |
| | 2 | 200 | 0.0357 | 0.0109 | 0.0122 |
| | 2 | 500 | 0.0261 | 0.0060 | 0.0067 |
| | 3 | 50 | 0.0398 | 0.0064 | 0.0079 |
| | 3 | 100 | 0.0263 | 0.0036 | 0.0043 |
| | 3 | 200 | 0.0173 | 0.0022 | 0.0025 |
| | 3 | 500 | 0.0141 | 0.0012 | 0.0014 |
| **Weibull** | | | | | |
| | 1.5 | 50 | 0.0090 | 0.0346 | 0.0347 |
| | 1.5 | 100 | 0.0072 | 0.0178 | 0.0178 |
| | 1.5 | 200 | 0.0030 | 0.0093 | 0.0093 |
| | 1.5 | 500 | 0.0066 | 0.0042 | 0.0043 |
| | 2 | 50 | 0.0067 | 0.0088 | 0.0088 |
| | 2 | 100 | 0.0022 | 0.0052 | 0.0052 |
| | 2 | 200 | 0.0006 | 0.0029 | 0.0029 |
| | 2 | 500 | 0.0023 | 0.0010 | 0.0010 |
| | 3 | 50 | 0.0087 | 0.0022 | 0.0023 |
| | 3 | 100 | 0.0051 | 0.0015 | 0.0015 |
| | 3 | 200 | 0.0020 | 0.0007 | 0.0007 |
| | 3 | 500 | 0.0011 | 0.0003 | 0.0003 |

consistent with the multiplicative factor $\alpha/(\alpha-1)$ appearing in the asymptotic expressions. For larger values such as $\alpha = 3$, the estimator exhibits noticeably more stable behavior due to the faster decay of $S^\alpha$ in the tail.

Comparing the two lifetime models, the exponential distribution generally produces smaller MSE values than the Weibull case. This difference is plausibly attributed to the lighter tail of the exponential model, which results in smoother kernel-based survival estimates. From a practical perspective, the results suggest that sample sizes of at least $n \geq 200$ are adequate when $\alpha \geq 2$, whereas values of $\alpha$ closer to one require larger samples for reliable inference.

4.5. **Results: Dynamic CRTE.** Table 3 shows the performance of $\widetilde{\Psi}_\alpha(X \mid t)$ at $t = 0.5$.

TABLE 3. Monte Carlo bias, variance, and MSE of $\widetilde{\Psi}_\alpha(X \mid t)$ at $t = 0.5$ under length-biased sampling.

| Model | $\alpha$ | $n$ | Bias | Var | MSE |
|---|---|---|---|---|---|
| **Exponential** | | | | | |
| | 1.5 | 50 | 0.0510 | 0.0370 | 0.0396 |
| | 1.5 | 100 | 0.0362 | 0.0209 | 0.0222 |
| | 1.5 | 200 | 0.0238 | 0.0106 | 0.0111 |
| | 1.5 | 500 | 0.0282 | 0.0043 | 0.0051 |
| | 2 | 50 | 0.0256 | 0.0065 | 0.0072 |
| | 2 | 100 | 0.0179 | 0.0034 | 0.0038 |
| | 2 | 200 | 0.0152 | 0.0020 | 0.0023 |
| | 2 | 500 | 0.0166 | 0.0008 | 0.0012 |
| | 3 | 50 | 0.0121 | 0.0011 | 0.0013 |
| | 3 | 100 | 0.0087 | 0.0005 | 0.0006 |
| | 3 | 200 | 0.0073 | 0.0003 | 0.0003 |
| | 3 | 500 | 0.0082 | 0.0001 | 0.0002 |
| **Weibull** | | | | | |
| | 1.5 | 50 | 0.0820 | 0.0071 | 0.0138 |
| | 1.5 | 100 | 0.0510 | 0.0040 | 0.0066 |
| | 1.5 | 200 | 0.0230 | 0.0023 | 0.0028 |
| | 1.5 | 500 | 0.0150 | 0.0009 | 0.0012 |
| | 2 | 50 | 0.0600 | 0.0015 | 0.0051 |
| | 2 | 100 | 0.0400 | 0.0008 | 0.0024 |
| | 2 | 200 | 0.0200 | 0.0004 | 0.0009 |
| | 2 | 500 | 0.0120 | 0.0002 | 0.0004 |
| | 3 | 50 | 0.2395 | 0.0002 | 0.0576 |
| | 3 | 100 | 0.2429 | 0.0001 | 0.0592 |
| | 3 | 200 | 0.2445 | 0.0008 | 0.0599 |
| | 3 | 500 | 0.2460 | 0.0004 | 0.0606 |

The results in Table 3 highlight the additional variability introduced by the dynamic structure of $\widehat{\Psi}_\alpha(X \mid t)$. When the evaluation time $t$ is relatively small and the survival probability $S(t)$ remains moderately large, the denominator $\widetilde{S}_n^\alpha(t)$ is well behaved, leading to comparatively small MSE values. As $t$ increases, however, the variability of the denominator becomes more pronounced and starts to dominate the overall estimation error.

This behavior is consistent with the representation in Proposition 3.3, which shows how estimation errors in both the numerator and denominator propagate through the ratio. Despite this added complexity, the empirical MSE still decreases at rate $O((nh_n)^{-1})$, in line with the asymptotic Theorems 3.4 and 3.5.

Based on these observations, it is advisable in practice to restrict attention to time points for which $S(t)$ is not too small. In our experiments, values satisfying $S(t) \geq 0.3$ provided stable and interpretable results, particularly for sample sizes $n \geq 200$.

4.6. **Asymptotic Normality Diagnostics.** Figures 1 and 2 show empirical distributions of the standardized test statistics

$$Z_\Psi = \sqrt{nh_n}\,\frac{\widetilde{\Psi}_\alpha(X) - \Psi_\alpha(X)}{\widehat{\sigma}_\Psi} \quad \Rightarrow \quad N(0,1), \tag{19}$$

and

$$Z_{\Psi,t} = \sqrt{nh_n}\,\frac{\widetilde{\Psi}_\alpha(X \mid t) - \Psi_\alpha(X \mid t)}{\widehat{\sigma}_{\Psi,t}} \quad \Rightarrow \quad N(0,1), \tag{20}$$

where $\widehat{\sigma}^2_\Psi$ and $\widehat{\sigma}^2_{\Psi,t}$ are sample-based estimates of (16) and (18).

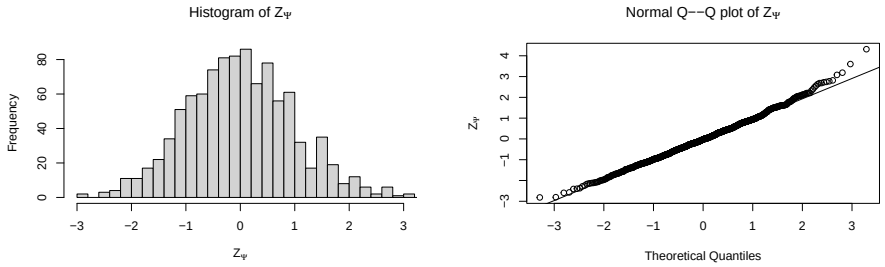


FIGURE 1. Normality diagnostics for the standardized CRTE statistic (Theorem 3.6). Left: Histogram of $Z_\Psi$. Right: Normal Q–Q plot.

Figures 1(Left) and 1(Right) provide diagnostic checks for the asymptotic normality of the standardized test statistic $Z_\Psi$ defined in equation (19). Figure 1(Left) displays a histogram of the $Z_\Psi$ values obtained from the Monte Carlo simulations, demonstrating a distribution closely resembling a standard normal distribution. Figure 1(Right) presents a quantile-quantile (Q-Q) plot, which further confirms the normality assumption. The close alignment of the points along the diagonal line indicates that the simulated data are consistent with a standard normal distribution, validating the theoretical results outlined in Theorem 3.6. This provides empirical support for the use of normal-based confidence intervals and hypothesis tests for the static cumulative residual Tsallis entropy.

Figures 2(Left) and 2(Right) offer normality diagnostics for the standardized test statistic $Z_{\Psi,t}$ associated with the dynamic cumulative residual Tsallis entropy, as defined in equation (20). Similar to the static case, the histogram in Figure 2(Left) exhibits a distribution that closely approximates a standard

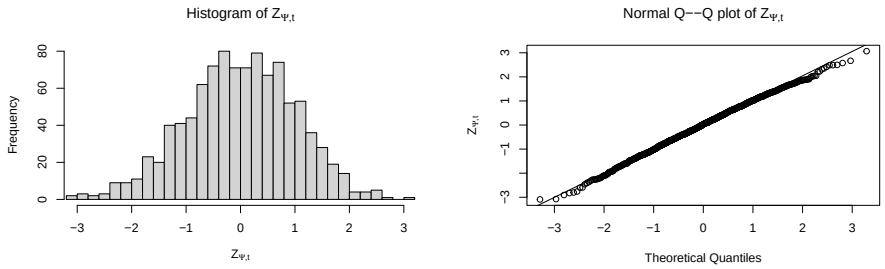


FIGURE 2. Normality diagnostics for the standardized dynamic CRTE statistic at $t = 0.1$ (Theorem 3.7). Left: Histogram of $Z_{\Psi,t}$ in (20). Right: Normal Q–Q plot of $Z_{\Psi,t}$ in (20).

normal distribution. The Q-Q plot in Figure 2(Right) further reinforces this observation, with points aligning well along the diagonal line. These results provide empirical evidence supporting the asymptotic normality of the dynamic cumulative residual Tsallis entropy estimator, as established in Theorem 3.7, and justify the use of normal-based inference procedures for this dynamic measure.

These visual assessments strongly support the theoretical findings presented in Theorems 3.6 and 3.7, demonstrating the validity of the asymptotic normality assumption under the simulated conditions. The histograms and Q–Q plots closely approximate the $N(0,1)$ distribution, even for moderate samples like $n = 200$. This provides strong numerical confirmation of Theorems 3.6–3.7. In practical terms, normal-based confidence intervals and hypothesis tests are expected to perform reliably when $n \geq 100$ and $\alpha \geq 2$, or when $n \geq 200$ for $\alpha$ closer to 1.

The histograms and Q–Q plots line up remarkably well with the $N(0,1)$ distribution, even for moderate samples like $n = 200$. This provides strong numerical confirmation of Theorems 3.6–3.7. In practical terms, normal-based confidence intervals and hypothesis tests should work reliably when $n \geq 100$ and $\alpha \geq 2$, or when $n \geq 200$ for $\alpha$ closer to 1.

## 5. **Empirical Application**

We apply our methodology to automotive brake pad durability data reported in [8]. The lifetimes (measured in thousands of kilometers) come from a sampling scheme that favors longer-lived components. If we treated these observations as a simple random sample, we'd get biased estimates of the survival function and, consequently, distorted entropy estimates. This makes the data a natural fit for the length-bias framework we've developed.

5.1. **Data Description and Model.** The dataset contains $n = 98$ observed lifetimes under length-biased sampling, collected from a renewal process where inspection times naturally favor components with longer service lives. Following standard practice in reliability analysis, we assume Weibull-type behavior and apply the kernel framework from Section 2. To check robustness, we report results for $\alpha = 1.2$ and $\alpha = 1.5$—both are commonly used values in entropy-based reliability analysis.

5.2. **Selection of Evaluation Points.** For the dynamic functional, we choose time points based on empirical quantiles. Figure 3 displays the estimated survival function $\widetilde{S}_n(t)$ at these points. We follow standard practice by restricting attention to values of $t$ where $\widetilde{S}_n(t) \geq 0.3$, ensuring numerical stability and focusing on time points where the survival probability remains reasonably high. The points where the survival curve crosses or exceeds the 0.3 threshold are selected as evaluation points for the dynamic cumulative residual Tsallis entropy analysis. The smoothness of the estimated survival curve is a direct result of the kernel density estimation technique employed, with the bandwidth parameter $h_n$ controlling the trade-off between bias and variance.

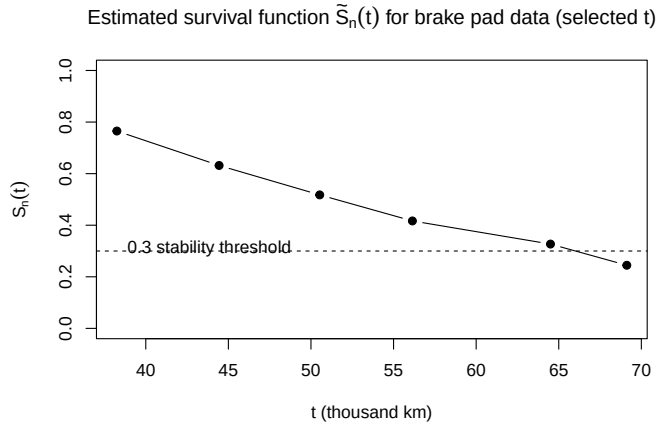


FIGURE 3. Estimated survival function $\widetilde{S}_n(t)$ for brake pad data at selected time points. The horizontal line at 0.3 marks a stability threshold.

5.3. **Dynamic CRTE Estimates.** Table 4 reports $\widetilde{\Psi}_\alpha(X \mid t)$ along with bootstrap standard errors and 95% confidence intervals.

It's worth noting that CRTE and DCRTE are scale-dependent measures that can be negative when $\alpha > 1$; we focus on the comparative trends across $t$ and $\alpha$

TABLE 4. Brake pad data: dynamic CRTE estimates $\widetilde{\Psi}_\alpha(X \mid t)$ with bootstrap standard errors and 95% confidence intervals.

| $\alpha$ | $t$ | $\widetilde{S}_n(t)$ | $\widetilde{\Psi}_\alpha(X \mid t)$ | SE | $\mathrm{CI}_{0.025}$ | $\mathrm{CI}_{0.975}$ |
|---|---|---|---|---|---|---|
| 1.2 | 38.25 | 0.765 | −107.199 | 10.108 | −127.010 | −87.389 |
| 1.2 | 44.44 | 0.632 | −100.634 | 10.142 | −120.512 | −80.755 |
| 1.2 | 50.53 | 0.518 | −94.786 | 10.062 | −114.507 | −75.065 |
| 1.2 | 56.14 | 0.417 | −91.966 | 10.155 | −111.869 | −72.064 |
| 1.2 | 64.50 | 0.327 | −76.330 | 10.385 | −96.685 | −55.976 |
| 1.5 | 38.25 | 0.765 | −35.370 | 3.906 | −43.024 | −27.715 |
| 1.5 | 44.44 | 0.632 | −33.076 | 3.778 | −40.481 | −25.670 |
| 1.5 | 50.53 | 0.518 | −31.109 | 3.666 | −38.295 | −23.923 |
| 1.5 | 56.14 | 0.417 | −30.348 | 3.793 | −37.781 | −22.914 |
| 1.5 | 64.50 | 0.327 | −24.430 | 3.674 | −31.631 | −17.228 |

rather than absolute magnitudes. Bootstrap resampling provides solid uncertainty quantification, complementing the asymptotic theory in Theorem 3.7.

The estimates show a clear systematic trend as $t$ increases. While the absolute values depend on scaling, the qualitative pattern holds for both $\alpha$ values. Notice that confidence intervals are wider when $\alpha = 1.2$, reflecting higher sensitivity to tail behavior when $\alpha$ is closer to one.

## 6. **Summary and Future Directions**

This paper developed nonparametric kernel-based estimators for Tsallis-type cumulative residual entropy functionals under length-biased lifetime sampling. By explicitly correcting for the biased sampling mechanism, we derived large-sample expressions for bias and variance, established consistency in both probability and integrated $L^2$ senses, and proved central limit theorems under standard regularity conditions.

Monte Carlo experiments based on exponential and Weibull models confirmed the theoretical findings and demonstrated that the proposed estimators perform reliably for moderate to large sample sizes. The empirical application to automotive component lifetime data further illustrated the practical relevance of the methodology and highlighted the distortions that arise when length bias is ignored in entropy-based inference.

Several directions merit further investigation. One natural extension is the development of fully data-driven bandwidth selection procedures tailored to entropy functionals under biased sampling. Another promising avenue is the adaptation of the proposed framework to alternative entropy measures beyond the Tsallis family. Finally, extending the methodology to censored data or to more complex biased sampling schemes, such as doubly truncated observations, would substantially broaden its applicability in reliability and survival analysis.

**Remark 2.** The empirical analysis shows that our approach produces stable inferences for length-biased lifetime data. The dynamic analysis underscores the importance of staying away from extreme tail regions when choosing evaluation points. Specifically, practitioners should avoid evaluation points where the survival function $S(t)$ is very small (e.g., below 0.2–0.3), as the variability of the denominator in the estimator becomes more pronounced as $t$ increases. Also, comparing results for $\alpha = 1.2$ and $\alpha = 1.5$ suggests the methodology is qualitatively robust to moderate changes in the parameter. These findings complement our simulation evidence and provide empirical backing for the asymptotic framework we've developed.

## 7. Author Contributions

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

## 8. Data Availability Statement

The data used in this study are publicly available and were obtained from Lawless [8]. The dataset consists of automotive brake pad lifetimes. No new empirical data were collected. Simulated data and computational procedures used to support the findings are available from the corresponding author upon reasonable request.

## 9. Aknowledgement

The authors would like to thank the reviewers for their valuable comments and suggestions, which helped improve the quality of this manuscript.

## 10. Ethical considerations

This study does not involve human participants or animals. The data used are publicly available, and no ethical approval was required. The authors confirm that no data fabrication or falsification was performed.

## 11. Funding

There is no funding available.

## 12. Conflict of interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References


[1] Ajami, M., Fakoor, V., & Jomhoori, S. (2013). Some asymptotic results of kernel density estimator in length-biased sampling. Journal of Sciences, Islamic Republic of Iran, 24(1), 55–62.

[2] Ajami, M., Zamini, R., & Amir Jahanshahi, S. M. (2024). The extended Glivenko–Cantelli property for kernel-smoothed estimator of the cumulative distribution function in length-biased sampling. Journal of Mahani Mathematical Research Center, 13(2), 535–545. https://doi.org/10.22103/jmmr.2024.22620.1546

[3] Asadi, M., & Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. Journal of Statistical Planning and Inference, 137(6), 1931–1941. https://doi.org/10.1016/j.jspi.2006.06.035

[4] Guillamón, A., Navarro, J., & Ruiz, J. M. (1998). Kernel density estimation using weighted data. Communications in Statistics – Theory and Methods, 27(9), 2123–2135. https://doi.org/10.1080/03610929808832126

[5] Irshad, M., Maya, R., Buono, F., & Longobardi, M. (2021). Kernel estimation of cumulative residual Tsallis entropy and its dynamic version under $\rho$-mixing dependent data. Entropy, 24(9), 1189. https://doi.org/10.3390/e24091189

[6] Jones, M. C. (1991). The performance of kernel density functions in kernel distribution function estimation. Statistics & Probability Letters, 9(2), 129–132. https://doi.org/10.1016/0167-7152(91)90110-V

[7] Khammar, A. H., & Jahanshahi, S. M. A. (2018). On weighted cumulative residual Tsallis entropy and its dynamic version. Physica A: Statistical Mechanics and its Applications, 491, 678–692. https://doi.org/10.1016/j.physa.2017.10.024

[8] Lawless, J. F. (2011). Statistical models and methods for lifetime data (2nd ed.). Wiley, New York.

[9] Misagh, F., & Yari, Gh. (2016). On shift-dependent cumulative entropy measures. International Journal of Mathematics and Mathematical Sciences, 2016, 9623465. https://doi.org/10.1155/2016/9623465

[10] Mohamed, S. (2022). On cumulative residual Tsallis entropy and its dynamic version of concomitants of generalized order statistics. Communications in Statistics – Theory and Methods, 51(8), 2534–2551. https://doi.org/10.1080/03610926.2020.1864812

[11] Rao, M., Chen, Y., Vemuri, B. C., & Wang, F. (2004). Cumulative residual entropy: A new measure of information. IEEE Transactions on Information Theory, 50(6), 1220–1228. https://doi.org/10.1109/TIT.2004.828057

[12] Rao, M. (2005). More on a new concept of entropy and information. Journal of Theoretical Probability, 18(4), 967–981. https://doi.org/10.1007/s10959-005-3517-6

[13] Sati, M. M., & Gupta, N. (2015). Some characterization results on dynamic cumulative residual Tsallis entropy. Journal of Probability and Statistics, 2015, 576203. https://doi.org/10.1155/2015/576203

[14] Sunoj, S. M., Krishnan, A. S., & Sankaran, P. G. (2018). A quantile-based study of cumulative residual Tsallis entropy measures. Physica A: Statistical Mechanics and its Applications, 494, 410–421. https://doi.org/10.1016/j.physa.2017.11.024

[15] Toomaj, A., & Agh Atabay, H. (2022). Some new findings on the cumulative residual Tsallis entropy. Journal of Computational and Applied Mathematics, 400, 113669. https://doi.org/10.1016/j.cam.2021.113669

[16] Wegman, E. J. (1972). Nonparametric probability density estimation: I. A summary of available methods. Technometrics, 14(3), 533–546. https://doi.org/10.1080/00401706.1972.10488907

[17] Wicksell, S. D. (1925). The corpuscle problem: A mathematical study of a biometric problem. Biometrika, 17(1–2), 84–99. https://doi.org/10.1093/biomet/17.1-2.84

[18] Zamini, R., Ajami, M., & Ghafouri, S. (2023). Kernel estimators for mean residual lifetime in length-biased sampling. Communications in Statistics – Theory and Methods, 8(1), 1–15. https://doi.org/10.1080/03610926.2023.2277129
[19] Zamini, R., Ajami, M., & Parvizi, S. (2024). Kernel estimation of Tsallis entropy and its generalization for length-biased data. Journal of the Iranian Statistical Society, 23(1), 131–152. https://doi.org/10.22034/jirss.2024.2016847.1046

Masoud Ajami
Orcid number: 0000-0002-6045-128x
Department of Statistics
Faculty of Mathematical Sciences
Vali-e-Asr University of Rafsanjan
Rafsanjan, Iran.
*Email address*: m.ajami@vru.ac.ir

Raheleh Zamini
Orcid number: 0009-0001-4502-7763
Department of Mathematics
Faculty of Mathematical Sciences and Computer
Kharazmi University
Tehran, Iran.
*Email address*: zamini@khu.ac.ir