JIRSS (2024)
Vol. 23, No. 01, pp 131-152
DOI: 10.22034/jirss.2024.2016847.1046

Kernel Estimation of Tsallis Entropy and its Generalization for Length-biased Data

Raheleh Zamini [1] **and Masoud Ajami** [2] **and Sepide Parvizi** [1]

[1] Department of Mathematics, Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran, Iran.
[2] Department of Statistics, Faculty of Mathematical Sciences, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.

Received: 30/11/2023, Accepted: 11/12/2024, Published online: 29/12/2024

Abstract. A new generalized Shannon entropy is Tsallis entropy. The Shannon entropy is additive and the Tsallis entropy is, however, non-additive. Due to the flexibility of the Tsallis entropy compared to the Shannon entropy, the non-additive entropy measures find their justification in many areas. In this paper, we propose two non-parametric kernel estimators for the Tsallis entropy and two non-parametric kernel estimators for the residual Tsallis entropy for the length-biased data. We investigate some asymptotic properties for these estimators such as the consistency and asymptotic normality. We obtain the bias, variance and the mean integrated squared error (MISE) of estimators. We also compare the behaviour of proposed estimators using the Monte Carlo simulation and plot some figures to see how close the fitted distribution is to the histogram of the data. In the end, we use a real dataset to show the performance of the proposed estimators.

Keywords. Asymptotic normality, Bandwidth, Kernel, Length-biased data, Residual Tsallis entropy, Strong consistency, Tsallis entropy.

MSC: 94A17, 62B10, 62G20.

Corresponding Author: Raheleh Zamini (rahelehzamini1361@yahoo.com)
Masoud Ajami (m.ajami@vru.ac.ir)
Sepide Parvizi (sepide13p@yahoo.com).

1 Introduction

The notion of entropy first appeared in physics and mathematics. The Shannon entropy of a continuous random variable X with density function f is

$$H(f) = -\int_{0}^{+\infty} f(x) \log f(x) \mathrm{d}x, \tag{1.1}$$

whenever this integral is meaningful. This concept was initially introduced by Shannon (1948). This measure is more applicable in goodness-of-fit test, quantization theory, statistical communication theory and econometrics.

This paper focuses on a particular generalization of the entropy, often called the Tsallis entropy (SeeTsallis (1988)). Tsallis entropy, denoted by $\tau_\alpha(X)$, given by (2.4), is a powerful tool that plays a fundamental role in information theory, physics, chemistry, and technology. For more applications of this concept one can see Cartwright (2014). Generally, $\tau_\alpha(X)$ can be negative, however, it can be non-negative if a proper value of α is chosen. Unlike Shannon entropy, which is additive, $\tau_\alpha(X)$ is non-additive. Due to the flexibility of the Tsallis entropy compared to the Shannon entropy, the non-additive entropy measures find their justification in many areas of thermodynamics, statistical mechanics, and quantum mechanics, where systems are inherently non-additive. The Tsallis entropy is closely related to the Rényi entropy. When $\alpha > 0$, the Tsallis entropy is a monotonic function of Rényi entropy which is additive. (For more details about Rényi entropy one can see Rényi (1961) and Section 9.3. of Naudts (2011)).

Suppose that X is the life length of a system. The Tsallis entropy is appropriate to measure the uncertainty of such a system. Now assume that the system remains alive until the age t and we are interested in measuring the uncertainty of its remaining lifetime, denoted by $X_t = X - t | X > t$. In this case $\tau_\alpha(X)$ is not appropriate to measure the remaining or residual uncertainty in the system's lifetime. To address this limitation, the residual Tsallis entropy, denoted by $\tau_\alpha(X;t)$, given by (2.5), is useful (See Nanda and Paul (2006)). Nanda and Paul (2006), Kumar and Taneja (2011), Rajesh and Sunoj (2019), Baratpour and Khammar (2016), Misagh and Yari (2012) and Chakraborty and Pradhan (2023) investigated some properties and applications of the residual Tsallis.

In this paper, for the first time, some estimators for $\tau_\alpha(X)$ and $\tau_\alpha(X;t)$ under the length-biased sampling model have been proposed and some asymptotic properties of these estimators, such as consistency and asymptotic normality, have been investigated. The rest of the paper is structured as follows. In Section 2, the concepts of the Tsallis entropy and length-biased sampling are briefly presented. In Section 3, estimators for $\tau_\alpha(X)$ and $\tau_\alpha(X;t)$ based on the length-biased data are proposed, respectively. Section 4 contains main results for proposed estimators. In this section, we investigate some asymptotic results for these estimators. Simulation studies of the proposed estimators is given in Section 5. Also application to the real data is reported in Section 6. The conclusion and detailed explanation of innovation are given in Section 7. At the end, the proofs of the lemmas and main theorems are given in the Appendix.

2 The Concept of Length-biased and Tsallis Entropy

Length-biased data appears in many fields of research. Such sampling schemes are usually employed in observational studies either due to their convenience or their cost-effectiveness. If the probability of an item selected in the sample is proportional to its length, then the distribution of the observed length is known as the length-biased distribution. Examples of such data can be found in Asgharian et al. (2002), Addona and Wolfson (2006) and Shen et al. (2009). Let X be a random variable with density function f and distribution function F. Associated with the random variable X, the length-biased random variable Y with density function g is given by:

$$g(x) = \frac{xf(x)}{\mu},\ \mu > 0, \tag{2.1}$$

where $\mu = E(X) < \infty$. A direct algebra shows that the distribution function of Y can be written as:

$$G(t) = \frac{1}{\mu}\int_0^t xf(x)\mathrm{d}x, \tag{2.2}$$

and F is determined uniquely by G, namely

$$F(t) = \mu\int_0^t y^{-1}\mathrm{d}G(y),\ t \geq 0. \tag{2.3}$$

A special case of the length-biased distribution introduced by Rajagopalan et al. (2019). They substitute the probability density function of Aradhana distribution $f_\theta(x) = \frac{\theta^3}{\theta^2+2\theta+2}(1+x)^2e^{-\theta x}$ and its mean $\mu = \frac{\theta^2+6+4\theta}{\theta(\theta^2+2\theta+2)}$ in Equation (2.2) and proposed the length-biased Aradhana (LBA) distribution. They investigated the various statistical properties of this distribution function.

The Tsallis entropy of order α is defined as

$$\begin{aligned}\tau_\alpha(X) =& \frac{1}{\alpha-1}\Big(1 - E\Big[\big(f(X)\big)^{\alpha-1}\Big]\Big)\\ =& \frac{1}{\alpha-1}\Big(1 - \int_0^{+\infty}\big(f(x)\big)^{\alpha}\mathrm{d}x\Big),\ \alpha \neq 1, \alpha > 0,\end{aligned} \tag{2.4}$$

which tends to the Shannon entropy, as $\alpha \to 1$. The residual Tsallis entropy is a generalization of the function Tsallis entropy defined by (2.4) for a unit surviving up to age t and is defined as

$$\tau_\alpha(X;t) = \frac{1}{\alpha-1}\left(1 - \frac{\int_t^{+\infty}\big(f(x)\big)^{\alpha}\mathrm{d}x}{(\bar{F}(t))^{\alpha}}\right),\ \alpha \neq 1, \alpha > 0, \tag{2.5}$$

where $\bar{F}(t) = 1 - F(t)$ is the survival function of X. Note that when $t = 0$, (2.5) reduces to (2.4).

The entropy of past lifetime $X_{[t]} = [X|X \leq t]$ with the density function $f_{[t]}(x) = \frac{f(x)}{F(t)}$ $0 < x < t$, is given by $\tau^{\alpha}_{[t]}(X) = \frac{1}{\alpha-1}(1 - \int_0^t \frac{f^{\alpha}(x)}{(F(t))^{\alpha}})$, which tends to the Shannon entropy of $X_{[t]}$ as $\alpha \to 1$. The Shannon entropy of $X_{[t]}$ is called the past entropy measure and has the following formula

$$\bar{H}(t) = -\int_0^t \log(f_{[t]}(x))f_{[t]}(x)dx = 1 - \frac{1}{F(t)}\int_0^t f(x)\log(\frac{f(x)}{F(t)})dx.$$

Under the length-biased sampling model, Iranmanesh et al. (2022) introduced some non-parametric estimators for $\bar{H}(t)$ and investigated some of their asymptotic properties.

In this paper we propose two estimators for $\tau_{\alpha}(X)$ given by (2.4) and $\tau_{\alpha}(X;t)$ given by (2.5), respectively, when data are under the length-biased data.

3 Proposed Estimators

Assume that $Y_1, Y_2, \ldots, Y_n$ are *i.i.d.* random variables with common density function $g(x)$ defined in (2.1). Cox (1969) defined the following estimator for F in (2.3)

$$F_n(t) = \mu_n \int_0^t y^{-1}\mathrm{d}G_n(y), \tag{3.1}$$

where

$$\mu_n^{-1} = \int_0^{\infty} y^{-1}\mathrm{d}G_n(y). \tag{3.2}$$

Based on (3.1), Jones (1991) proposed the following kernel estimator

$$f_n(t) = \frac{1}{h_n}\int_R K\Big(\frac{t-y}{h_n}\Big)dF_n(y) = \frac{\mu_n}{nh_n}\sum_{i=1}^{n}\frac{1}{Y_i}K\Big(\frac{t-Y_i}{h_n}\Big), \tag{3.3}$$

where $K(x)$ is a symmetric kernel and h_n is a positive bandwidth satisfying $h_n \to 0$ and $nh_n \to \infty$ when $n \to \infty$. Ajami et al. (2013) investigated the consistency and asymptotic normality of $f_n(t)$ under some conditions. We use these asymptotic properties in the proof of main results.

Since f is not known, a natural estimator for $\tau_{\alpha}(X)$ can be obtained by substituting f by f_n, given by

$$\begin{aligned}\widehat{\tau}_{\alpha_1}(X) &= \frac{1}{(\alpha-1)}\left(1 - \int_0^{\infty}(f_n(x))^{\alpha-1}dF_n(x)\right)\\ &= \frac{1}{(\alpha-1)}\left(1 - \frac{\mu_n}{n}\sum_{i=1}^{n}Y_i^{-1}(\tilde{f}(Y_i))^{\alpha-1}\right),\end{aligned} \tag{3.4}$$

where $\tilde{f}(Y_i)$ is the density estimator obtained from the sample without Y_i, given by

$$\tilde{f}(Y_i) = \frac{\mu_{n-1}}{(n-1)h_{n-1}} \sum_{j=1, j\neq i}^{n} \frac{1}{Y_j} K\Big(\frac{Y_i - Y_j}{h_{n-1}}\Big). \tag{3.5}$$

Another estimator of $\tau_\alpha(X)$ is directly obtained by plug-in f_n in the $\tau_\alpha(X)$, and is given by

$$\hat{\tau}_{\alpha_2}(X) = \frac{1}{\alpha-1}\Big(1 - \int_0^{+\infty} \Big(f_n(x)\Big)^{\alpha} \mathrm{d}x\Big), \; \alpha \neq 1, \alpha > 0, \tag{3.6}$$

where the kernel estimator $f_n(x)$ is defined in (3.3). In a similar way we propose the following kernel estimators for $\tau_\alpha(X;t)$ as below:

$$\begin{aligned}\widehat{\tau}_{\alpha_1}(X;t) &= \frac{1}{(\alpha-1)}\left(1 - \frac{\int_t^{\infty} (f_n(x))^{\alpha-1} dF_n(x)}{(\bar{F}_n(t))^{\alpha}}\right)\\ &= \frac{1}{(\alpha-1)}\left(1 - \frac{\mu_n}{n(\bar{F}_n(t))^{\alpha}} \sum_{i=1}^{n} Y_i^{-1} I[Y_i \geq t] (\tilde{f}(Y_i))^{\alpha-1}\right),\end{aligned} \tag{3.7}$$

and

$$\widehat{\tau}_{\alpha_2}(X;t) = \frac{1}{(\alpha-1)}\left(1 - \frac{\int_t^{\infty} (f_n(x))^{\alpha} dx}{(\bar{F}_n(t))^{\alpha}}\right). \tag{3.8}$$

4 Main Results

In this section, some asymptotic properties of the proposed estimators are investigated under a set of conditions. For the sake of simplicity, the assumptions used in this paper are listed below.

Assumptions.

A1. The kernel function $K(\cdot)$ is symmetric, of bounded variation on $(-1,1)$. In addition $K(x) = 0$ if $t \notin (-1,1)$.

A2. $\int_{-1}^{1} K(x)dx = 1$.

A3. $\int_{-1}^{1} xK(x)dx = 0$.

A4. $\int_{-1}^{1} x^2 K(x)dx = m < \infty$.

A5. $\int_{-1}^{+1} |\, \mathrm{d}K(x)| = V_K < \infty$

A6. $\tau := \inf\{x; F(x) < 1\} < \infty$.

A7. $\int_0^{+\infty} u^{-2} G^{\frac{1}{r}}(u)\mathrm{d}u < \infty$, for some $r > 2$.

A8. $\mu_{-1} = E\left[X^{-1}\right] < \infty$.

A9. $\lim_{n\to\infty} h_n = 0$ and $\lim_{n\to\infty} nh_n = \infty$.

A10. $f(x) = O(x)$ as $x \to 0$.

4.1 Asymptotic Results for the Kernel Estimators of $\tau_\alpha(X)$

In this subsection, we present some asymptotic properties for $\widehat{\tau}_{\alpha_1}(X)$ and $\widehat{\tau}_{\alpha_2}(X)$ in the from of the following theorems.

Theorem 4.1. *Let Assumptions* **A1** - **A10** *hold and* $\lim_{n\to+\infty} \frac{\log n}{n^{\frac{1}{2}+\lambda} h_n} = 0$, *for any* $0 < \lambda < \frac{1}{2} - \frac{1}{r}$, *and some* $r > 2$. *Then for any* $\alpha > 1$, $\alpha \in N$, $\hat{\tau}_{\alpha_1}(X)$ *is a consistent estimator for* $\tau_\alpha(X)$, *i.e., as* $n \longrightarrow \infty$

$$\lim_{n\to\infty} \sup_{0\le x<\infty} \left|\hat{\tau}_{\alpha_1}(X) - \tau_\alpha(X)\right| = 0 \quad a.s.$$

Theorem 4.2. *Suppose that* $\hat{\tau}_{\alpha 2}(X)$ *is the non-parametric estimator of Tsallis entropy defined in* (3.6). *If* $f''(t)$ *exists, then under the Assumptions* **A8** *and* **A9**, $\hat{\tau}_{\alpha 2}(X)$ *is a consistent estimator of* $\tau_\alpha(X)$.

Theorem 4.3. *Suppose that* $\hat{\tau}_{\alpha 2}(X)$ *is the non-parametric estimator of Tsallis entropy defined in* (3.6). *Then, under the conditions of Theorem 4.2, the bias and variance of* $\hat{\tau}_{\alpha 2}(X)$ *are given for* $\alpha > \frac{1}{2}$ *as*

$$\mathrm{Bias}(\hat{\tau}_{\alpha 2}(X)) \simeq \frac{-\alpha}{2(\alpha-1)} h_n^2 \mu_2(K) \int_0^{+\infty} f^{\alpha-1}(x) f''(x)\mathrm{d}x, \tag{4.1}$$

$$\mathrm{Var}(\hat{\tau}_{\alpha 2}(X)) \simeq \frac{\alpha^2}{(\alpha-1)^2 n h_n} \mu \|K\|_2^2 \int_0^{+\infty} f^{2\alpha-1}(x) x^{-1}\mathrm{d}x. \tag{4.2}$$

Theorem 4.4. *Suppose that* $\hat{\tau}_{\alpha 2}(X)$ *is the non-parametric estimator of* $\tau_\alpha(X)$ *defined in*(3.6), *Then, under the conditions of Theorem 4.2, for* $\alpha > \frac{1}{2}$ *and* $\alpha \neq 1$

$$\lim_{n\to\infty} MISE(\hat{\tau}_{\alpha 2}(X)) = 0.$$

Theorem 4.5. *Suppose that* f' *and* f'' *exist and* $f'(t)$ *is bounded at* $t \in (0, \tau]$. *Let* $nh_n^3 \to 0$ *and* $n^{2\lambda} h_n \to \infty$ *as* $n \to \infty$, *for any* $0 < \lambda < \frac{1}{2} - \frac{1}{r}$, *for some* $r > 2$. *Then as* $n \to \infty$, *under the Assumptions* **A1** - **A9**, *for* $\hat{\tau}_{\alpha 2}(X)$ *defined in* (3.6) *with* $\alpha > \frac{1}{2}$,

$$(nh_n)^{\frac{1}{2}} \left\{ \frac{\hat{\tau}_{\alpha 2}(X) - \tau_\alpha(X)}{\sigma_\tau} \right\}, \tag{4.3}$$

has a standard normal distribution where

$$\sigma_\tau^2 \simeq \frac{\alpha^2}{(\alpha-1)^2} \mu \|K\|_2^2 \int_0^{+\infty} f^{2\alpha-1}(x) x^{-1}\mathrm{d}x. \tag{4.4}$$

4.2 Asymptotic Results for the Kernel Estimators of $\tau_\alpha(X;t)$

In this subsection, some asymptotic properties for $\widehat{\tau}_{\alpha_1}(X;t)$ and $\widehat{\tau}_{\alpha_2}(X,t)$ are presented in the following Theorems.

Theorem 4.6. *Under conditions of Theorem 4.1 for every, it can be shown that*

$$\lim_{n\to\infty} \sup_{0\le t<\infty} \left|\widehat{\tau}_{\alpha_1}(X;t) - \tau_\alpha(X;t)\right| = 0 \quad a.s.$$

Theorem 4.7. *Under the conditions of Theorem 4.2, $\widehat{\tau}_{\alpha_2}(X;t)$ for $\alpha \neq 1, \alpha > \frac{1}{2}$ is a consistent estimator of $\tau_\alpha(X;t)$.*

Theorem 4.8. *Under the conditions of Theorem 4.2, for every $\alpha > \frac{1}{2}, \alpha \neq 1$, it can be shown that* $\mathrm{MISE}(\widehat{\tau}_{\alpha_2}(X;t)) \longrightarrow 0, \quad as \quad n \to \infty.$

5 Simulation

In this section, in order to illustrate the behaviour of the proposed estimators, we present the result of a preliminary small-sample simulation study, especially those of Monte Carlo method. For this purpose, three family of distributions are considered to generate the length-biased data. These families of distributions are beta(2,1), gamma(2,1) and log-normal(1,1). It is easy to see that when X has a beta(α,β) or gamma(α,β) distribution, the corresponding length-biased distribution has a beta$(\alpha+1,\beta)$ or gamma$(\alpha+1,\beta)$ distribution respectively. Also when X has a log-normal(μ,σ) distribution, the corresponding length-biased distribution has a density

$$g_Y(y) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} e^{-\left(\mu+\frac{\sigma^2}{2}\right)} \exp\left\{-\frac{(\log y-\mu)^2}{2\sigma^2}\right\}.$$

500 samples of size n=10, 20, 30, 50, 100, 200, 300 and 400 are taken into consideration. There are a number of options for choosing the bandwidth h_n. We apply the rule-of-thumb bandwidth selection method mentioned by Borrajo et al. (2017). Also, the Epanechnikov kernel $K(x) = \frac{3}{4}(1-x^2)I(x \le 1)$ is considered. Tables 1-4 contain the root of mean squared error (RMSE) values and the absolute value of the biases of the proposed estimators for $\alpha = 3$ and $\alpha = 4$ at different sample sizes for each of the selected distributions. According to the results in Tables 1 and 2, $\hat{\tau}_{\alpha_1}(X)$ defined in (3.4) performs consistently better than $\hat{\tau}_{\alpha_2}(X)$ defined in (3.6) respect to the bias and RMSE. Furthermore, one can see from Tables 3 and 4 that $\hat{\tau}_{\alpha_1}(X;t)$ in (3.7) has a better performance as compared with $\hat{\tau}_{\alpha_2}(X;t)$ in (3.8).

Table 1: Performance of $\hat{\tau}_{\alpha_1}(X)$ against $\hat{\tau}_{\alpha_2}(X)$ with $\alpha = 4$.

n	$bias\,(\hat{\tau}_{\alpha_1}(X))$	$RMSE\,(\hat{\tau}_{\alpha_1}(X))$	$bias\,(\hat{\tau}_{\alpha_2}(X))$	$RMSE\,(\hat{\tau}_{\alpha_2}(X))$
beta(2,1)				
10	0.903	0.905	0.959	0.959
20	0.798	0.803	0.871	0.872
30	0.725	0.734	0.805	0.808
50	0.651	0.626	0.728	0.716
100	0.564	0.571	0.635	0.625
200	0. 462	0.482	0.521	0.531
300	0.391	0.399	0.493	0.414
400	0.365	0.374	0.418	0.401
gamma(2,1)				
10	0.004	0.013	0.009	0.014
20	0.003	0.009	0.007	0.012
30	0.002	0.008	0.006	0.010
50	0.002	0.007	0.006	0.009
100	0.002	0.005	0.005	0.006
200	0.001	0.004	0.004	0.005
300	0.001	0.003	0.003	0.004
400	0.001	0.003	0.003	0.004
lognormal(1,1)				
10	0.002	0.007	0.008	0.012
20	0.002	0.007	0.007	0.011
30	0.002	0.006	0.007	0.010
50	0.001	0.005	0.006	0.009
100	0.001	0.004	0.004	0.006
200	0.001	0.002	0.003	0.004
300	0.001	0.001	0.001	0.001
400	0.001	0.001	0.001	0.001

Table 2: Performance of $\hat{\tau}_{\alpha_1}(X)$ against $\hat{\tau}_{\alpha_2}(X)$ with $\alpha = 3$.

n	$bias\,(\hat{\tau}_{\alpha_1}(X))$	$RMSE\,(\hat{\tau}_{\alpha_1}(X))$	$bias\,(\hat{\tau}_{\alpha_2}(X))$	$RMSE\,(\hat{\tau}_{\alpha_2}(X))$
beta(2,1)				
10	0.697	0.700	0.801	0.801
20	0.581	0.589	0.695	0.696
30	0.512	0.524	0.627	0.629
50	0.501	0.491	0.618	0.581
100	0.464	0.334	0.587	0.471
200	0.391	0.315	0.423	0.372
300	0.278	0.271	0.315	0.331
400	0.264	0.259	0.291	0.291
gamma(2,1)				
10	0.004	0.032	0.020	0.031
20	0.002	0.023	0.017	0.026
30	0.004	0.021	0.016	0.025
50	0.004	0.020	0.014	0.024
100	0.003	0.017	0.013	0.018
200	0.002	0.014	0.009	0.016
300	0.002	0.010	0.008	0.009
400	0.002	0.008	0.006	0.007
lognormal(1,1)				
10	0.001	0.020	0.023	0.032
20	0.003	0.019	0.020	0.028
30	0.003	0.014	0.018	0.024
50	0.003	0.014	0.017	0.022
100	0.003	0.012	0.014	0.018
200	0.002	0.009	0.012	0.010
300	0.002	0.006	0.008	0.009
400	0.001	0.005	0.006	0.006

Table 3: Performance of $\hat{\tau}_{\alpha_1}(X;t)$ against $\hat{\tau}_{\alpha_2}(X;t)$ with $\alpha = 4$.

n	$bias\left(\hat{\tau}_{\alpha_1}(X;t)\right)$	$RMSE\left(\hat{\tau}_{\alpha_1}(X;t)\right)$	$bias\left(\hat{\tau}_{\alpha_2}(X;t)\right)$	$RMSE\left(\hat{\tau}_{\alpha_2}(X;t)\right)$
beta(2,1)(t=.5)				
10	2.821	2.836	2.941	2.959
20	2.547	2.562	2.716	2.740
30	2.370	2.382	2.565	2.582
50	2.191	2.141	2.381	2.354
100	1.840	1.791	1.952	1.931
200	1.421	1.461	1.602	1.696
300	1.249	1.270	1.481	1.492
400	1.151	1.155	1.302	1.341
gamma(2,1) (t=2)				
10	0.019	0.097	0.115	0.507
20	0.006	0.045	0.031	0.092
30	0.002	0.025	0.018	0.043
50	0.003	0.022	0.018	0.022
100	0.003	0.018	0.013	0.015
200	0.002	0.012	0.008	0.014
300	0.002	0.005	0.004	0.010
400	0.001	0.003	0.004	0.008
lognormal(1,1) (t=2)				
10	0.006	0.008	0.016	0.049
20	0.006	0.007	0.007	0.016
30	0.005	0.004	0.005	0.009
50	0.005	0.004	0.004	0.007
100	0.004	0.003	0.003	0.005
200	0.004	0.003	0.002	0.003
300	0.003	0.002	0.001	0.002
400	0.001	0.001	0.001	0.001

Table 4: Performance of $\hat{\tau}_{\alpha_1}(X;t)$ against $\hat{\tau}_{\alpha_2}(X;t)$ with $\alpha = 3$.

n	$bias\,(\hat{\tau}_{\alpha_1}(X;t))$	$RMSE\,(\hat{\tau}_{\alpha_1}(X;t))$	$bias\,(\hat{\tau}_{\alpha_2}(X;t))$	$RMSE\,(\hat{\tau}_{\alpha_2}(X;t))$
beta(2,1)(t=.5)				
10	1.618	1.648	1.795	1.849
20	1.406	1.418	1.635	1.654
30	1.273	1.282	1.508	1.522
50	1.189	1.170	1.387	1.463
100	0.837	0.843	1.081	1.078
200	0.639	0.685	0.841	0.846
300	0.523	0.590	0.736	0.705
400	0.510	0.518	0.662	0.656
gamma(2,1)(t=2)				
10	0.006	0.085	0.084	0.250
20	0.005	0.057	0.036	0.080
30	0.005	0.047	0.026	0.057
50	0.003	0.034	0.023	0.047
100	0.003	0.023	0.011	0.037
200	0.002	0.017	0.009	0.018
300	0.002	0.012	0.005	0.014
400	0.002	0.011	0.003	0.014
lognormal(1,1)(t=2)				
10	0.005	0.032	0.035	0.078
20	0.004	0.017	0.018	0.029
30	0.004	0.015	0.015	0.023
50	0.003	0.014	0.012	0.018
100	0.003	0.008	0.008	0.013
200	0.003	0.006	0.007	0.005
300	0.002	0.004	0.003	0.005
400	0.001	0.004	0.002	0.004

Since the Tsallis entropy tends to the Shannon entropy, as $\alpha \to 1$, we expect our estimators to be close to the true value of Shannon entropy for α values close to one, for example, $\alpha = 0.99$ and $\alpha = 1.01$. In the following, we will repeat the simulation results for these two α values. Tables $5-10$ contain RMSE and the absolute values of biases of the proposed estimators for $\alpha = 0.99$ and $\alpha = 1.01$ at different sample sizes. The bold values in the Tables $5-10$ are the absolute value of the biases and RMSE of the estimators, in which the true Shannon entropy is used instead of the true Tsallis entropy. As can be seen in Tables $5-10$, the bold values are close to their corresponding the biases and RMSEs. According to the results in Tables $5-10$, except for lognormal(1,1) and gamma (2,1) (for $\alpha = 0.99$), $\hat{\tau}_{\alpha_1}(X)$ performs consistently better than $\hat{\tau}_{\alpha_2}(X)$ respect to bias and RMSE values.

Table 5: Performance of $\hat{\tau}_{\alpha_1}(X)$ against $\hat{\tau}_{\alpha_2}(X)$ with $\alpha = 1.01$ for $beta(2,1)$.

n	bias ($\hat{\tau}_{\alpha_1}(X)$)	RMSE ($\hat{\tau}_{\alpha_1}(X)$)	bias($\hat{\tau}_{\alpha_2}(X)$),	RMSE ($\hat{\tau}_{\alpha_2}(X)$)
10	0.478	0.499	0.914	1.279
	(0.476)	**(0.497)**	**(0.912)**	**(1.278)**
20	0.328	0.367	0.861	1.094
	(0.326)	**(0.366)**	**(0.857)**	**(1.093)**
30	0.276	0.313	0.777	0.943
	(0.275)	**(0.312)**	**(0.775)**	**(0.914)**
50	0.238	0.276	0.765	0.896
	(0.236)	**(0.275)**	**(0.762)**	**(0.895)**
100	0.160	0.187	0.631	0.872
	(0.159)	**(0.186)**	**(0.628)**	**(0.871)**
200	0.121	0.144	0.614	0.760
	(0.119)	**(0.143)**	**(0.612)**	**(0.759)**
300	0.106	0.125	0.601	0.635
	(0.105)	**(0.124)**	**(0.599)**	**(0.634)**
400	0.094	0.108	0.434	0.473
	(0.092)	**(0.107)**	**(0.432)**	**(0.432)**

Table 6: Performance of $\hat{\tau}_{\alpha_1}(X)$ against $\hat{\tau}_{\alpha_2}(X)$ with $\alpha = 0.99$ for $beta(2,1)$.

n	bias ($\hat{\tau}_{\alpha_1}(X)$)	RMSE ($\hat{\tau}_{\alpha_1}(X)$)	bias($\hat{\tau}_{\alpha_2}(X)$),	RMSE ($\hat{\tau}_{\alpha_2}(X)$)
10	0.469	0.488	1.508	1.625
	(0.471)	**(0.489)**	**(1.506)**	**(1.651)**
20	0.321	0.347	1.167	1.312
	(0.323)	**(0.349)**	**(1.165)**	**(1.311)**
30	0.284	0.326	0.829	0.845
	(0.286)	**(0.327)**	**(0.826)**	**(0.844)**
50	0.225	0.263	0.736	0.754
	(0.227)	**(0.264)**	**(0.735)**	**(0.753)**
100	0.159	0.186	0.641	0.618
	(0.160)	**(0.187)**	**(0.640)**	**(0.617)**
200	0.129	0.155	0.606	0.615
	(0.131)	**(0.157)**	**(0.604)**	**(0.615)**
300	0.108	0.129	0.518	0.512
	(0.110)	**(0.130)**	**(0.518)**	**(0.513)**
400	0.095	0.119	0.503	0.511
	(0.096)	**(0.120)**	**(0.503)**	**(0.511)**

Table 7: Performance of $\hat{\tau}_{\alpha_1}(X)$ against $\hat{\tau}_{\alpha_2}(X)$ with $\alpha = 1.01$ for $gamma(2, 1)$.

n	bias ($\hat{\tau}_{\alpha_1}(X)$)	RMSE ($\hat{\tau}_{\alpha_1}(X)$)	bias($\hat{\tau}_{\alpha_2}(X)$),	RMSE ($\hat{\tau}_{\alpha_2}(X)$)
10	1.382	1.380	2.217	2.215
	(1.490)	**(1.488)**	**(2.325)**	**(2.324)**
20	0.736	0.735	1.843	1.840
	(0.845)	**(0.843)**	**(1.951)**	**(1.949)**
30	0.633	0.623	1.175	1.172
	(0.742)	**(0.731)**	**(1.067)**	**(0.067)**
50	0.509	0.506	0.677	0.674
	(0.617)	**(0.615)**	**(0.785)**	**(0.785)**
100	0.240	0.239	0.426	0.420
	(0.348)	**(0.346)**	**(0.534)**	**(0.531)**
200	0.151	0.150	0.332	0.331
	(0.260)	**(0.259)**	**(0.441)**	**(0.440)**
300	0.069	0.068	0.222	0.221
	(0.178)	**(0.177)**	**(0.330)**	**(0.328)**
400	0.059	0.058	0.127	0.126
	(0.168)	**(0.166)**	**(0.235)**	**(0.233)**

Table 8: Performance of $\hat{\tau}_{\alpha_1}(X)$ against $\hat{\tau}_{\alpha_2}(X)$ with $\alpha = 0.99$ for $gamma(2, 1)$.

n	bias ($\hat{\tau}_{\alpha_1}(X)$)	RMSE ($\hat{\tau}_{\alpha_1}(X)$)	bias($\hat{\tau}_{\alpha_2}(X)$),	RMSE ($\hat{\tau}_{\alpha_2}(X)$)
10	2.137	2.135	1.015	1.017
	(2.029)	**(2.027)**	**(1.123)**	**(1.125)**
20	1.490	1.489	0.888	0.889
	(1.382)	**(1.379)**	**(0.996)**	**(0.997)**
30	0.642	0.641	0.303	0.305
	(0.534)	**(0.531)**	**(0.411)**	**(0.413)**
50	0.335	0.352	0.095	0.096
	(0.247)	**(0.245)**	**(0.112)**	**(0.113)**
100	0.266	0.264	0.074	0.075
	(0.158)	**(0.155)**	**(0.108)**	**(0.109)**
200	0.208	0.206	0.019	0.020
	(0.117)	**(0.115)**	**(0.088)**	**(0.089)**
300	0.176	0.174	0.014	0.016
	(0.168)	**(0.165)**	**(0.022)**	**(0.023)**
400	0.067	0.065	0.002	0.002
	(0.041)	**(0.040)**	**(0.011)**	**(0.012)**

Table 9: Performance of $\hat{\tau}_{\alpha_1}(X)$ against $\hat{\tau}_{\alpha_2}(X)$ with $\alpha = 1.01$ for $lognormal(1,1)$.

n	bias ($\hat{\tau}_{\alpha_1}(X)$)	RMSE ($\hat{\tau}_{\alpha_1}(X)$)	bias($\hat{\tau}_{\alpha_2}(X)$),	RMSE ($\hat{\tau}_{\alpha_2}(X)$)
10	5.501	7.445	0.622	2.838
	(5.464)	**(7.418)**	**(0.586)**	**(2.830)**
20	2.860	3.534	0.235	1.658
	(2.824)	**(3.505)**	**(0.199)**	**(1.653)**
30	2.221	2.622	0.130	1.014
	(2.184)	**(2.591)**	**(0.125)**	**(1.010)**
50	1.485	1.839	0.128	0.862
	(1.449)	**(1.810)**	**(0.108)**	**(0.857)**
100	0.980	1.073	0.117	0.597
	(0.943)	**(1.040)**	**(0.080)**	**(0.595)**
200	0.873	1.031	0.053	0.171
	(0.837)	**(1.001)**	**(0.40)**	**(0.168)**
300	0.743	0.923	0.049	0.092
	(0.711)	**(0.897)**	**(0.031)**	**(0.089)**
400	0.685	0.802	0.036	0.065
	(0.622)	**(0.798)**	**(0.027)**	**(0.061)**

Table 10: Performance of $\hat{\tau}_{\alpha_1}(X)$ against $\hat{\tau}_{\alpha_2}(X)$ with $\alpha = 0.99$ for $lognormal(1,1)$.

n	bias ($\hat{\tau}_{\alpha_1}(X)$)	RMSE ($\hat{\tau}_{\alpha_1}(X)$)	bias($\hat{\tau}_{\alpha_2}(X)$),	RMSE ($\hat{\tau}_{\alpha_2}(X)$)
10	3.675	4.683	1.422	2.864
	(3.640)	**(4.665)**	**(1.385)**	**(2.846)**
20	3.587	4.342	1.156	2.112
	(3.553)	**(4.313)**	**(1.119)**	**(2.092)**
30	2.942	3.829	0.875	1.677
	(2.890)	**(3.793)**	**(0.838)**	**(1.658)**
50	1.756	2.580	0.663	1.142
	(1.745)	**(2.551)**	**(0.626)**	**(1.121)**
100	0.983	1.719	0.439	0.753
	(0.979)	**(1.676)**	**(0.402)**	**(0.732)**
200	0.851	0.975	0.238	0.253
	(0.849)	**(.971)**	**(0.200)**	**(0.218)**
300	0.713	0.894	0.181	0.089
	(0.709)	**(0.889)**	**(0.155)**	**(0.046)**
400	0.682	0.721	0.097	0.034
	(0.677)	**(0.719)**	**(0.062)**	**(0.019)**

At the end, a Monte Carlo simulation is carried out to support the asymptotic normality of the estimator $H_n = (nh_n)^{\frac{1}{2}}\left\{\frac{\hat{\tau}_{\alpha 2}(X)-\tau_{\alpha}(X)}{\sigma_{\tau}}\right\}$ given in Theorem 4.5. Our goal is to check whether H_n has a standard normal distribution. For chekings 500 samples of size $n = 400$, whose parent distribution is beta(2,1), are generated and H_n is obtained for each sample. This procedure is repeated by choosing $\alpha = 1.5, 2$ and 3. Then, in order to check the asymptotic normality of H_n, the histograms displayed, in the following figure, for different values of α.

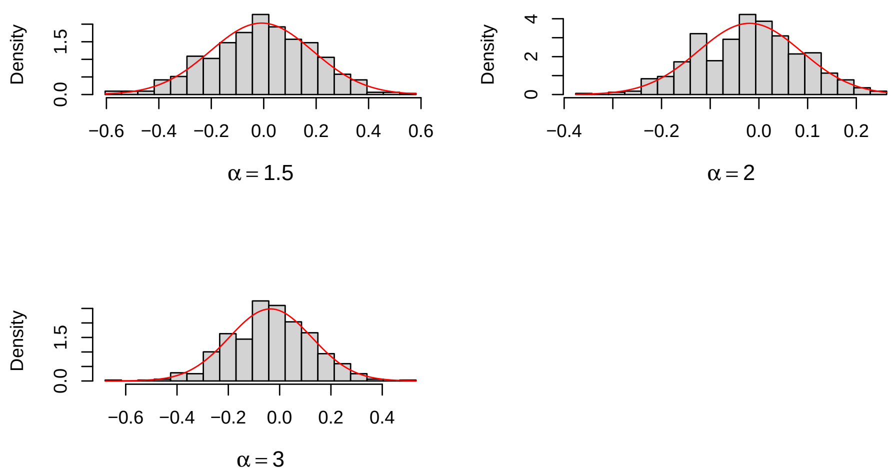

Figure 1: Histograms of H_n for different choices of α.

6 Real Data

Helu et al. (2020) used the failure times of aircraft windshields to illustrate the behaviour of the kernel density estimators based on progressive type-II censoring. They showed that these data are fitted using a gamma model with parameters $\alpha = 7.75$ and $\beta = 0.285$. Rajesh et al. (2022) modified the data to make it length-biased by selecting a sample ($n = 50$) without replacement by weighting each unit by its value. Here, we use these modified data to illustrate the proposed methods. The values of the proposed estimators corresponding to the real data are given in Table 11 From this table, it can be seen that $\hat{\tau}_{\alpha_1}(X)$ is closer to $\tau_{\alpha}(X)$ more than $\hat{\tau}_{\alpha_2}(X)$ and $\hat{\tau}_{\alpha_1}(X;t)$ is closer to $\tau_{\alpha}(X;t)$ more than $\hat{\tau}_{\alpha_2}(X;t)$.

Table 11: Values of $\tau_\alpha(X)$ and $\tau_\alpha(X;t)$ and their estimators for real data with (t=.5).

α	$\tau_\alpha(X)$	$\hat{\tau}_{\alpha_1}(X)$	$\hat{\tau}_{\alpha_2}(X)$	$\tau_\alpha(X;t)$	$\hat{\tau}_{\alpha_1}(X;t)$	$\hat{\tau}_{\alpha_2}(X;t)$
$\alpha=3$	0.499	0.407	0.401	0.499	0.407	0.401
$\alpha=4$	0.333	0.301	0.297	0.333	0.301	0.297

7 Conclusion

In this paper, we proposed two estimators for the Tsallis entropy of order α under the length-biased sampling model. The first estimator is based on the estimator defined for the distribution function F by Cox (1969). We showed that under some assumptions this estimator is an a.s. consistent estimator for the Tsallis entropy. The second estimator is obtained by the plug-in Jone's estimator in the Tsallis entropy formula. We investigated the a.a. consistency for this estimator. Also, we obtained the bias and variance of this estimator and proved the MISE of this estimator tends to zero for the large enough n. We also proved the asymptotic normality for this estimator. In the following, we proposed two non-parametric kernel estimators for the residual Tsallis entropy and investigated their a.s. consistency. Also, we showed that the MISE for the second estimator tends to zero for the large enough n. Next, to illustrate the behaviour of the proposed estimators, we exhibited a numerical evaluation of the estimators and carried out a Monte Carlo simulation study. We obtained bias and RMSE values of estimators for different α values and showed which estimator performs better for each of the selected distributions. In the end, a Monte Carlo simulation was carried out to support the asymptotic normality of the second estimator for the Tsallis entropy.

Appendix

Proof of Theorem 4.1 Using integration by parts, we can write:

$$\begin{aligned}\int_0^\infty (f_n(x))^{\alpha-1}dF_n(x) - \int_0^\infty (f(x))^{\alpha-1}dF(x) &= \int_0^\infty \Big((f_n(x))^{\alpha-1} - (f(x))^{\alpha-1}\Big)dF_n(x)\\ &+ \int_0^\infty (f(x))^{\alpha-1}d(F_n(x)-F(x))\\ &= \int_0^\infty \Big((f_n(x))^{\alpha-1} - (f(x))^{\alpha-1}\Big)dF_n(x)\\ &+ \int_0^\infty (F_n(x)-F(x))d(f(x))^{\alpha-1}. \qquad (7.1)\end{aligned}$$

A direct algebra shows that

$$\Big((f_n(x))^{\alpha-1} - (f(x))^{\alpha-1}\Big) = \sum_{i=1}^{\alpha-1}(f(x))^{i-1}(f_n(x))^{\alpha-1-i}(f_n(x)-f(x)). \qquad (7.2)$$

If $\lim_{n\to+\infty} \frac{\log n}{n^{\frac{1}{2}+\lambda} h_n} = 0$ for any $0 < \lambda < \frac{1}{2} - \frac{1}{r}$, and some $r > 2$, then under the Assumptions **A1** - **A10** from Theorem 2 of Ajami et al. (2013), we can write

$$\lim_{n\to\infty} \sup_{0\leq x<\tau} |f_n(x) - f(x)| = 0 \quad a.s. \tag{7.3}$$

Also (7.2) and (7.3) imply that

$$\lim_{n\to\infty} \sup_{0\leq x<\tau} \Big|(f_n(x))^{\alpha-1} - (f(x))^{\alpha-1}\Big| = 0 \quad a.s. \tag{7.4}$$

Chaubey et al. (2010) showed that

$$\lim_{n\to\infty} \sup_{0\leq x<\infty} |F_n(x) - F(x)| = 0 \quad a.s. \tag{7.5}$$

The desired conclusion now follows the a.s. convergence of $(f_n(x))^{\alpha-1}$ and $F_n(x)$ given in (7.4) and (7.5), respectively, and the expression given in (7.1).

Proof of Theorem 4.2. By using the Taylor's series expansion, we obtain:

$$\int_0^{+\infty} f_n^{\alpha}(x)\mathrm{d}x \simeq \int_0^{+\infty} f^{\alpha}(x)\mathrm{d}x + \alpha \int_0^{+\infty} f^{\alpha-1}(x)\Big(f_n(x) - f(x)\Big)\mathrm{d}x.$$

Under the conditions of the existence of $f''(t)$ and the finiteness of $E\Big[X^{-1}\Big]$, the bias and variance of $f_n(x)$ are given by (see Guillamòn et al. (1998))

$$\mathrm{Bias}\Big(f_n(x)\Big) \simeq \frac{1}{2} h_n^2 \mu_2(K) f''(x), \tag{7.6}$$

and

$$\mathrm{Var}\Big(f_n(x)\Big) = \frac{1}{nh_n} \mu x^{-1} f(x) ||K||_2^2, \tag{7.7}$$

where $\mu_2(K) = \int s^2 K(s)\mathrm{d}s$ and $||K||_2^2 = \int K^2(s)\mathrm{d}s$. Using the equation (7.6) and (7.7), the bias and the variance of $\int_0^{+\infty} f_n^{\alpha}(x)\mathrm{d}x$ are given by

$$\mathrm{Bias}\Big(\int_0^{+\infty} f_n^{\alpha}(x)\mathrm{d}x\Big) \simeq \frac{\alpha}{2} h_n^2 \mu_2(K) \int_0^{+\infty} f^{\alpha-1}(x) f''(x)\mathrm{d}x, \tag{7.8}$$

and

$$\mathrm{Var}\Big(\int_0^{+\infty} f_n^{\alpha}(x)\mathrm{d}x\Big) \simeq \frac{\alpha^2}{nh_n} \mu ||K||_2^2 \int_0^{+\infty} f^{2\alpha-1}(x) x^{-1}\mathrm{d}x. \tag{7.9}$$

The corresponding MSE is given by

$$\begin{aligned} \mathrm{MSE}\Big(\int_0^{+\infty} f_n^{\alpha}(x)\mathrm{d}x\Big) \simeq & \Big(\frac{\alpha}{2} h_n^2 \mu_2(K) \int_0^{+\infty} f^{\alpha-1}(x) f''(x)\mathrm{d}x\Big)^2 \\ & + \frac{\alpha^2}{nh_n} \mu ||K||_2^2 \int_0^{+\infty} f^{2\alpha-1}(x) x^{-1}\mathrm{d}x. \end{aligned} \tag{7.10}$$

From (7.10), as $n \to \infty$

$$\mathrm{MSE}(\int_0^{\infty} f_n^{\alpha}(x)\mathrm{d}x) \to 0.$$

Therefore

$$\hat{\tau}_{\alpha 2}(X) \longrightarrow \tau_{\alpha}(X).$$

Proof of Theorem 4.3. By using equations (7.8) and (7.9), the bias and variance of $\hat{\tau}_{\alpha 2}(X)$ have been obtained.

Proof of Theorem 4.4. Consider the mean integrated squared error (MISE) of the estimator $\hat{\tau}_{\alpha 2}(X)$. That is

$$\begin{aligned} \mathrm{MISE}\Big(\hat{\tau}_{\alpha 2}(X)\Big) &= E\int_0^{+\infty}\Big[\hat{\tau}_{\alpha 2}(X) - \tau_{\alpha}(X)\Big]^2 \mathrm{d}x, \\ &= \int_0^{+\infty} E\Big[\hat{\tau}_{\alpha 2}(X) - \tau_{\alpha}(X)\Big]^2 \mathrm{d}x, \\ &= \int_0^{+\infty} MSE\Big(\hat{\tau}_{\alpha 2}(X)\Big)\mathrm{d}x. \end{aligned} \tag{7.11}$$

Using (4.1) and (4.2), we obtain

$$\begin{aligned} \mathrm{MSE}\Big(\hat{\tau}_{\alpha 2}(X)\Big) \simeq & \Big(\frac{-\alpha}{2(\alpha-1)} h_n^2 \mu_2(K)\int_0^{+\infty} f^{\alpha-1}(x) f''(x)\mathrm{d}x\Big)^2 \\ & +(\frac{\alpha^2}{(\alpha-1)^2 n h_n}\mu ||K||_2^2 \int_0^{+\infty} f^{2\alpha-1}(x) x^{-1}\mathrm{d}x. \end{aligned} \tag{7.12}$$

From (7.12), one can derive $\lim_{n\to\infty} \mathrm{MSE}\Big(\hat{\tau}_{\alpha 2}(X)\Big) = 0$. Therefore, from (7.11)

$$\mathrm{MISE}\Big(\hat{\tau}_{\alpha 2}(X)\Big) \longrightarrow 0, \ as\, n \to \infty. \tag{7.13}$$

Proof of Theorem 4.5

$$\begin{aligned} (nh_n)^{\frac{1}{2}}\Big(\hat{\tau}_{\alpha 2}(X) - \tau_{\alpha}(X)\Big) &= \frac{-(nh_n)^{\frac{1}{2}}}{(\alpha-1)}\Big\{\int_0^{+\infty} f_n^{\alpha}(x)\mathrm{d}x - \int_0^{+\infty} f^{\alpha}(x)\mathrm{d}x\Big\} \\ &\simeq \frac{-\alpha(nh_n)^{\frac{1}{2}}}{(\alpha-1)}\Big\{\int_0^{+\infty} f^{\alpha-1}(x)\Big(f_n(x) - f(x)\Big)\mathrm{d}x\Big\}. \end{aligned} \tag{7.14}$$

By using the asymptotic normality of $f_n(x)$ given in Ajami et al. (2013), it is immediate that

$$(nh_n)^{\frac{1}{2}}\Big\{\frac{\hat{\tau}_{\alpha 2}(X) - \tau_{\alpha}(X)}{\sigma_{\tau}}\Big\}$$

is asymptotically normal with a mean of zero, variance of 1 and σ_τ^2 given in (4.4).

Proof of Theorem 4.6 First note that we have

$$\widehat{\tau}_{\alpha_1}(X;t) - \tau_\alpha(X;t) = \frac{1}{(\alpha-1)(\overline{F}_n(t))^\alpha}\left(\int_t^\infty (f(x))^{\alpha-1}dF(x) - \int_t^\infty (f_n(x))^{\alpha-1}dF_n(x)\right) + \frac{1}{(\alpha-1)}\left(\int_t^\infty (f(x))^{\alpha-1}dF(x)\right)\left(\frac{(\overline{F}_n(t))^\alpha - (\overline{F}(t))^\alpha}{(\overline{F}_n(t))^\alpha(\overline{F}(t))^\alpha}\right). \tag{7.15}$$

Now, a direct algebra shows that

$$(\overline{F}_n(t))^\alpha - (\overline{F}(t))^\alpha = \sum_{i=1}^{\alpha}(\overline{F}(t))^{i-1}(\overline{F}_n(t))^{\alpha-i}(\overline{F}_n(t) - \overline{F}(t)). \tag{7.16}$$

From (7.5) and (7.16) one can see that

$$\lim_{n\to\infty}\sup_{0\le t<\infty}\left|(\overline{F}_n(t))^\alpha - (\overline{F}(t))^\alpha\right| = 0 \quad a.s. \tag{7.17}$$

Also according to the proof of Theorem 4.1 we can write

$$\lim_{n\to\infty}\sup_{0<t<\infty}\left|\int_t^\infty (f(x))^{\alpha-1}dF(x) - \int_t^\infty (f(x))^{\alpha-1}dF_n(x)\right| = 0 \quad a.s. \tag{7.18}$$

So the proof is completed by using (7.15), (7.17) and (7.18) .

Proof of Theorem 4.7. By (7.5), we have

$$\overline{F}_n(y) = \frac{\mu_n}{n}\sum_{i=1}^{n}\frac{I(Y_i \ge y)}{Y_i} = \int_y^\infty dF_n(y) \longrightarrow \int_y^\infty dF(y) = \overline{F}(y) \quad a.s, \tag{7.19}$$

and $\frac{1}{\mu_n} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i} \longrightarrow \frac{1}{\mu} \quad a.s$, which implies

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i}I(Y_i \ge y) \longrightarrow \frac{\overline{F}(y)}{\mu} \quad a.s.$$

Consider the expansion

$$\begin{aligned}\overline{F}_n(y) &\simeq \overline{F}(y) + \mu\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i}I(Y_i \ge y) - \frac{\overline{F}(y)}{\mu}\right] \\ &- \mu\overline{F}(y)\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i} - \frac{1}{\mu}\right] \\ &\simeq \overline{F}(y) + \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\mu}{Y_i}I\left(Y_i \ge y\right) - \frac{\mu\overline{F}(y)}{Y_i}\right].\end{aligned} \tag{7.20}$$

From (7.20) it can be shown that

$$\text{Bias}(\overline{F}_n(t)) \simeq o(1), \tag{7.21}$$

and

$$\text{Var}(\overline{F}_n(t)) \simeq \frac{1}{n}\,\text{Var}\left[\frac{\mu}{Y} I\,(Y \geq t) - \frac{\mu \overline{F}(t)}{Y}\right]. \tag{7.22}$$

Using Taylor's series expansion and also (7.21) , (7.22), the expression for the bias of $\int_t^{\infty} (f_n(x))^{\alpha}\, dx$ and $\left(\overline{F}_n(t)\right)^{\alpha}$ are given by

$$\text{Bias}\left(\int_t^{\infty} (f_n(x))^{\alpha}\, dx\right) \simeq \frac{\alpha}{2} h_n^2 \mu_2(k) \int_t^{\infty} (f(x))^{\alpha-1} f''(x) dx, \tag{7.23}$$

$$\text{Bias}\left(\left(\overline{F}_n(t)\right)^{\alpha}\right) \simeq \alpha \left(\overline{F}(t)\right)^{\alpha-1} \text{Bias}\left(\overline{F}_n(t)\right) = o(1), \tag{7.24}$$

whereas the variances are given by

$$Var\Big(\int_t^{+\infty} f_n^{\alpha}(x)\mathrm{d}x\Big) \simeq \frac{\alpha^2}{nh_n} \mu \|K\|_2^2 \int_t^{+\infty} f^{2\alpha-1}(x) x^{-1}\mathrm{d}x, \tag{7.25}$$

and

$$\text{Var}\left(\left(\overline{F}_n(t)\right)^{\alpha}\right) \simeq \alpha^2 \left(\overline{F}(t)\right)^{2\alpha-2} \text{Var}\left(\overline{F}_n(t)\right). \tag{7.26}$$

Using (7.21)-(7.26) similar to the proof of Theorem 4.2 one can see that as $n \longrightarrow \infty$

$$\text{MSE}\left(\int_t^{\infty} (f_n(x))^{\alpha}\, dx\right) \longrightarrow 0,$$

and

$$\text{MSE}\left(\left(\overline{F}_n(t)\right)^{\alpha}\right) \longrightarrow 0.$$

Therefore

$$\widehat{\tau}_{\alpha_2}(X;t) = \frac{1}{\alpha-1}\left\{1 - \frac{\int_t^{+\infty} (f_n(x))^{\alpha} dx}{\overline{F}_n^{\alpha}(t)}\right\} \xrightarrow{\mathcal{P}} \frac{1}{\alpha-1}\left\{1 - \frac{\int_t^{+\infty} (f(x))^{\alpha} dx}{\overline{F}^{\alpha}(t)}\right\} = \tau_{\alpha}(X;t).$$

Proof of Theorem 4.8. Let $\widehat{M}_{\alpha}(t) = \int_t^{+\infty} (f_n(x))^{\alpha} dx$, $M_{\alpha}(t) = \int_t^{+\infty} (f(x))^{\alpha} dx$, $\widehat{A}_{\alpha}(t) = \left(\overline{F}_n(t)\right)^{\alpha}$, and $A_{\alpha}(t) = \left(\overline{F}(t)\right)^{\alpha}$. We can write:

$$\frac{\widehat{M}_{\alpha}(t)}{\widehat{A}_{\alpha}(t)} - \frac{M_{\alpha}(t)}{A_{\alpha}(t)} = \frac{1}{A_{\alpha}(t)}\left(\widehat{M}_{\alpha}(t) - \widehat{A}_{\alpha}(t)\frac{M_{\alpha}(t)}{A_{\alpha}(t)}\right)\left(1 + O_p(1)\right),$$

with $\left(\frac{A_{\alpha}(t)}{\widehat{A}_{\alpha}(t)} - 1\right) = O_p(1)$, since $\widehat{A}_{\alpha}(t) \xrightarrow{\mathcal{P}} A_{\alpha}(t)$. Therefore

$$\widehat{\tau}_{\alpha 2}(X;t) - \tau_{\alpha}(X;t) \simeq \frac{-1}{(\alpha-1)A_{\alpha}(t)}\left(\widehat{M}_{\alpha}(t) - \widehat{A}_{\alpha}(t)\frac{M_{\alpha}(t)}{A_{\alpha}(t)}\right). \tag{7.27}$$

By using (7.27) and Equations (7.23)- (7.26) the bias and variance of $\widehat{\tau}_{\alpha_2}(X;t)$ are obtained as follows

$$\text{Bias}\left(\widehat{\tau}_{\alpha_2}(X;t)\right) \simeq \frac{-\alpha h_n^2}{2(\alpha-1)} \frac{\mu_2(k)}{\left(\bar{F}(t)\right)^{\alpha}} \int_t^{+\infty} (f(x))^{\alpha-1} f''(x)dx, \tag{7.28}$$

and

$$\begin{aligned}\text{Var}\left(\hat{\tau}_{\alpha_2}(X;t)\right) \simeq & \frac{\alpha^2}{(\alpha-1)^2} \frac{1}{\bar{F}^{2\alpha}(t)} \left\{ \frac{\mu \| K \|_2^2}{n h_n} \int_t^{+\infty} x^{-1} f^{2\alpha-1}(x)dx \right. \\ & \left. + \frac{1}{n(\bar{F}(t))^2} \text{Var}\left(\frac{\mu}{Y} I(Y \geq t) - \mu \frac{\bar{F}(t)}{Y} \right) \left(\int_t^{+\infty} f^{\alpha}(x)dx \right)^2 \right\}.\end{aligned} \tag{7.29}$$

Using (7.28) and (7.29), it can be shown that $\lim_{n\to\infty} \text{MSE}\left(\widehat{\tau}_{\alpha_2}(X,t)\right) = 0$, and hence $\lim_{n\to\infty} \text{MISE}\left(\widehat{\tau}_{\alpha_2}(X,t)\right) = 0$.

References

Addona, V., and Wolfson, DB. (2006). A formal test for the stationarity of the incidence rate using data from a prevalent cohort study with follow-up. *Lifetime Data Anal*, **12**, 267 – 284.

Asgharian, M., M'Lan, CE., Wolfson, DB. (2002). Length-biased sampling with right censoring: an unconditional approach. *JASA*, **97**(457), 201 – 209.

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.

Baratpour, S., Khammar, A. (2016). Results on Tsallis entropy of order statistics and record values. *Istat. J. Turk. Stat. Assoc*, **8**, 60-73.

Borrajo, M. I., Gonzàlez-Manteiga, W., Martìnez-Miranda, M. D. (2017). Bandwidth selection for kernel density estimation with length-biased data. *Journal of Nonparametric Statistics*, **29**(3), 636 – 668.

Chaubey, YP., Sen, PK., Li, J. (2010). Smooth density estimation for length-biased data. *J. Indian Soc. Agric. Stat*, **64**(2), 145 – 155.

Cox, D. R. (1969). *Some sampling problems in technology. In New Developments in Survey Sampling*. N. L. Johnson and H. Smith (eds.). Wiley Interscience, 506-527.

Chakraborty, S., Pradhan, B. (2023). Generalized weighted survival and failure entropies and their dynamic versions. *Commun.-Stat.-Theory Methods*, **52**, 730-750.

Cartwright, J., and Roll, J. (2014). over Boltzmann. Phys.World. **27** 31 – 35.

Guillamòn, A., Navarro, J., Ruiz, JM. (1998). Kernel density estimation using weighted data. *Comm.Statist. Theory Methods*, **27**(9), 2123 – 2135.

Helu, A., Samawi, H., Rochani, H., Yin, J., Vogel, R. (2020). Kernel density estimation based on progressive type-II censoring. *J. Korean Stat. Soc.*, **49**(2), 475–498.

Iranmanesh, A., Oliazadeh, F., Fakoor, V. (2022). Estimation of Past Entropy in Length-Biased Data. *Journal of Statistical Sciences*, **15**(2), 363–380.

Jones, MC. (1991). Kernel density estimation for length-biased data. *Biometrika*, **78**(3), 511 − 519.

Kumar, V., and Taneja, HC. (2011). A generalized entropy-based residual lifetime distributions. *Int.J.Biomath*, **4**(2), 171 − 184.

Misagh, F., and Yari, G. (2012). Interval entropy and informative distance. *Entropy*, **14**, 480-490.

Naudts, J. (2011) Generalised thermostatistics. *Springer Science & Business Media*.

Nanda, AK., and Paul, P. (2006). Some results on generalized residual entropy. *Inform.Sci*, **176**, 27 − 47.

Rényi, A. (1961). On measures of entropy and information. booktitleProceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics, *The Regents of the University of California,1961*, **4**, 547-562.

Rajesh, G., and Sunoj, S. (2019). Some properties of cumulative Tsallis entropy of order α. *Stat. Pap*, **60**, 933-943.

Rajesh, R., Rajesh, G., Sunoj, SM. (2022). Kernel estimation of extropy function under length-biased sampling. *Statistics and Probability Letters*, **181**, 109290.

Rajagopalan, V., Ganaie, R. A., Rather, A. A. (2019). A New Length biased distribution with Applications. *Science, Technology and Development*, **8**(8), 161-174.

Shannon, C. (1948). A mathematical theory of communication. *Bell Syst. Tech. J*, **27**, 379 − 432.

Shen, Y., Ning, J., Qin, J. (2009). Analyzing length-biased data with semiparametric transformation and accelerated failure time models. *Journal of the American Statistical Association*, **104**(487), 1192 − 1202.

Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys*, **52**, 479 − 487.