Technical Report No. 2/20, April 2020

ON AN ENTROPY ESTIMATOR BASED ON A NON-PARAMETRIC DENSITY ESTIMATOR FOR NON-NEGATIVE DATA

Yogendra P. Chaubey and Nhat Linh Vu

On An Entropy Estimator Based On a Non-parametric Density Estimator For Non-negative Data

Yogendra P. Chaubey
Department of Mathematics and Statistics
Concordia University, Montreal, Canada H4B 1R6
E-mail: yogen.chaubey@concordia.ca

Nhat Linh Vu
Department of Mathematics and Statistics
Concordia University, Montreal, Canada H4B 1R6
E-mail: nhatlinh.vu@concordia.ca

Abstract

In the recent decades, entropy has become more and more essential in statistics and machine learning. It features in many applications involving data transmission, cryptography, signal processing, network theory, bio-informatics, and so on. A large number of estimators for entropy have been proposed in the past ten years. Here we focus on entropy estimation for non-negative random variables. Specifically, the use of entropy estimator based on Poisson-weights density estimator is found to be of interest. We establish some asymptotic properties of the resulting estimators and present a simulation study comparing these with well known estimators in literature.

Keywords: information theory, entropy estimator, non-parametric density estimator, asymptotic properties.

1 Introduction

Let X be a continuous random variable with the probability density function (pdf) $f(x)$. The entropy of the random variable X is defined as

$$H_f = -\int_{-\infty}^{\infty} f(x) \log f(x) dx. \tag{1.1}$$

The concept of *entropy*, also known as *differential entropy* was introduced by Shannon (1948) in information theory and consequently, it has attracted a lot of interests in many areas of statistical applications such as data transmission, cryptography, signal processing, network theory, bio-informatics, and so on.

Jaynes (1957a, 1957b) introduced the *maximum entropy principle* for statistical inference, while establishing a connection of differential entropy to the concept of thermodynamic entropy in statistical mechanics. Chaubey and Mudholkar (2013) provided an entropy based rationale for maximum likelihood principle. We refer to Beirlant *et al.* (2001) for an excellent review of various estimation methods of H_f. The reader is also directed to the excellent text by Cover and Thomas (1991) covering various aspects of entropy theory and applications. This paper concerns the estimator of entropy H_f, that is obtained by plugging in a non-parametric estimator $\hat{f}_n(x)$ for $f(x)$ that is based on a random sample $\{X_1, ..., X_n\}$ obtained from the density $f(x)$. This estimator was proposed by Ahmad and Lin (1989) when $\hat{f}_n(x)$ is obtained by kernel smoothing. Such an estimator is constructed implicitly for densities on the whole real line. These methods, however, breakdown for estimating the densities over a bounded interval or over a subset of the real line (see e.g. Karunamuni and Albert 2005). In case we have a density supported on R^+, Chaubey and Sen (1996) and Chaubey *et al.* (2012) proposed non-parametric density estimators using Poisson weights and asymmetric kernels, respectively. Asymmetric kernel density estimators have also been proposed and studied by other authors, e.g. Bagai and Prakasa Rao (1996), Chen (2000) and Scaillet (2004) among others. Bouezmarni and Scaillet (2005) have studied the consistency of such estimators.

Eggermont and LaRiccia (1999) established *best asymptotic normality* of the kernel based entropy estimator and Schwartz *et al.* (2005) proposed an efficient algorithm for kernel entropy estimation using Gaussian kernel. However, since the underlying density estimator may not be appropriate in the current setting of non-negative variables, we would like to consider the plug-in estimator of entropy using the density estimator from Chaubey and Sen (1996). This estimator is known to be better than asymmetric kernel estimator as it handles the estimation of density at zero in a better manner (see Bouezmarni and Scaillet 2005). The original Poisson weights estimator, as proposed in Chaubey and Sen (1996) has been modified somewhat (see Chaubey *et al.* 2010) where we did not truncate the Poisson distribution at a fixed index N as considered in Chaubey and Sen (1996).

There exist quite a few entropy estimators, as seen in the paper by Beirlant *et al.* (2001). However, we have not seen plug-in estimators for non-negative random variables, a case that presents itself in many applications such as in reliability and life testing. Furthermore, the Poisson based estimator is faster to compute as opposed to the asymmetric kernel based estimator. Our purpose in this paper is to establish some asymptotic properties of the resulting entropy estimator and present a numerical study comparing it to some standard estimators in literature. Some important papers studying asymptotic properties of entropy estimators may be mentioned. Györfi and van der Meulen (1987, 1990) studied convergence properties of various entropy estimators including those based kernel plug-in estimator; Van Es (1992) studied asymptotic properties of the so-called "*spacing*" approach; Hall and Morton (1993) studied the asymptotic properties when an empirical version of the plug-in estimator is employed; Eggermont and LaRiccia (1999) used double exponential kernel on density estimation; and Bouzebda *et al.* (2013) relied on the quantile density estimation and Brownian bridge.

Our goal in this paper is to first present two new entropy estimators, and then study their asymptotic properties. The organization of the paper is as follows. The new entropy estimators are presented in Section 2 along with their asymptotic properties. Section 3 is dedicated to a brief review on existing entropy estimators along with a simulation study comparing our estimators with the existing ones.

2 New Entropy Estimators for Non-negative Support

One straightforward approach for entropy estimation is to estimate the underlying density function $f(x)$ by some well-known density estimator $\hat{f}_n(x)$, then plug it into (1.1) to obtain the entropy estimator

$$\hat{H}_f^{Plugin} = -\int_0^\infty \hat{f}_n(x) \log \hat{f}_n(x) dx. \tag{2.1}$$

On the other hand, motivated by the representation of entropy as an expected value

$$H_f = -\int_0^\infty f(x) \log f(x) dx = -\mathbb{E}[\log f(X)], \tag{2.2}$$

it follows by the strong law of large numbers that

$$-\frac{1}{n}\sum_{i=1}^{n} \log f(X_i) \overset{a.s.}{\to} H_f \text{ as } n \to \infty. \tag{2.3}$$

Thus we obtain a new entropy estimator if we replace $f(\cdot)$ by an appropriate density estimator $\hat{f}_n(\cdot)$, as given by

$$\hat{H}_f^{Meanlog} = -\frac{1}{n}\sum_{i=1}^{n} \log \hat{f}_n(X_i). \tag{2.4}$$

Hence, with a well-performing density estimator on hand, we may obtain a good entropy estimator. The fixed symmetric kernel density estimator is a well-known and popular approach for estimating the density function with an unbounded support, but it results in a heavy bias near the boundary when dealing with densities of non-negative random variable, as mentioned earlier. In order to alleviate this problem, we focus on the Poisson smoothed histogram density estimator, that has the following form (see Chaubey and Sen 2009),

$$\hat{f}_n^{Pois}(x) = k \sum_{i=0}^{\infty} \left[F_n\left(\frac{i+1}{k}\right) - F_n\left(\frac{i}{k}\right) \right] e^{-kx} \frac{(kx)^i}{i!}, \tag{2.5}$$

where $F_n(\cdot)$ is the empirical distribution function, and $k := k(n)$ can be viewed as the smoothing parameter. The estimator $\hat{f}_n^{Pois}(x)$ can be interpreted as a random weighted sum of Poisson probabilities. The asymptotic properties of $\hat{f}_n^{Pois}(x)$ have been studied and its weak convergence was proven by Bouezmarni and Scaillet (2005) under the assumptions $\lim_{n\to\infty} k = \infty$ and $\lim_{n\to\infty} nk^{-2} = \infty$. They also obtained the weak convergence for the case of unbounded pdf f at $x = 0$. Later on, Chaubey *et al.* (2010) provided further results by working out the asymptotic bias, asymptotic variance, strong consistency and the asymptotic normality of the estimator. Particularly, under the assumptions that $k = cn^h$ for some constant $c > 0$ and $0 < h < 1$, and $f'(x)$ satisfies the Lipschitz condition of order $\alpha > 0$, the asymptotic bias and variance of $\hat{f}_n^{Pois}(\cdot)$ are given by

$$Bias\left[\hat{f}_n^{Pois}(x)\right] \approx \frac{f'(x)}{2cn^h}, \tag{2.6}$$

$$Var\left[\hat{f}_n^{Pois}(x)\right] \approx \frac{\mathbb{E}[X]}{2}\sqrt{\frac{c}{2\pi x^3}} f(x) n^{h/2-1}. \tag{2.7}$$

The strong consistency of the estimator $\hat{f}_n^{Pois}(x)$,

$$||\hat{f}_n^{Pois}(x) - f(x)|| \stackrel{a.s.}{\to} 0 \tag{2.8}$$

is established under the conditions $\mathbb{E}[X^{-2}] < \infty$ and $f'(x)$ is bounded and $k_n = O(n^h)$. On the other hand, the asymptotic normality of $\hat{f}_n^{Pois}(x)$ is established under the conditions $\mathbb{E}[X^{-2}] < \infty$, $k_n = O(n^{2/5})$, and $f'(x)$ satisfies the Lipschitz order α condition. Namely, under these conditions, for x in a compact set $I \subset \mathbb{R}^+$, we have

$$n^{2/5}(\hat{f}_n^{Pois}(x) - f(x)) - \frac{1}{2\delta^2} f'(x) \stackrel{D}{\to} \mathcal{G}, \tag{2.9}$$

where $\mathcal{G}$ is a zero-mean Gaussian process with covariance function $\gamma_x^2 \delta_{xs}$, where

$$\gamma_x^2 = \frac{\mathbb{E}[X]}{2}(2\pi x^3)^{-1/2} f(x)\delta,$$

with δ and δ_{xs} are given by

$$\delta = \lim_{n\to\infty}(n^{-1/5} k_n^{1/2}), \qquad \text{and} \qquad \delta_{xs} = \begin{cases} 0 \text{ for } x \neq s, \\ 1 \text{ for } x = s. \end{cases}$$

With these nice asymptotic properties of $\hat{f}_n^{Pois}(\cdot)$, we propose the following two entropy estimators:

$$\hat{H}_f^{Plugin-Pois} = -\int_0^\infty \hat{f}_n^{Pois}(x) \log \hat{f}_n^{Pois}(x) dx, \tag{2.10}$$

$$\hat{H}_f^{Meanlog-Pois} = -\frac{1}{n}\sum_{i=1}^n \log \hat{f}_n^{Pois}(X_i). \tag{2.11}$$

We are able to establish the asymptotic consistency and normality of these estimators as stated in the theorems that follow. The following theorem concerns of the asymptotic properties of $\hat{H}_f^{Meanlog-Pois}$.

Theorem 2.1. *Assume the following conditions to hold:*

- $\mathbb{E}[X^{-2}] < \infty$,
- $\mathbb{E}[(\log f(X))^2] < \infty$,
- $f'(x)$ *is bounded with* $\int_0^\infty f'(x)dx < \infty$, *and satisfies Lipschitz order of* α *condition,*
- $f(x)$ *is twice differentiable and* $\int_0^\infty \frac{f''(x)}{f(x)}dx < \infty$,
- $k_n = o(n^h)$ *for some* $h \in (1/2, 1)$,

then

$$\hat{H}_f^{Meanlog-Pois} = -\frac{1}{n}\sum_{i=1}^n \log f(X_i) + o(n^{-1/2}) \ a.s. \text{ as } n \to \infty. \tag{2.12}$$

Consequently, we get

$$\left| \hat{H}_f^{Meanlog-Pois} - H_f \right| \xrightarrow{a.s.} 0, \tag{2.13}$$

and

$$\sqrt{n}\left(\hat{H}_f^{Meanlog-Pois} - H_f \right) \xrightarrow{D} \mathcal{N}(0, Var[\log f(X)]). \tag{2.14}$$

Proof. Writing $\hat{H}_f^{Meanlog-Pois}$ as an integral with respect to the empirical distribution function $F_n(x)$, we have

$$\begin{aligned} -\frac{1}{n}\sum_{i=1}^n \log \hat{f}_n^{Pois}(X_i) &= -\int_0^\infty \log \hat{f}_n^{Pois}(x) dF_n(x) \\ &= -\int_0^\infty \log f(x) dF_n(x) - \int_0^\infty \Big(\log \hat{f}_n^{Pois}(x) - \log f(x) \Big) dF_n(x) \\ &= -\frac{1}{n}\sum_{i=1}^n \log f(X_i) - I_n, \end{aligned}$$

where

$$I_n := \int_0^\infty \Big(\log \hat{f}_n^{Pois}(x) - \log f(x) \Big) dF_n(x).$$

On the other hand, we know that $-\frac{1}{n}\sum_{i=1}^n \log f(X_i)$ is an unbiased, strongly consistent estimator for H_f by the strong law of large numbers, and is root-n asymptotic normal by the *central limit theorem* if $\mathbb{E}[\log f(X)^2] < \infty$. That is

$$\left| -\frac{1}{n}\sum_{i=1}^n \log f(X_i) - H_f \right| \overset{a.s.}{\to} 0,$$

and

$$\sqrt{n}\Big(-\frac{1}{n}\sum_{i=1}^n \log f(X_i) - H_f \Big) \overset{D}{\to} \mathcal{N}(0, \mathrm{Var}[\log f(X)]),$$

as $n \to \infty$. Therefore, it is sufficient to prove that

$$I_n = o(n^{-1/2}) \text{ a.s. as } n \to \infty.$$

In order to study the asymptotic behavior of I_n, we decompose it into two parts as,

$$\begin{aligned} I_n &= \int_0^\infty \Big(\log \hat{f}_n^{Pois}(x) - \log f(x) \Big) d(F_n(x) - F(x)) + \int_0^\infty \Big(\log \hat{f}_n^{Pois}(x) - \log f(x) \Big) dF(x) \\ &= I_{n,1} + I_{n,2}, \text{ say.} \end{aligned}$$

- Analysis of $I_{n,2}$:

  Since the function $\log z$ is continuous and differentiable for all $z > 0$, we can apply the Taylor expansion centering at a to get

  $$\log z = \log a + \frac{z - a}{tz + (1-t)a},$$

  where $t \in (0, 1)$. By letting $z = \hat{f}_n^{Pois}(x)$ and $a = f(x)$, we obtain

  $$\log \hat{f}_n^{Pois}(x) - \log f(x) = \frac{\hat{f}_n^{Pois}(x) - f(x)}{t\hat{f}_n^{Pois}(x) + (1-t)f(x)}.$$

  As $||\hat{f}_n^{Pois}(x) - f(x)|| \overset{a.s.}{\to} 0$ uniformly, it implies $t\hat{f}_n^{Pois}(x) + (1-t)f(x) \overset{a.s.}{\to} f(x)$. Thus,

  $$\log \hat{f}_n^{Pois}(x) - \log f(x) \overset{a.s.}{\to} \frac{\hat{f}_n^{Pois}(x) - f(x)}{f(x)}.$$

As a result, $I_{n,2}$ can be expressed as

$$\begin{aligned} I_{n,2} &\overset{a.s.}{\to} \int_0^\infty \left(\frac{\hat{f}_n^{Pois}(x) - f(x)}{f(x)} \right) f(x)dx \\ &= \int_0^\infty \left(\hat{f}_n^{Pois}(x) - f(x) \right) dx \\ &= \int_0^\infty \hat{f}_n^{Pois}(x)dx - \int_0^\infty f(x)dx \\ &= 0, \end{aligned}$$

that follows, since the Poisson smooth density estimator integrates to unity.

<u>Analysis of $I_{n,1}$:</u>

Using integration by part we have

$$\begin{aligned} I_{n,1} &= \big(F_n(x) - F(x)\big) \log \left(\frac{\hat{f}_n^{Pois}(x)}{f(x)} \right) \Bigg|_0^\infty \\ &\quad - \int_0^\infty \left(\frac{\hat{f}_n'^{Pois}(x)}{\hat{f}_n^{Pois}(x)} - \frac{f'(x)}{f(x)} \right) \big(F_n(x) - F(x)\big) dx. \end{aligned}$$

It is well-known that by the law of the iterated logarithm, we have

$$||F_n - F||_\infty = O\big(n^{-1/2}(\log\log n)^{1/2}\big) \text{ a.s.}$$

Meanwhile, Recall from Chaubey *et al.* (2010) that if $f'(x)$ satisfies the Lipschitz of order $\alpha > 0$, i.e. there exists a finite positive K such that

$$|f'(s) - f'(t)| \leq K|s-t|^\alpha \qquad \forall s, t \in \mathbb{R}^+,$$

then for fixed $x \in \mathbb{R}$ we have

$$\hat{f}_n^{Pois}(x) - f(x) = \frac{1}{2k} f'(x) + O(k^{-1-\alpha}), \tag{2.15}$$

which implies that

$$\frac{\hat{f}_n^{Pois}(x)}{f(x)} = 1 + \frac{1}{2k} \frac{f'(x)}{f(x)} + O(k^{-k-\alpha}) = 1 + O(k^{-1}). \tag{2.16}$$

If we restrict $k = o(n^h)$ for $1/2 < h < 1$, then for fixed $x \in \mathbb{R}$ we get

$$\sqrt{n}\Big(F_n(x) - F(x)\Big) \log\left(\frac{\hat{f}_n^{Pois}(x)}{f(x)}\right) = O\big((\log\log n)^{1/2}\big)\log(1 + o(n^{-h})).$$

By L'Hôpital's rule, we get

$$\begin{aligned}
\lim_{n\to 0}(\log\log n)^{1/2}\log(1+n^{-h}) &= \lim_{n\to 0} 2h\frac{\big(\log\log n\big)^{3/2}\log n}{n^h} \\
&= \lim_{n\to 0}\frac{2\big(\log\log n\big)^{3/2} + 3\big(\log\log n\big)^{1/2}}{n^h} \\
&= \lim_{n\to 0}\frac{3\big(\log\log n\big)^{1/2} + \frac{3}{2}\big(\log\log n\big)^{-1/2}}{hn^h\log n} \\
&= 0.
\end{aligned}$$

Thus, we obtain

$$\sqrt{n}\left[\big(F_n(x) - F(x)\big)\log\left(\frac{\hat{f}_n^{Pois}(x)}{f(x)}\right)\Bigg|_0^\infty\right] \overset{a.s.}{=} o(1),$$

which means that

$$\left[\big(F_n(x) - F(x)\big)\log\left(\frac{\hat{f}_n^{Pois}(x)}{f(x)}\right)\Bigg|_0^\infty\right] \overset{a.s.}{=} o(n^{-1/2}).$$

On the other hand, since $||\hat{f}_n^{Pois}(x) - f(x)|| \overset{a.s.}{\to} 0$ uniformly, we can bound the second term of $I_{n,1}$ as

$$\begin{aligned}
&\left|\int_0^\infty \left(\frac{\hat{f}_n'^{Pois}(x)}{\hat{f}_n^{Pois}(x)} - \frac{f'(x)}{f(x)}\right)\big(F_n(x) - F(x)\big)dx\right| \\
\overset{a.s.}{\to}\ &\left|\int_0^\infty \left(\frac{\hat{f}_n'^{Pois}(x)}{f(x)} - \frac{f'(x)}{f(x)}\right)\big(F_n(x) - F(x)\big)dx\right| \\
\leq\ &||F_n - F||_\infty \int_0^\infty \frac{\left|\hat{f}_n'^{Pois}(x) - f'(x)\right|}{f(x)}dx.
\end{aligned}$$

By differentiating on both sides of (2.15) and dividing by $f(x)$, we get

$$\frac{\hat{f}_n'^{Pois}(x) - f'(x)}{f(x)} = \frac{1}{2k}\frac{f''(x)}{f(x)} + O(k^{-1-\alpha}).$$

So, if $k = o(n^h)$ where $1/2 < h < 1$ and $\int_0^\infty \frac{f''(x)}{f(x)} dx < \infty$, then

$$\begin{aligned} \left| \int_0^\infty \left(\frac{\hat{f}_n'^{Pois}(x)}{\hat{f}_n^{Pois}(x)} - \frac{f'(x)}{f(x)} \right) \big(F_n(x) - F(x)\big) dx \right| &\leq O\big(n^{-1/2} (\ln \ln n)^{1/2}\big) o(n^{-1/2} \\ &= o(n^{-1/2}). \end{aligned}$$

Therefore, $I_{n,1} = o(n^{-1/2})$ a.s. Putting everything together, we get

$$\hat{H}_f^{Meanlog-Pois} = -\frac{1}{n} \sum_{i=1}^n \log f(X_i) + o(n^{-1/2}) \text{ almost surely.}$$

This completes the proof of the theorem. □

Next we establish asymptotic properties of the plug-in estimator $\hat{H}_f^{Plugin-Pois}$. First, we establish its asymptotic consistency as given in the following theorem.

Theorem 2.2. *Assume the following conditions:*

- $\mathbb{E}[X^{-2}] < \infty$,
- $f'(x)$ *is bounded, satisfies Lipschitz order of* α *condition,*
- $\int_0^\infty f'(x) \log f(x) dx < \infty$,
- $k_n = o(n^h)$ *for some* $0 < h < 1$,

then,

$$|\hat{H}_f^{Plugin-Pois} - H_f| \stackrel{a.s.}{\to} 0. \tag{2.17}$$

Proof. We have

$$\begin{aligned} &\hat{H}_f^{Plugin-Pois} \\ &= -\int_0^\infty \hat{f}_n^{Pois}(x) \log \hat{f}_n^{Pois} dx \\ &= -\int_0^\infty \left(\hat{f}_n^{Pois}(x) - f(x) + f(x) \right) \log \hat{f}_n^{Pois}(x) dx \\ &= -\int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x) \big) \log \hat{f}_n^{Pois}(x) dx - \int_0^\infty f(x) \big(\log \hat{f}_n^{Pois}(x) - \log f(x) \big) dx + H_f \\ &= H_f - U_n - I_{n,2}, \end{aligned} \tag{2.18}$$

where $U_n := \int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big) \log \hat{f}_n^{Pois}(x) dx$. From the proof of the Theorem 2.1, we already have

$$I_{n,2} \stackrel{a.s.}{\to} 0.$$

For analysing U_n, we can decompose it as

$$U_n = \int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\big(\log \hat{f}_n^{Pois}(x) - \log f(x)\big)dx + \int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\log f(x)dx.$$

Since under the assumptions in the theorem, $||\hat{f}_n^{Pois}(x) - f(x)|| \overset{a.s.}{\to} 0$ uniformly, we get $\hat{f}_n^{Pois}(x)/f(x) \overset{a.s.}{\to} 1$ uniformly. Thus, by the dominant convergence theorem (DCT), the first term of U_n becomes

$$\begin{aligned} & \int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\big(\log \hat{f}_n^{Pois}(x) - \log f(x)\big)dx \\ = & \int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\log\left(\frac{\hat{f}_n^{Pois}(x)}{f(x)}\right)dx \\ \overset{a.s.}{\to} & \ 0. \end{aligned}$$

And for the second term of U_n, recall that if $f'(x)$ satisfies the Lipschitz of order $\alpha > 0$, i.e. there exists a finite positive K such that

$$|f'(s) - f'(t)| \leq K|s-t|^\alpha \qquad \forall s,t \in \mathbb{R}^+,$$

then for fixed $x \in \mathbb{R}$ we have

$$\hat{f}_n^{Pois}(x) - f(x) = \frac{1}{2k}f'(x) + O(k^{-1-\alpha}),$$

Thus, we get

$$\int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\log f(x)dx = \frac{1}{2k}\int_0^\infty f'(x)\log f(x)dx + O(k^{-1-\alpha}) = o(1),$$

given that $k = o(n^h)$ and $\int_0^\infty f'(x)\log f(x)dx < \infty$. Therefore,

$$|\hat{H}_f^{Plugin-Pois} - H_f| \leq |U_n| + |I_{n,2}| \overset{a.s.}{\to} 0.$$

This establishes the strong-consistency of $\hat{H}_f^{Plugin-Pois}$. □

For the asymptotic normality of $\hat{H}_f^{Plugin-Pois}$, we present the following theorem.

Theorem 2.3. *If these conditions hold*

- $\mathbb{E}[X^{-2}] < \infty$,
- $f'(x)$ *is bounded, satisfies Lipschitz order of* α *condition,*

- $\int_0^\infty f'(x)\log f(x)dx < \infty$,
- $\int_0^\infty x^{-3/2} f(x)\big(\log f(x)\big)^2 dx < \infty$,
- $k_n = cn^{2/5}$ *for a constant* $c > 0$,

then, we have

$$n^{2/5}\left(\hat{H}_f^{Plugin-Pois} - H_f\right) \xrightarrow{D} \mathcal{N}, \tag{2.19}$$

where $\mathcal{N}$ *is a normal random variable with mean* $\frac{1}{2c}\int_0^\infty f'(x)\log f(x)dx$ *and variance* $\frac{\mathbb{E}[X]}{2}\sqrt{\frac{c}{2\pi}}\int_0^\infty x^{-3/2} f(x)(\log f(x))^2 dx$.

Proof. From the proof of the previous theorem, we have

$$\hat{H}_f^{Plugin-Pois} - H_f = U_n + I_{n,2},$$

where $I_{n,2} \xrightarrow{a.s.} 0$ identically, and

$$U_n = \int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\log\left(\frac{\hat{f}_n^{Pois}(x)}{f(x)}\right)dx + \int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\log f(x)dx.$$

Recall that if $f'(x)$ satisfies the Lipschitz of order $\alpha > 0$, i.e. there exists a finite positive K such that

$$|f'(s) - f'(t)| \le K|s-t|^\alpha \qquad \forall s,t \in \mathbb{R}^+,$$

then for fixed $x \in \mathbb{R}$ we have

$$\hat{f}_n^{Pois}(x) - f(x) = \frac{1}{2k}f'(x) + O(k^{-1-\alpha}).$$

Consider the quantity $n^{2/5}U_n$, then, under the assumption that $k = cn^{2/5}$ for a constant $c > 0$, we first get

$$n^{2/5}\int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\log\left(\frac{\hat{f}_n^{Pois}(x)}{\log f(x)}\right)dx = o(1),$$

due the dominant convergence theorem and the fact that

$$\lim_{n\to\infty} n^{2/5}cn^{-2/5}\log(1 + cn^{-2/5}) = \lim_{n\to\infty} c\log(1 + cn^{-2/5}) = 0.$$

As a result, the limiting distribution of $n^{2/5}\big(\hat{H}_f^{Plugin-Pois} - H_f\big)$ follows the same as the limiting distribution of $n^{2/5}\int_0^\infty \big(\hat{f}_n^{Pois}(x) - f(x)\big)\log f(x)dx$.

Recall that, if $\mathbb{E}[X^{-2}] < \infty$, $k_n = cn^{2/5}$, $c > 0$, and $f'(x)$ satisfies the Lipschitz order α condition, then for x in a compact set $I \subset \mathbb{R}^+$,

$$n^{2/5}\bigg(\hat{f}_n^{Pois}(x) - f(x)\bigg) - \frac{1}{2c}f'(x) \xrightarrow{D} \mathcal{G},$$

where $\mathcal{G}$ is the Gaussian process with covariance function $\gamma_x^2 \delta_{xs}$, where $\gamma_x^2 = \frac{\mathbb{E}[X]}{2}\sqrt{\frac{c}{2\pi}}x^{-3/2}f(x)$, $\delta_{xs} = 0$ for $x \neq s$ and $\delta_{xs} = 1$ for $x = s$.

On the other hand, consider $Y(t)$ as a Gaussian process with mean $m(t)$ and covariance function $\gamma(t, t')$ and $g(t)$ as a non-random measurable function, then by rewriting the integral as a Riemann sum, it is obvious that

$$\int_0^t Y(s)g(s)ds = \lim_{\Delta s \to 0} \sum_0^t Y(s)g(s)\Delta s$$

follows a Gaussian distribution with mean $\int_0^t m(s)g(s)ds$ and variance $\int_0^t \int_0^t k(s, s')g(s)g(s')dsds'$. Thus, by letting $t \to \infty$ we get

$$\int_0^\infty Y(s)g(t)dt \to \mathcal{N}(\mu, \sigma^2),$$

where

$$\mu = \int_0^\infty m(s)g(s)ds, \qquad \text{and} \qquad \sigma^2 = \int_0^t \int_0^\infty k(s, s')g(s)g(s')dsds.$$

As a result, we obtain

$$n^{2/5}\bigg(\hat{H}_f^{Plugin-Pois} - H_f\bigg) = \int_0^\infty n^{2/5}\bigg(\hat{f}_n^{Pois}(x) - f(x)\bigg)\log f(x)dx + o(1) \xrightarrow{D} \mathcal{N},$$

where $\mathcal{N}$ is a Normal random variable with mean

$$\lim_{n\to\infty} \mathbb{E}\int_0^\infty n^{2/5}\bigg(\hat{f}_n^{Pois}(x) - f(x)\bigg)\log f(x)dx = \frac{1}{2c}\int_0^\infty f'(x)\log f(x)dx$$

and the variance

$$\lim_{n\to\infty} Var\left[\int_0^\infty n^{2/5}\bigg(\hat{f}_n^{Pois}(x) - f(x)\bigg)\log f(x)dx\right] = \frac{\mathbb{E}[X]}{2}\sqrt{\frac{c}{2\pi}}\int_0^\infty x^{-3/2}f(x)\big(\log f(x)\big)^2 dx.$$

For $\mathcal{N}$ to be well-defined, we need the following conditions

$$\int_0^\infty f'(x)\log f(x)dx < \infty \qquad \text{and} \qquad \int_0^\infty x^{-3/2}f(x)\big(\log f(x)\big)^2 dx < \infty,$$

as assumed in the theorem. This completes the proof of the theorem. □

Remark 2.1. The asymptotic results established here using the Poisson weight smoothing estimator of the density function follows very closely to those established in Hall and Morton (1993), though under some stringent smoothness conditions. Such results may also be established using alternative asymmetric kernel density estimators such as those proposed in Chaubey *et al.* (2012), Chen (2000), Cheng and Parzen (1997) and others. However, Poisson weights based entropy estimator, especially $\hat{H}_f^{Meanlog-Pois}$ may be computationally preferable over others. This is important as there may not exist an uniformly best estimator, as demonstrated through numerical studies in the next section.

3 Comparison of Estimators: A Simulation Study

In this section, we present a simulation study on our proposed estimators and some other existing entropy estimators which are given in the sub-section below.

3.1 Entropy estimators

3.1.1 Van Es entropy estimator (Van Es, 1992)

This entropy is motivated by an approach called "*spacing*", initiated by Vasicek (1976). By the change of variable $p = F(x)$, the entropy can be expressed in the form

$$H_f = -\int_{-\infty}^{\infty} f(x) \log f(x) dx = \int_0^1 \log \left\{ \frac{d}{dp} F^{-1}(p) \right\} dp.$$

To estimate H_f, the distribution F is replaced by the empirical distribution F_n, and the differential operator is replaced by the difference operator. As a result, the derivative of $F^{-1}(p)$ is estimated by $\frac{n}{2m}(X_{(i+m)} - X_{(i-m)})$ for $(i-1)/n < p \leq i/n$, $i = m+1, m+2, ..., n-m$, where $X_{(i)}$'s are the order statistics and m is a positive integer smaller than $n/2$. When $p \leq m/n$ or $p > (n-m)/n$, one-sided differences are used. That is, $(X_{(i+m)} - X_{(1)})$ and $(X_{(n)} - X_{(i-m)})$ are in place of $(X_{(i+m)} - X_{(i-m)})$ respectively. All together this leads to the following estimator of entropy

$$\hat{H}_f^{Vasicek} = \frac{1}{n} \sum_{i=1}^{n} \log \left\{ \frac{n}{2m} (X_{(i+m)} - X_{(i-m)}) \right\}. \tag{3.1}$$

Since $\hat{H}_f^{Vasicek}$ was shown to produce a large bias. Van Es (1992) modified and improved the performance of $\hat{H}_f^{Vasicek}$ by his estimator that is given as

$$\hat{H}_f^{VanEs} = -\frac{1}{n-m} \sum_{i=1}^{n-m} \log \left\{ \frac{n+1}{m} (X_{(i+m)} - X_{(i)}) \right\} + \sum_{j=m}^{n} \frac{1}{j} + \log(m) - \log(n+1). \tag{3.2}$$

Van Es (1992) established, under certain conditions, the strong consistency and asymptotic normality of $\hat{H}_f^{VanEs}$.

3.1.2 Entropy estimator by means of quantile density estimation

Bouzebda *et al.* (2013) presented an estimator of entropy based on smooth estimator of quantile density function. Their idea again starts with the expression of the entropy

$$H_f = \int_0^1 \log\left\{\frac{d}{dp}F^{-1}(p)\right\}dp = \int_0^1 \log\left\{\frac{d}{dp}Q(p)\right\}dp = \int_0^1 \log q(p)dp,$$

where $Q(p) := \inf\{t : F(t) \geq p\}$ for $0 \leq p \leq 1$ is the quantile function and $q(p) := dQ(p)/dp = 1/f(Q(p))$ is the quantile density function. Then a new entropy estimator can be obtained by substituting $q(\cdot)$ by its appropriate estimator $\hat{q}_n(\cdot)$.

$$\hat{H}_f = \int_0^1 \log \hat{q}_n(p)dp.$$

Bouzebda *et al.* (2013) were motivated by the work of Cheng and Parzen (1997), which introduced a kernel type estimator of $q(\cdot)$.

$$\tilde{q}_n^{CP}(p) := \frac{d}{dp}\tilde{Q}_n^{CP}(p) = \frac{d}{dp}\int_0^1 \hat{Q}_n(t)K_n(p,t)d\mu_n(t),$$

where $\hat{Q}_n(\cdot)$ is the empirical quantile function, $K_n(p,x)$ is the sequence density kernel functions defined on $(0,1)\times[0,1]$, and $\mu_n(x)$ is a sequence of σ-finite measure on $[0,1]$. In this paper, we use one special form of $\tilde{q}_n^{CP}(p)$, which coincides to the method of Bernstein polynomial. First, the quantile function $Q(\cdot)$ is estimated by Bernstein polynomial of degree m.

$$\tilde{Q}_n(p) = \sum_{i=0}^{m} \hat{Q}_n\left(\frac{i}{m}\right)b(i,m,p) \qquad p \in [0,1],$$

where $b(i,m,p) = P[Y = i]$, and Y follows the binomial(m,p) distribution, and m is a function of n such that $m \to \infty$ as $n \to \infty$. Then the estimator of $q(\cdot)$ can be obtained by differentiating $\tilde{Q}_n(\cdot)$

$$\tilde{q}_n(p) = \frac{d\tilde{Q}_n(p)}{dp} = \sum_{i=0}^{m} \hat{Q}_n\left(\frac{i}{m}\right)b(i,m,p)\left[\frac{i-mp}{p(1-p)}\right]. \tag{3.3}$$

Therefore, the entropy estimator by means of quantile density estimation is of the form

$$\hat{H}_f^{Quantile} = \int_0^1 \log \tilde{q}_n(p)dp. \tag{3.4}$$

Bouzebda *et al.* (2013) studied and obtained the strong asymptotic consistency and asymptotic normality for their estimator.

3.1.3 Entropy estimator by means of asymmetric kernel density estimation

Lastly, to compare the performance of our entropy estimators with the one using asymmetric kernel density estimation, we also run simulation on these two estimators using Gamma kernel estimators as given in Chen (2000).

$$\hat{H}_f^{Plugin-Gam} = -\int_0^{\infty} \hat{f}_n^{Gam}(x) \log \hat{f}_n^{Gam}(x) dx, \tag{3.5}$$

$$\hat{H}_f^{Meanlog-Gam} = -\sum_{i=1}^{n} \log \hat{f}_n^{Gam}(X_i), \tag{3.6}$$

where $\hat{f}_n^{Gam}(x)$ is the asymmetric kernel density estimator with the Gamma kernel defined as

$$\hat{f}_n^{Gam}(x) = \frac{1}{n}\sum_{i=1}^{n} K_{x/b+1,b}^{Gam}(X_i) = \sum_{i=1}^{n} \frac{X_i^{x/b}\exp(-X_i/b)}{b^{x/b+1}\Gamma(x/b+1)}. \tag{3.7}$$

3.2 Simulation results

The simulation is performed on selected densities which are

- Standard Exponential(1) with true $H_f = 1$.
- Uniform(0,1) with true $H_f = 0$.
- Weibull(2,2) with true $H_f \approx 1.2886$.
- Gamma(2,2) with true $H_f \approx 0.8841$.

The simulation study is organized as follows. For each density, 500 replicated data are generated for different sample sizes: 10, 20, 40, 80, 160, 320, 640 and 1000. To obtain the entropy estimators, we need to assign value to smoothing parameters. Particularly, the choice of m in Van Es estimators is set to $\lfloor \sqrt{n} + 0.5 \rfloor$; the bandwidth selection $b = sn^{-2/5}$ is used in $\hat{f}_n^{Gam}(\cdot)$ where s is the sample standard deviation; the choice of the polynomial degree m in the quantile density estimator $\tilde{q}(\cdot)$ is fixed to $m = n/\log n$; and lastly for our estimators, we applied the smoothing parameter $k = n^{2/5} + 1$. To compare the performance between estimators, for each density and each estimator, we compute the point estimate and its MSE shown in the parentheses. The simulation results are shown in the Table 1 to Table 4. Each row in the table corresponds to one sample size, and the bold value in that row indicates the best estimator with the smallest MSE for that sample size. Furthermore, for better visual comparison, we convert the tables into graphs in which we present the graphs of true entropy along with entropy estimator and the mean squared error (MSE) of each estimator. Note that, deal to the round-up in R software, the integrand $f(x)\log f(x)$ may produce non-finite value for large value of x, so this results in a problem when computing plugin entropy estimators. However, since the contribution of the right tail of the integrand to the entropy is insignificant for sufficiently large x, we obtain the approximately true value of entropy by cutting off the negligible right tail of the integration. That is, instead of integrating over the entire support of $f(x)\log f(x)$, we integrate up to a certain value at which the right tail is negligible. By observing the simulation results below, we have some important remarks.

Remark 3.1. *There does not exist the uniquely best entropy estimator in all cases.*

It is clear from the simulation results that, depending on the density and the sample size, the best entropy estimator switches from one to another. However, in most cases the MSE of $\hat{H}_f^{Meanlog-Pois}$ is the smallest comparing to that of others estimators from small to large sample size.

Exp(1) with H=1						
n	Van	Gam Plugin	Gam Meanlog	Pois Plugin	Pois Meanlog	Quantile
10	0.88692(0.15032)	1.03736(0.08266)	1.01024(0.11782)	**1.06104(0.06758)**	0.86768(0.09830)	0.72719(0.21404)
20	0.90369(0.06882)	1.06464(0.04512)	1.39239(0.27253)	**1.02245(0.02677)**	0.92819(0.04920)	0.94326(0.06576)
40	0.90975(0.03749)	1.06563(0.02532)	1.05875(0.03029)	**1.06993(0.02494)**	0.95946(0.02542)	1.01170(0.02919)
80	0.91354(0.02061)	1.05590(0.01353)	1.05148(0.01511)	1.05815(0.01344)	**0.97348(0.01217)**	1.02228(0.01406)
160	0.92048(0.01335)	1.04263(0.00772)	1.03973(0.00834)	1.04367(0.00773)	**0.97876(0.00691)**	1.01126(0.00719)
320	0.93601(0.00731)	1.03691(0.00424)	1.03528(0.00445)	1.03769(0.00421)	**0.98818(0.00317)**	1.02038(0.00380)
640	0.95099(0.00398)	1.03193(0.00248)	1.03116(0.00255)	1.03256(0.00250)	**0.99495(0.00156)**	1.01621(0.00190)
1000	0.95532(0.00303)	1.02577(0.00160)	1.02522(0.00164)	1.02626(0.00162)	**0.99469(0.00101)**	1.01174(0.00111)

Table 1: Simulation results for Exp(1). The point estimate and MSE (in parentheses) are computed for each estimator. The bold value is the best estimator with the smallest MSE.

Uniform(0,1) with H=0						
n	Van	Gam Plugin	Gam Meanlog	Pois Plugin	Pois Meanlog	Quantile
10	0.00769(0.04625)	**0.08813(0.01535)**	0.09916(0.03164)	0.14767(0.02580)	0.21800(0.06163)	-0.28010(0.10542)
20	-0.00516(0.01648)	**0.08265(0.01001)**	0.10628(0.01932)	0.13774(0.02146)	0.20724(0.05009)	-0.14574(0.02800)
40	**0.00129(0.00449)**	0.07954(0.00790)	0.11189(0.01562)	0.12830(0.01787)	0.19737(0.04241)	-0.07994(0.00824)
80	**0.00013(0.00167)**	0.07288(0.00600)	0.10478(0.01222)	0.11711(0.01438)	0.18013(0.03388)	-0.04452(0.00249)
160	**-0.00043(0.00060)**	0.06433(0.00444)	0.09414(0.00938)	0.10561(0.01146)	0.16206(0.02688)	-0.02613(0.00082)
320	**0.00004(0.00018)**	0.05740(0.00343)	0.08470(0.00739)	0.09403(0.00900)	0.14371(0.02093)	-0.01567(0.00029)
640	**0.00005(0.00006)**	0.05085(0.00265)	0.07496(0.00571)	0.08474(0.00726)	0.12859(0.01666)	-0.00942(0.00010)
1000	**0.00016(0.00003)**	0.04637(0.00219)	0.06858(0.00475)	0.23047(0.05322)	0.11770(0.01393)	-0.00703(0.00005)

Table 2: Simulation results for Uniform(0,1). The point estimate and MSE (in parentheses) are computed for each estimator. The bold value is the best estimator with the smallest MSE.

Weibull(2,2) with H=1.288608						
n	Van	Gam Plugin	Gam Meanlog	Pois Plugin	Pois Meanlog	Quantile
10	1.10575(0.11478)	1.41383(0.04666)	1.26877(0.04891)	1.42878(0.03868)	**1.21830(0.03531)**	0.91140(0.21121)
20	1.11085(0.06779)	1.40301(0.02891)	1.27566(0.02284)	1.41451(0.02720)	**1.24453(0.01914)**	1.11519(0.05976)
40	1.13968(0.03680)	1.39623(0.01956)	1.28998(0.01086)	1.40614(0.02015)	**1.27008(0.00946)**	1.21977(0.01871)
80	1.15698(0.02440)	1.38271(0.01322)	1.29238(0.00578)	1.39027(0.01396)	**1.28011(0.00517)**	1.26866(0.00742)
160	1.17739(0.01546)	1.36726(0.00829)	1.29183(0.00263)	1.37379(0.00905)	**1.28478(0.00243)**	1.28819(0.00306)
320	1.20003(0.00934)	1.35354(0.00536)	1.29212(0.00136)	1.35890(0.00595)	**1.28808(0.00128)**	1.29801(0.00155)
640	1.21771(0.00571)	1.34003(0.00319)	1.29061(0.00063)	1.34422(0.00359)	**1.28822(0.00061)**	1.29667(0.00073)
1000	1.22664(0.00428)	1.3323(0.00236)	1.29050(0.00042)	1.33689(0.00267)	**1.28882(0.00040)**	1.29654(0.00050)

Table 3: Simulation results for Weibull(2,2). The point estimate and MSE (in parentheses) are computed for each estimator. The bold value is the best estimator with the smallest MSE.

Gamma(2,2) with H=0.884068						
n	Van	Gam Plugin	Gam Meanlog	Pois Plugin	Pois Meanlog	Quantile
10	0.68667(0.13542)	0.51642(0.31349)	0.84063(0.07423)	1.02446(0.05316)	**0.81202(0.05043)**	0.50401(0.23319)
20	0.72870(0.07030)	0.44159(0.35513)	0.88528(0.03654)	1.02871(0.04049)	**0.85915(0.02737)**	0.75141(0.06294)
40	0.74058(0.04216)	0.95076(0.01776)	0.89397(0.01803)	1.01353(0.02765)	**0.87418(0.01456)**	0.84924(0.02292)
80	0.75324(0.02707)	0.94141(0.01022)	0.89129(0.00876)	0.99356(0.01783)	**0.87847(0.00755)**	0.88047(0.00975)
160	0.77418(0.01658)	0.93428(0.00578)	0.89224(0.00413)	0.97651(0.01147)	**0.88345(0.00367)**	0.89297(0.00450)
320	0.79502(0.01026)	0.92585(0.00349)	0.89142(0.00218)	0.96047(0.00745)	**0.88550(0.00198)**	0.90258(0.00260)
640	0.81324(0.00617)	0.91769(0.00206)	0.88939(0.00114)	0.94522(0.00463)	**0.88546(0.00106)**	0.89770(0.00134)
1000	0.82046(0.00484)	0.91201(0.00142)	0.88725(0.00076)	0.93605(0.00332)	**0.88435(0.00072)**	0.89501(0.00087)

Table 4: Simulation results for Gamma(2,2). The point estimate and MSE (in parentheses) are computed for each estimator. The bold value is the best estimator with the smallest MSE.

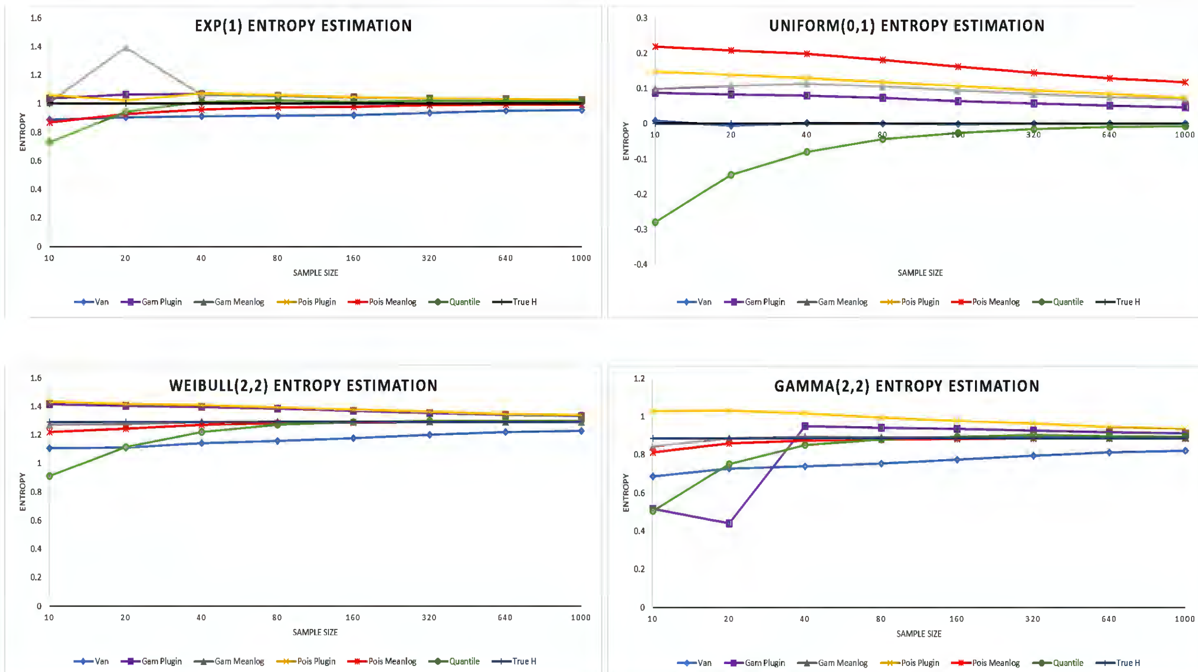

Figure 1: Plot of the true entropy (black) and entropy estimators: Van Es estimator (blue), Gamma plugin (purple), Gamma mean of log (grey), Poisson plug-in (orange), Poisson mean of log (red), quantile (green) for different sample sizes.

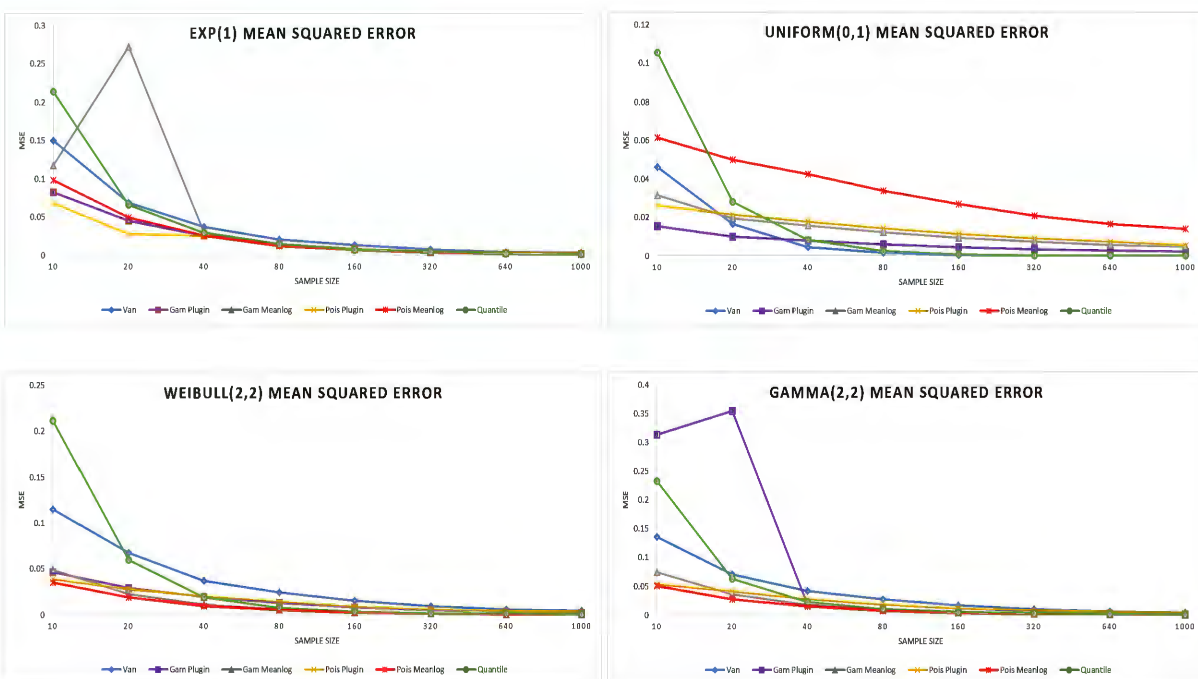

Figure 2: Plot of MSE of entropy estimators: Van Es estimator (blue), Gamma plugin (purple), Gamma mean of log (grey), Poisson plug-in (orange), Poisson mean of log (red), quantile (green) for different sample sizes.

Remark 3.2. *In The case of Uniform(0,1), The Van Es estimator outperforms all other estimators for sufficiently large sample size.*

The Uniform(0,1) is also the only case in which our estimators have a poor performance while the Van Es estimator gets very close to the true one. This makes sense because the Van Es estimator is built on the principle of approximating differential operator by a difference operator of distribution function, and the uniform distribution has a constant rate on its support.

Remark 3.3. *The rate of convergence of the estimator $\hat{H}_f^{Quantile}$ seems to be faster than other estimators.*

Although for a small sample size, $\hat{H}_f^{Quantile}$ estimator does not show up to be a good estimator, and it is not recommended. As the sample size increases, it becomes one of the best potential choices for entropy estimator due to its consistently fast rate of convergence to the true entropy.

References

[1] Ahmad, I.A. and Lin, P.E. (1989). A nonparametric estimation of the entropy for absolutely continuous distributions. *IEEE Trans. Inform. Theory*, **36,** 688–692.

[2] Bagai, I. and Prakasa Rao, B.L.S. (1996). Kernel type density estimates for positive valued random variables. *Sankhya Ser. A*, **57**, 56–67.

[3] Beirlant, J., Dudewicz, E.J., Györfi, L.,van de Meulen, E.C. (1997). Nonparametric entropy estimation. An overview. *International Journal of Mathematical and Statistical Sciences*, **6**, 17–39.

[4] Bouezmarni, T. and Scaillet, O. (2005). Consistency of asymmetric kernel density estimators and smoothed histograms with application to income data. *Econometric Theory*, **21**, 390–412.

[5] Bouzebda, S., Elhattab, I., Keziou, A., and Lounis, T. (2013). New entropy estimator with an application to test of normality. *Communication in Statistics-Theory and Methods*, **42**, 2245–2270.

[6] Chaubey, Y.P., Li, J., Sen, A. and Sen, P.K. (2012). A new smooth density estimator for non-negative random variables. *Journal of the Indian Statistical Association*, **50**, 83–104.

[7] Chaubey, Y.P. and Mudholkar, G.S. (2013). A rationale for maximizing likelihood and related alternatives to maximum likelihood estimator. *Investigations in Mathematical Sciences*, **3**, 1–15.

[8] Chaubey, Y.P. and Sen, P.K. (1996). On smooth estimation of survival and density functions. *Statistics and Decisions*, **14**, 1–22.

[9] Chaubey, Y.P. and Sen, P.K. (2009). On the selection of the smoothing parameter in Poisson smoothing of histogram estimator: Computational aspects. *Pakistan Journal of Statistics*, **25**, 385–401.

[10] Chaubey, Y.P., Sen, P.K. and Li, J. (2010). Smooth density estimation for length-biased data. *Journal of The Indian Society of Agricultural Statistics*, **64**, 145–155.

[11] Chen, S.X. (2000). Probability Density Function Estimation Using Gamma Kernels. *Annals of the Institute of Statistical Mathematics*, **52**, 471–480.

[12] Cheng, C. and Parzen, E. (1997). Unified estimators of smooth quantile and quantile density funnctions. *Journal of Statistical Planning and Inference*, **59**, 291–307.

[13] Cover, T.M. and Thomas, J.A. (1991). *Elements of Information Theory*, Wiley, New York.

[14] Eggermont, P.P.B. and LaRiccia, V.N. (1999). Best asymptotic normality of the kernel density entropy estimator for smooth densities. *IEEE Transactions On Information Theory*, **45**, 1321–1324.

[15] Györfi, L. and van der Meulen, E.C. (1987). Density-free convergence properties of various estimators of the entropy. *Comput. Statist. Data Anal.*, **5**, 425–436.

[16] Györfi, L. and van der Meulen, E.C. (1990). An entropy estimate based on a kernel density estimation, *Colloq. Math. Soc. J. Bolyai* no. 57. In *Limit Theorems in Probability and Statistics,* I. Berkes, E. Csáki and P. Révész, Eds., North-Holland: Amsterdam, The Netherlands.

[17] Hall, P. and Morton, S.C. (1993). On the estimation of entropy. *Ann. Inst. Statist. Math.*, **45**, 69–88.

[18] Jaynes, E.T. (1957a). Information theory and statistical mechanics-I. *The Physical Review,* **106**, 620–630.

[19] Jaynes, E.T. (1957b). Information theory and statistical mechanics-II. *The Physical Review,* **108**, 171–190.

[20] Karunamuni, R.J. and Alberts, T. (2005). On boundary correction in kernel density estimation. *Statistical Methodology*, **2**, 191–212.

[21] Scaillet, O. (2004). Density estimation using inverse and reciprocal inverse Gaussian kernels. *Nonparametric Statistics*, **16**, 217–226.

[22] Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Tech. J.*, **27**, 379-423. Reprinted in C. E . Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana.

[23] Shwartz, S., Zibulevsky, M. and Schechner, Y.Y. (2005). Fast kernel entropy estimation and optimization. *Signal Processing*, **85**, 1045–1058.

[24] Van Es, B. (1992). Estimating functionals related to a density by a class of statistics based on spacings. *Scandinavian Journal of Statistics*, **19**, 61–72.

[25] Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society*, **38**, 54–59.

List of Recent Technical Reports

90. Md. Sharif Mozumder and José Garrido, *On the Relation between the Lévy Measure and the Jump Function of a Lévy Process*, October 2007

91. Arusharka Sen and Winfried Stute, *A Bi-Variate Kaplan-Meier Estimator via an Integral Equation*, October 2007

92. C. Sangüesa, *Uniform Error Bounds in Continuous Approximations of Nonnegative Random Variables Using Laplace Transforms*, January 2008

93. Yogendra P. Chaubey, Naâmane Laib and Arusharka Sen, *A Smooth Estimator of Regression Function for Non-negative Dependent Random Variables*, March 2008

94. Alejandro Balbás, Beatriz Balbás and Antonio Heras, *Optimal Reinsurance with General Risk Functions*, March 2008

95. Alejandro Balbás, Raquel Balbás and José Garrido, *Extending Pricing Rules with General Risk Functions*, April 2008

96. Yogendra P. Chaubey and Pranab K. Sen, *On the Selection of the Smoothing Parameter in Poisson Smoothing of Histogram Estimator: Computational Aspects*, December 2008

97. Runhuan Feng and José Garrido, *Actuarial Applications of Epidemiological Models*, December 2008

98. Alejandro Balbás, Beatriz Balbás and Raquel Balbás, *CAPM and APT Like Models with Risk Measures*, June 2009

99. Jun Zhou and José Garrido, *A Loss Reserving Method Based on Generalized Linear Models*, June 2009

100. Yogendra P. Chaubey and Isha Dewan, *Smooth Estimation of Survival and Density Functions for a Stationary Associated Process Using Poisson Weights*, September 2009

101. Krishna K. Saha and Debaraj Sen, *Improved Confidence Interval for the Dispersion Parameter in Count Data Using Profile Likelihood*, September 2009

102. Araceli Reyes, *Difficulties Teaching Mathematics with the ACE Cycle*, September 2009

103. Mansi Khurana, Yogendra P. Chaubey and Shalini Chandra, *Jackknifing the Ridge Regression Estimator: A Revisit*, February 2012.

104. Yogendra P. Chaubey and Rui Zhang, *Survival Distributions with Bathtub Shaped Hazard: A New Distribution Family*, November 2013.

105. Yogendra P. Chaubey, *Smooth Kernel Estimation of a Circular Density Function: A Connection to Orthogonal Polynomials on the Unit Circle*, January 2016.

106. Nasim Ejlali, Hamid Pezeshk, Yogendra P. Chaubey and Mehdi Sadeghi, *Parrondo's Paradox for Games with Three Players*, April 2018.

107. Yogendra P. Chaubey and Shamal C. Karmaker, *On Some Circular Distributions Induced by Inverse Stereographic Projection*, April 2018.

108. Patrice Gaillardetz and Saeb Hachem, *Risk-control Strategies*, August 2019.

109. Yogendra P. Chaubey and Nhat Linh Vu, *A Numerical Study of Entropy and Residual Entropy Estimators Based on Smooth Density Estimators for Non-negative Random Variables*, March 2020.

110. Yogendra P. Chaubey and Nhat Linh Vu, *On an Entropy Estimator Based on a Non-parametric Density Estimator for Non-negative Data*, April 2020.

Copies of technical reports can be requested from:

Dr. Wei Sun
Department of Mathematics and Statistics
Concordia University
1455 de Maisonneuve Blvd. West,
Montreal, QC, H3G 1M8 CANADA