© Statistical Research and Training Center

مجلهٔ پژوهش‌های آماری ایران

J. Statist. Res. Iran **12** (2015): 57–70

دوره‌ی ۱۲، شماره‌ی ۱، بهار و تابستان ۱۳۹۴، صص ۵۷–۷۰

پژوهشکده‌ی آمار

On the Estimation of Shannon Entropy

Hadi Alizadeh Noughabi

University of Birjand

Abstract. Shannon entropy is increasingly used in many applications. In this article, an estimator of the entropy of a continuous random variable is proposed. Consistency and scale invariance of variance and mean squared error of the proposed estimator is proved and then comparisons are made with Vasicek's (1976), van Es (1992), Ebrahimi et al. (1994) and Correa (1995) entropy estimators. A simulation study is performed and the results indicate that the proposed estimator has smaller mean squared error than competing estimators.

Keywords. Information theory; entropy estimator; exponential distribution; normal distribution; uniform distribution.

MSC 2010: 94A17, 94A24, 62B10, 62F10.

1 Introduction

Suppose that a random variable X has a distribution function $F(x)$ with a continuous density function $f(x)$. The entropy $H(f)$ of the random variable X is defined by Shannon (1948) to be

$$H(f) = -\int_{-\infty}^{\infty} f(x)\,\log f(x)\,dx.$$

Vasicek (1976) showed that

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

and then by replacing the distribution function F by the empirical distribution function F_n , and using a difference operator instead of the differential operator, an estimator constructed. The derivative of $F^{-1}(p)$ is then estimated by a function of the order statistics. If $X_1, \ldots, X_n$ is a random sample, then the Vasicek estimator is given by

$$HV_{mn} = \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{n}{2m}\{X_{(i+m)} - X_{(i-m)}\}\right],$$

where the window size m is a positive integer smaller than $n/2$, $X_{(i)} = X_{(1)}$ if $i < 1$, $X_{(i)} = X_{(n)}$ if $i > n$ and $X_{(1)} \leqslant X_{(2)} \leqslant \ldots \leqslant X_{(n)}$ are order statistics based on a random sample of size n. Vasicek proved that $HV_{mn} \overset{pr.}{\to} H(f)$ as $n \to \infty$, $m \to \infty$, $\frac{m}{n} \to 0$.

Van Es (1992) based on spacings proposed another estimator of entropy and proved its consistency. Van Es'estimator is given by

$$HVE_{mn} = \frac{1}{n-m}\sum_{i=1}^{n-m}\left[\frac{n+1}{m}\{X_{(i+m)} - X_{(i)}\}\right] + \sum_{k=m}^{n}\frac{1}{k} + \log(m) - \log(n+1) .$$

Ebrahimi et al. (1994) modified the Vasicek's estimator and proposed their estimator as

$$HE_{mn} = \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{n}{c_i m}\{X_{(i+m)} - X_{(i-m)}\}\right],$$

where

$$c_i = \begin{cases} 1 + \frac{i-1}{m}, & 1 \leqslant i \leqslant m, \\ 2, & m+1 \leqslant i \leqslant n-m, \\ 1 + \frac{n-i}{m}, & n-m+1 \leqslant i \leqslant n. \end{cases}$$

They proved that $HE_{mn} \overset{pr.}{\to} H(f)$ as $n \to \infty$, $m \to \infty$, $\frac{m}{n} \to 0$.

Correa (1995) proposed a modification of Vasicek estimator which has a smaller mean square error (MSE) based on a local linear model. The estimator of entropy proposed by Correa is given by

$$HC_{mn} = -\frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{\sum_{j=i-m}^{i+m}\{X_{(j)} - \bar{X}_{(i)}\}(j-i)}{n\sum_{j=i-m}^{i+m}\{X_{(j)} - \bar{X}_{(i)}\}^2}\right],$$

where

$$\bar{X}_{(i)} = \frac{1}{2m+1} \sum_{j=i-m}^{i+m} X_{(j)}\,.$$

He showed that his estimator has smaller MSE than Vasicek entropy estimator. Also, for some of m his estimator has smaller MSE than van Es estimator.

Because

1) Entropy is a useful measure of uncertainty and dispersion and has been widely employed in many pattern analysis applications;

2) Many researchers have used the estimators of entropy for developing entropy based statistical procedure. See, for example, Esteban et al. (2001), Park and Park (2003), Choi et al. (2004), Goria et al. (2005), Choi (2008), Vexler et al. (2011), and Alizadeh Noughabi and Arghami (2011, 2012, 2013);

3) In nonparametric statistics, we often have a random sample of an unknown population and inferences are based on the observations;
it will be of interest to introduce an entropy estimator based on observations of an unknown population. Our goal in this paper is to introduce an entropy estimator with a better performance than the existing estimators.

In Section 2, a new estimator of entropy is introduced and its consistency is proved. Scale invariance of variance and mean squared error of the proposed estimator is established. In Section 3, results of a comparison of our estimator with the competing estimators by a simulation study are given.

2 The New Estimator

It is clear that

$$S_i(m,n) = \frac{\frac{2m}{n}}{X_{(i+m)} - X_{(i-m)}}\,, \qquad (1)$$

is not a correct formula for the slope when $i \leqslant m$ or $i \geqslant n-m+1$. Therefore, in order to correctly estimate the slopes at these points the denominator and/or the numerator should be modified.

We propose an estimator by modifying the numerator of (1). In fact when $i \leqslant m$ the denominator of (1) is $X_{(i+m)} - X_{(i-m)} = X_{(i+m)} - X_{(1)}$ and this distance is less than the actual distance. For example, let $n = 10$ and $m = 3$, we have

$$\begin{array}{ll}
i=1 & X_{(1+3)} - X_{(1-3)} = X_{(4)} - X_{(-2)} \geqslant X_{(4)} - X_{(1)} \\
i=2 & X_{(2+3)} - X_{(2-3)} = X_{(5)} - X_{(-1)} \geqslant X_{(4)} - X_{(1)} \\
i=3 & X_{(3+3)} - X_{(3-3)} = X_{(6)} - X_{(0)} \geqslant X_{(4)} - X_{(1)} \\
i=4 & X_{(4+3)} - X_{(4-3)} = X_{(7)} - X_{(1)} = X_{(7)} - X_{(1)} \\
i=5 & X_{(5+3)} - X_{(5-3)} = X_{(8)} - X_{(2)} = X_{(8)} - X_{(2)} \\
i=6 & X_{(6+3)} - X_{(6-3)} = X_{(9)} - X_{(3)} = X_{(9)} - X_{(3)} \\
i=7 & X_{(7+3)} - X_{(7-3)} = X_{(10)} - X_{(4)} = X_{(10)} - X_{(4)} \\
i=8 & X_{(8+3)} - X_{(8-3)} = X_{(11)} - X_{(5)} \geqslant X_{(10)} - X_{(5)} \\
i=9 & X_{(9+3)} - X_{(9-3)} = X_{(12)} - X_{(6)} \geqslant X_{(10)} - X_{(6)} \\
i=10 & X_{(10+3)} - X_{(10-3)} = X_{(12)} - X_{(7)} \geqslant X_{(10)} - X_{(7)}
\end{array}$$

Toward this end, in numerator of (1) a value from $\frac{2m}{n}$ is subtracted.

Ebrahimi et al. (1994) modified the numerator of (1) as

$$\begin{cases}
\frac{2m}{n} - \frac{1+m-i}{n} = \frac{2m}{n} - \frac{1-(i-m)}{n} = \frac{m+i-1}{n}, & 1 \leqslant i \leqslant m, \\
\frac{2m}{n}, & m+1 \leqslant i \leqslant n-m, \\
\frac{2m}{n} - \frac{i+m-n}{n} = \frac{2m}{n} - \frac{(i+m)-n}{n} = \frac{n+m-i}{n}, & n-m+1 \leqslant i \leqslant n,
\end{cases}$$

Again, consider the above example, we have

i	Denominator	Numerator
$i=1$	$X_{(4)} - X_{(-2)}$	$\frac{2m}{n} - \frac{3}{n}$
$i=2$	$X_{(5)} - X_{(-1)}$	$\frac{2m}{n} - \frac{2}{n}$
$i=3$	$X_{(6)} - X_{(0)}$	$\frac{2m}{n} - \frac{1}{n}$
$i=4$	$X_{(7)} - X_{(1)}$	$\frac{2m}{n}$
$i=5$	$X_{(8)} - X_{(2)}$	$\frac{2m}{n}$
$i=6$	$X_{(9)} - X_{(3)}$	$\frac{2m}{n}$
$i=7$	$X_{(10)} - X_{(4)}$	$\frac{2m}{n}$
$i=8$	$X_{(11)} - X_{(5)}$	$\frac{2m}{n} - \frac{1}{n}$
$i=9$	$X_{(12)} - X_{(6)}$	$\frac{2m}{n} - \frac{2}{n}$
$i=10$	$X_{(13)} - X_{(7)}$	$\frac{2m}{n} - \frac{3}{n}$

Based on the simulations obtained by Ebrahimi et al. (1994), we can see their estimator has smaller root mean square error (MSE) than Vasicek estimator. Here, we improve this estimator and then show the proposed estimator has a good MSE respect to Vasicek and Ebrahimi et al. estimators.

We modify the numerator of (1) as

$$\begin{cases} \frac{2m}{n} - \frac{3}{2}\frac{1+m-i}{n} = \frac{2m}{n} - \frac{3-3(i-m)}{2n} = \frac{m+3(i-1)}{2n}, & 1 \leqslant i \leqslant m, \\ \frac{2m}{n}, & m+1 \leqslant i \leqslant n-m, \\ \frac{2m}{n} - \frac{3}{2}\frac{i+m-n}{n} = \frac{2m}{n} - \frac{3(i+m)-3n}{2n} = \frac{m+3(n-i)}{2n}, & n-m+1 \leqslant i \leqslant n, \end{cases}$$

Consider the above example, we have

i	Denominator	Numerator
$i=1$	$X_{(4)} - X_{(-2)}$	$\frac{2m}{n} - \frac{9}{2n}$
$i=2$	$X_{(5)} - X_{(-1)}$	$\frac{2m}{n} - \frac{6}{2n}$
$i=3$	$X_{(6)} - X_{(0)}$	$\frac{2m}{n} - \frac{3}{2n}$
$i=4$	$X_{(7)} - X_{(1)}$	$\frac{2m}{n}$
$i=5$	$X_{(8)} - X_{(2)}$	$\frac{2m}{n}$
$i=6$	$X_{(9)} - X_{(3)}$	$\frac{2m}{n}$
$i=7$	$X_{(10)} - X_{(4)}$	$\frac{2m}{n}$
$i=8$	$X_{(11)} - X_{(5)}$	$\frac{2m}{n} - \frac{3}{2n}$
$i=9$	$X_{(12)} - X_{(6)}$	$\frac{2m}{n} - \frac{6}{2n}$
$i=10$	$X_{(13)} - X_{(7)}$	$\frac{2m}{n} - \frac{9}{2n}$

Therefore, we propose entropy estimator of Shannon entropy $H(f)$ of an unknown continuous probability density function by

$$HN_{mn} = \frac{1}{n}\sum_{i=1}^{n} \log\left[\frac{n}{a_i m}\{X_{(i+m)} - X_{(i-m)}\}\right],$$

where

$$a_i = \begin{cases} \frac{m+3(i-1)}{2m}, & 1 \leqslant i \leqslant m, \\ 2, & m+1 \leqslant i \leqslant n-m, \\ \frac{m+3(n-i)}{2m}, & n-m+1 \leqslant i \leqslant n. \end{cases}$$

and $X_{(i-m)} = X_{(1)}$ for $i \leqslant m$ and $X_{(i+m)} = X_{(n)}$ for $i \geqslant n-m$.

The reason that we propose the above estimator is as follows. We can see from the simulation study presented in Ebrahimi et al. (1994) that the estimated value of entropy by Vasicek (1976)'s and Ebrahimi et al. (1994)'s estimators are less than the value of entropy of the considered population. For example, let the distribution of the population is the standard exponential. Therefore, $H(f) = 1$. Now, if we estimate the entropy by the observations generated from the exponential distribution (we repeat this experiment several times), we can see that almost all the estimated values of entropy are less than one. This show that these estimators are underestimate. Therefore, we use the coefficient 3/2 so that this problem can be solved. Actually, we modify the numerator so that the estimator produce larger values than the values of the existing estimators. We will show, in Theorem 2, that $HN_{mn} > HV_{mn}$ and $HN_{mn} > HE_{mn}$. This show that the problem of underestimate can be solved by the proposed estimator.

***Theorem* 1.** *Let $X_1, \ldots, X_n$ be a random sample from distribution $F(x)$. Then*

$$HN_{mn} = HV_{mn} + \frac{2}{n}\left[m\log(4m) - \sum_{i=1}^{m}\log\{m + 3(i-1)\}\right],$$

and

$$HN_{mn} = HE_{mn} + \frac{2}{n}\left[m\log(2) - \log\frac{(m-1)!}{(2m-1)!} - \sum_{i=1}^{m}\log\{m + 3(i-1)\}\right],$$

where HV_{mn} and HE_{mn} are Vasicek (1976) and Ebrahimi et al. (1994) estimators, respectively.

Proof. By comparing the proposed estimator with Vasicek estimator, we obtain

$$
\begin{aligned}
HN_{mn} &= \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{n}{a_i m}\{X_{(i+m)}-X_{(i-m)}\}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{2n}{2a_i m}\{X_{(i+m)}-X_{(i-m)}\}\right] \\
&= HV_{mn}+\frac{1}{n}\sum_{i=1}^{n}\log\frac{2}{a_i} \\
&= HV_{mn}+\frac{1}{n}\left\{\sum_{i=1}^{m}\log\frac{4m}{m+3(i-1)}+\sum_{i=n-m+1}^{n}\log\frac{4m}{m+3(n-i)}\right\} \\
&= HV_{mn}+\frac{2}{n}\left[m\log(4m)-\sum_{i=1}^{m}\log\{m+3(i-1)\}\right]. \qquad (2)
\end{aligned}
$$

Also, from Ebrahimi et al. (1994), we have

$$
HE_{mn} = HV_{mn}+\frac{2}{n}\left\{m\log(2m)+\log\frac{(m-1)!}{(2m-1)!}\right\}. \qquad (3)
$$

Therefore, we obtain from (2) and (3)

$$
HN_{mn} = HE_{mn}+\frac{2}{n}\left[m\log(2)-\log\frac{(m-1)!}{(2m-1)!}-\sum_{i=1}^{m}\log\{m+3(i-1)\}\right].
$$

□

Remark 1. Theil (1980) computed the entropy $H(f_n^{ME})$ of an empirical maximum entropy density f_n^{ME}, which is related to HV_{1n}, HE_{1n} and HA_{1n} as follows.

$$
\begin{aligned}
H(f_n^{ME}) &= HV_{1n}+\frac{2-2\log 2}{n} \\
&= HE_{1n}+\frac{2-4\log 2}{n} \\
&= HN_{1n}+\frac{2-6\log 2}{n}.
\end{aligned}
$$

Here, f_n^{ME} is

$$f_n^{ME}(x) = \begin{cases} n^{-1}\frac{4}{x_{(2)}-x_{(1)}}\exp\left[\frac{x-(\frac{1}{2})\{x_{(1)}+x_{(2)}\}}{(\frac{1}{4})\{x_{(2)}-x_{(1)}\}}\right] & \text{if } x \leqslant \frac{1}{2}\{x_{(1)}+x_{(2)}\}, \\ n^{-1}\frac{2}{x_{(i+1)}-x_{(i-1)}} & \text{if } \frac{1}{2}\{x_{(i-1)}+x_{(i)}\} < x \leqslant \frac{1}{2}\{x_{(i)}+x_{(i+1)}\}, \\ & i=2,\ldots,n \\ n^{-1}\frac{4}{x_{(n)}-x_{(n-1)}}\exp\left[-\frac{x-(\frac{1}{2})\{x_{(n-1)}+x_{(n)}\}}{(\frac{1}{4})\{x_{(n)}-x_{(n-1)}\}}\right] & \text{if } x > \frac{1}{2}\{x_{(n-1)}+x_{(n)}\}. \end{cases}$$

Theorem 2. *Let $X_1, \ldots, X_n$ be a random sample from distribution $F(x)$. Then*

$$\begin{aligned} &i) \quad HN_{mn} > HV_{mn} \\ &ii) \quad HN_{mn} > HE_{mn} \end{aligned}$$

Proof. i). From (2), we have

$$HN_{mn} = HV_{mn} + \frac{2}{n}\left[m\log(4m) - \sum_{i=1}^{m}\log\{m+3(i-1)\}\right],$$

then it is enough to establish

$$\frac{2}{n}\left[m\log(4m) - \sum_{i=1}^{m}\log\{m+3(i-1)\}\right] > 0\,.$$

We can write

$$\begin{aligned} m\log(4m) - \sum_{i=1}^{m}\log\{m+3(i-1)\} &= \sum_{i=1}^{m}[\log(4m) - \log\{m+3(i-1)\}] \\ &= \sum_{i=1}^{m}\left[\log\left\{\frac{4m}{m+3(i-1)}\right\}\right]. \end{aligned}$$

Since

$$4m > \ m+3(i-1) \qquad \forall i = 1, 2, \ldots, m,$$

then (i) holds.

ii). From (2) and (3), it is enough to show

$$\frac{2}{n}\left[m\log(4m)-\sum_{i=1}^{m}\log\{m+3(i-1)\}\right]>\frac{2}{n}\left\{m\log(2m)+\log\frac{(m-1)!}{(2m-1)!}\right\},$$

or equivalently

$$\begin{aligned} & m\log(2)+\sum_{i=1}^{m}\log\left\{\frac{m+i-1}{m+3(i-1)}\right\}>0 \\ \Longleftrightarrow & \sum_{i=1}^{m}\log\left\{\frac{2(m+i-1)}{m+3(i-1)}\right\}>0\,. \end{aligned}$$

Since $m>\ (i-1),\quad \forall i=1,2,\ldots,m$, then (ii) holds. □

The next theorem states that the scale of the random variable X has no effect on the accuracy of HN_{mn} in estimating $H(f)$. Similar results have been obtained for HV_{mn} and HE_{mn} by Mack (1988) and Ebrahimi et al. (1994), respectively.

Theorem 3. *Let $X_1,\ldots,X_n$ be a sequence of i.i.d. random variables with entropy $H^X(f)$ and let $Y_i=kX_i, i=1,\ldots,n$, where $k>0$. Let HN^X_{mn} and HN^Y_{mn} be entropy estimators for $H^X(f)$ and $H^Y(g)$ respectively. (here g is pdf of $Y=kX$). Then the following properties hold.*

$$\begin{aligned} i)&\ \ \mathrm{E}\left(HN^Y_{mn}\right)=\mathrm{E}\left(HN^X_{mn}\right)+\log k, \\ ii)&\ \ \mathrm{Var}\left(HN^Y_{mn}\right)=\mathrm{Var}\left(HN^X_{mn}\right), \\ iii)&\ \ \mathrm{MSE}\left(HN^Y_{mn}\right)=\mathrm{MSE}\left(HN^X_{mn}\right). \end{aligned}$$

Proof. Since

$$HV^{kX}_{mn}=HV^X_{mn}+\log(k),$$

then from (2), we have

$$\begin{aligned} E(HN^{kX}_{mn}) &= E(HV^{kX}_{mn})+\frac{2}{n}\left[m\log(4m)-\sum_{i=1}^{m}\log\{m+3(i-1)\}\right] \\ &= E(HV^X_{mn})+\log(k)+\frac{2}{n}\left[m\log(4m)-\sum_{i=1}^{m}\log\{m+3(i-1)\}\right] \\ &= E(HN^X_{mn})+\log(k). \end{aligned}$$

Moreover,

$$\mathrm{Var}(HN_{mn}^{kX}) = \mathrm{Var}(HV_{mn}^{kX}) = \mathrm{Var}(HV_{mn}^{X}) = \mathrm{Var}(HN_{mn}^{X}),$$

and

$$\begin{aligned}\mathrm{MSE}(HN_{mn}^{kX}) &= \mathrm{Var}(HN_{mn}^{kX}) + \left\{\mathrm{E}(HN_{mn}^{kX}) - H^{KX}(g)\right\}^2\\ &= \mathrm{Var}(HN_{mn}^{X}) + \left\{\mathrm{E}(HN_{mn}^{X}) + \log(k) - H^{X}(f) - \log(k)\right\}^2\\ &= \mathrm{Var}(HN_{mn}^{X}) + \left\{\mathrm{E}(HN_{mn}^{X}) - H^{X}(f)\right\}^2\\ &= \mathrm{MSE}(HN_{mn}^{X}).\end{aligned}$$

Therefore, the proof of this theorem is complete. □

The following theorem establishes the consistency of HN_{mn}.

Theorem 4. *Let C be the class of continuous densities with finite entropies and let $X_1, \ldots, X_n$ be a random sample from $f \in C$. If $n \to \infty$, $m \to \infty$ and $\frac{m}{n} \to 0$, then*

$$HN_{mn} \longrightarrow H(f).$$

Proof. We have

$$HN_{mn} = HV_{mn} + \frac{2}{n}\left[m\log(4m) - \sum_{i=1}^{m}\log\{m+3(i-1)\}\right],$$

and

$$HV_{mn} \longrightarrow H(f),$$

(Vasicek, 1976). Moreover,

$$\frac{2}{n}\left[m\log(4m) - \sum_{i=1}^{m}\log\{m+3(i-1)\}\right] = \frac{2}{n}\sum_{i=1}^{m}\log\left\{\frac{4m}{m+3(i-1)}\right\},$$

and

$$0 \leqslant \frac{2}{n}\sum_{i=1}^{m}\log\left\{\frac{4m}{m+3(i-1)}\right\} \leqslant \frac{m}{n}\{2\log(4)\}.$$

Since the terms 0 and $\frac{m}{n}(2\log(4))$ go to zero if $n \to \infty$, $m \to \infty$ and $\frac{m}{n} \to 0$, the term

$$\frac{2}{n}\left[m\log(4m) - \sum_{i=1}^{m}\log\{m+3(i-1)\}\right]$$

goes to zero and consequently the theorem is hold. □

3 Simulation Study

A simulation study was performed to analyze the behavior of the proposed estimator. The proposed estimator compared with Vasicek's estimator, van Es's estimator, Correa's estimator and Ebrahimi et al.'s estimator. For each sample size 20000 samples were generated and the RMSEs of the estimators were computed. Similar to Correa (1995), we considered normal, exponential and uniform distributions. The formula for computing MSE is

$$\text{MSE} = \frac{1}{B}\sum_{i=1}^{B}\{HN_{mn}(i) - H(f)\}^2,$$

where B is number of iterations (here, 20,000) and $HN_{mn}(i)$ is the value of the proposed estimator for ith iteration. Further, $H(f)$ is the value of the population entropy. For example, for the statndard normal, the standard exponential and uniform distributions the value of $H(f)$ is $\log\sqrt{2\pi e}$, 1, 0, respectively.

Still an open problem in entropy estimation is the optimal choice of m for given n. The following heuristic formula for computing the competitors estimators is considered. (see Grzegorzewski and Wieczorkowski, 1999)

$$m = \left[\sqrt{n} + 0.5\right] .$$

Table 1 reports the values of m which the proposed estimator obtains reasonably good (not best) RMSE. With increasing n the optimal choice of m increases.

Table 1. Proposed values of m for different values of n

sample size n	window size m
$n \leqslant 7$	2
$8 \leqslant n \leqslant 15$	3
$16 \leqslant n \leqslant 25$	4
$26 \leqslant n \leqslant 40$	5
$41 \leqslant n \leqslant 60$	6
$61 \leqslant n \leqslant 90$	7
$91 \leqslant n \leqslant 120$	8

Tables 2-4 give the RMSE values of the five estimators at different sample size for each of the three considered distributions.

Table 2. Root of mean square error of estimators in estimate of entropy $H(f)$ for standard normal distribution.

	RMSE				
n	HV_{mn}	HVE_{mn}	HC_{mn}	HE_{mn}	HN_{mn}
5	0.994	0.509	0.793	0.665	0.452
10	0.618	0.366	0.470	0.402	0.277
15	0.474	0.318	0.348	0.301	0.231
20	0.373	0.276	0.265	0.247	0.183
30	0.282	0.243	0.194	0.186	0.145
50	0.198	0.212	0.135	0.128	0.109

Table 3. Root of mean square error of estimators in estimate of entropy $H(f)$ for exponential distribution with mean one.

	RMSE				
n	HV_{mn}	HVE_{mn}	HC_{mn}	HE_{mn}	HN_{mn}
5	0.930	0.596	0.743	0.652	0.556
10	0.570	0.392	0.435	0.401	0.360
15	0.421	0.310	0.328	0.310	0.286
20	0.356	0.274	0.272	0.261	0.245
30	0.276	0.227	0.208	0.203	0.199
50	0.194	0.179	0.155	0.151	0.151

Table 4. Root of mean square error of estimators in estimate of entropy $H(f)$ for uniform distribution on (0,1).

	RMSE				
n	HV_{mn}	HVE_{mn}	HC_{mn}	HE_{mn}	HN_{mn}
5	0.774	0.407	0.566	0.456	0.345
10	0.455	0.216	0.295	0.234	0.173
15	0.343	0.155	0.208	0.160	0.117
20	0.274	0.121	0.157	0.135	0.091
30	0.210	0.086	0.110	0.097	0.065
50	0.156	0.058	0.076	0.063	0.040

We observe that the proposed estimator performs better than the competitor estimators. In fact, it is evident from the simulation results that our proposed estimator can be said to dominate the other estimators.

Acknowledgement

The author wishes to thank the two anonymous referees whose comments helped to improve the presentation of this article.

References

Alizadeh Noughabi, H. and Arghami, N.R. (2011). Monte Carlo Comparison of Five Exponentiality Tests Using Different Entropy Estimates. *Journal of Statistical Computation and Simulation*, **81**, 1579-1592.

Alizadeh Noughabi, H. and Arghami, N.R. (2012). General Treatment of Goodness-of-Fit Tests Based on Kullback-Leibler Information. *Journal of Statistical Computation and Simulation*, **83**, 1556-1569.

Alizadeh Noughabi, H. and Arghami, N.R. (2013). Goodness-of-Fit Tests Based on Correcting Moments of Entropy Estimators. *Communications in Statistics - Simulation and Computation*, **42**, 499-513.

Choi, B. (2008). Improvement of Goodness of Fit Test for Normal Distribution Based on Entropy and Power Comparison. *Journal of Statistical Computation and Simulation*, **78**, 781-788.

Choi, B., Kim, K. and Song, S.H. (2004). Goodness of Fit Test for Exponentiality Based on Kullback-Leibler Information. *Communications in Statistics-Simulation and Computation*, **33**, 525-536.

Correa, J.C. (1995). A New Estimator of Entropy. *Communications in Statistics - Theory and Methods*, **24**, 2439-2449.

Ebrahimi, N., Pflughoeft, K. and Soofi, E. (1994). Two Measures of Sample Entropy. *Statistics and Probability Letters*, **20**, 225-234.

Esteban, M.D., Castellanos, M.E., Morales, D. and Vajda I. (2001). Monte Carlo Comparison of Four Normality Tests Using Different Entropy Estimates. *Communications in Statistics-Simulation and computation*, **30**, 761-785.

Goria, M.N., Leonenko, N.N., Mergel, V.V. and Novi Inverardi, P.L. (2005). A New Class of Random Vector Entropy Estimators and Its Applications in Testing Statistical Hypotheses. *Nonparametric Statistics*, **17**, 277-297.

Grzegorzewski, P. and Wieczorkowski, R. (1999). Entropy-based Goodness-of-fit Test for Exponentiality. *Communications in Statistics-Theory and Methods*, **28**, 1183-1202.

Mack, S.P. (1988). A Comparative Study of Entropy Estimators and Entropy Based Goodness-of-fit Tests. Ph.D. Dissertation, University of California, Riverside.

Park, S. and Park, D. (2003). Correcting Moments for Goodness of Fit Tests Based on Two Entropy Estimates. *Journal of Statistical Computation and Simulation*, **73**, 685-694.

Shannon, C.E. (1948). A Mathematical Theory of Communications. *Bell System Technical Journal*, **27**, 379-423; 623-656.

Theil, J. (1980). The Entropy of Maximum Entropy Distribution. *Economic Letters*, **5**, 145-148.

van Es, B. (1992). Estimating Functionals Related to a Density by Class of Statistics Based on Spacings. *Scandinavian Journal of Statistics*, **19**, 61-72.

Vasicek, O. (1976). A Test for Normality Based on Sample Entropy. *Journal of the Royal Statistical Society, Series B*, **38**, 54-59.

Vexler, A., Shan, G., Kim, S., Tsai, W.M., Tian, L. and Hutson, A.D. (2011). An Empirical Likelihood Ratio Based Goodness-of-fit Test for Inverse Gaussian Distributions. *Journal of Statistical Planning and Inference*, **141**, 2128-2140.

Hadi Alizadeh Noughabi
Department of Statistics,
University oh Birjand,
Birjand, Iran.
email: *alizadehhadi@birjand.ac.ir*