Improving Sample Entropy Based on Nonparametric Quantile Estimation

Sangun Park\textsuperscript{a}, Dongryeon Park\textsuperscript{1,b}

\textsuperscript{a}Department of Applied Statistics, Yonsei University
\textsuperscript{b}Department of Statistics, Hanshin University

Abstract

Sample entropy (Vasicek, 1976) has poor performance, and several nonparametric entropy estimators have been proposed as alternatives. In this paper, we consider a piecewise uniform density function based on quantiles, which enables us to evaluate entropy in each interval, and study the poor performance of the sample entropy in terms of the poor estimation of lower and upper quantiles. Then we propose some improved entropy estimators by simply modifying the quantile estimators, and compare their performances with some existing estimators.

Keywords: Kernel density estimator, maximum entropy, order statistics, sample entropy.

1. Introduction

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

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

The nonparametric entropy estimator of (1.1) can be instantly obtained as

$$H_n(f) = -\int_{-\infty}^{\infty} f_n(x) \log f_n(x) dx,$$  \hfill (1.2)

where $f_n(x)$ is a nonparametric density estimator.

Dmitriev and Tarasenko (1973), and Ahmad and Lin (1976) considered the kernel-type density estimator. Theil (1980) considered the mean and mass preserving maximum entropy distribution, and Dudewicz and van der Meulen (1987) considered the empirical probability density estimator. There have also been some approaches based on the direct estimation of (1.1). Vasicek (1976) introduced the sample entropy based on the spacings of order statistics, whose underlying nonparametric density estimator was later discovered by Park and Park (2003). Ebrahimi \textit{et al.} (1994) modified the sample entropy, and an approach based on linear regression was also provided by Correa (1995). Entropy estimator itself can be considered as a goodness of fit test statistic based on the maximum entropy principle (Vasicek, 1976; Choi and Kim, 2006).

Sample entropy has been widely considered as an entropy estimate but its performance on tails is not so satisfactory. In this paper, we consider a piecewise uniform density function in each interval.
where the endpoints of each interval are defined to be quantiles. Hence, a nonparametric density estimator can be instantly established by estimating the quantiles. Then we can discuss the poor performance of the sample entropy in terms of its underlying nonparametric density estimator. We can consider the amount of the entropy in each interval and confirm the poor performance of the sample entropy in both tails. We further propose some modified entropy estimators by improving the lower and upper quantile estimators and compare their performances with some existing estimators.

2. Tail Behaviors of Sample Entropy

Suppose that a random variable $X$ has a distribution function $F(x)$ with a continuous density function $f(x)$. We will consider the following approximation of $f(x)$ as

$$f_0(x) = \frac{1}{n(\hat{\xi}_i^L - \hat{\xi}_i^U)}, \quad \text{if } \hat{\xi}_i^U < x \leq \hat{\xi}_i^L, \quad i = 1, \ldots, n, \quad (2.1)$$

where $\hat{\xi}_{i/n} = F^{-1}(i/n)$, which is the $(i/n)^{th}$ quantile of $F(x)$.

We can interpret $f_0(x)$ to be a piecewise uniform density function with a mass $1/n$ at each interval $(\hat{\xi}_{i-1/n}, \hat{\xi}_{i/n})$, $i = 1, \ldots, n$. We can easily show that $f_0(x) \to f(x)$ almost everywhere as $n \to \infty$.

Suppose now that we have a sample of size $n$, $x_{i:n}, \ldots, x_{n:n}$, from a continuous density function $f(x)$ where $x_{i:n}$ is the $i^{th}$ ordered sample value. Then we can establish a nonparametric density estimator just by replacing $\hat{\xi}_{i/n}$'s in (2.1), $i = 0, \ldots, n$, with their nonparametric quantile estimators. Theil (1980), and Dudewicz and van der Meulen (1987) considered the similar type of density estimators and established the corresponding nonparametric entropy estimators.

Many nonparametric quantile estimators have been suggested in a function of the ordered sample (Reiss, 1981; Yang, 1985). However, we confine our attention to the average of adjacent order statistics to $(i/n)^{th}$ quantile since it can later provide a simple nonparametric entropy estimator. We define $\hat{\xi}_{i/n}$ to be the average of $2m$ adjacent order statistics to $(i/n)^{th}$ quantile, $x_{i-m+1:n}, \ldots, x_{i+m:n}$ as

$$\hat{\xi}_i = \frac{x_{i-m+1:n} + \cdots + x_{i+m:n}}{2m},$$

where $x_{i:n}$ if $i < 1$ and $x_{n:n}$ if $i > n$.

Then the density function estimator can be written as

$$f_{mn}(x) = \frac{2m}{n(x_{i+m:n} - x_{i-m:n})}, \quad \text{if } \hat{\xi}_i^U < x \leq \hat{\xi}_i^L, \quad i = 1, \ldots, n,$$

where $x_{i+m:n} = x_{n:n}$ if $i + m > n$ and $x_{i-m:n} = x_{1:n}$ if $i - m < 1$.

$f_{mn}(x)$ is the underlying density estimator of the sample entropy (Park and Park, 2003). We can instantly see the defectiveness in both tails of $f_{mn}(x)$ for a finite sample size; The $(i/n)^{th}$ quantile is overestimated for $i < m + 1$ and underestimated for $i > n - m$ since $x_{i:n}$'s are replaced with $x_{1:n}$ and $x_{n:n}$ for $i < 1$ and $i > n$, respectively. For example, we can see that $(1/n)^{th}$ quantile is overestimated with $(m \cdot x_{1:n} + x_{2:n} + \cdots + x_{m+1:n})/2m$. This defectiveness comes from replacing $x_{i:n}$’s for $i < 1$ and $i > n$ with $x_{1:n}$ and $x_{n:n}$, respectively, since we do not have any more data points beyond $(x_{1:n}, x_{n:n})$.

In Figure 1, we draw $f_{mn}(x)$ based on a sample of size 100 from standard normal and exponential distributions where the sample is taken to be $\hat{\xi}_{i/101}$, $i = 0, \ldots, 100$’s to avoid the sample variation. Comparing with true density lines, Figure 1 shows that $f_{mn}(x)$ has a serious defectiveness in both tails and suggests that $f_{mn}(x)$ needs to be modified to be monotone decreasing in both ends.
Improving Sample Entropy Based on Nonparametric Quantile Estimation

Several nonparametric entropy estimators have been suggested by many authors including Vasicek (1976), Theil (1980), Dudewicz and van der Meulen (1987), Bowman (1992) and Correa (1995), Ebrahimi et al. (1994). While Vasicek (1976) and Ebrahimi et al. (1994) directly obtained the nonparametric density function and obtained its entropy by evaluating (1.2). The sample entropy (Vasicek, 1976), which is the entropy of $f_{mn}(x)$, has been most popular among the mentioned nonparametric entropy estimators, as

$$H_{mn} = H(f_{mn}) = \frac{1}{n} \sum_{i=1}^{n} \log \left( \frac{n}{2m} (x_{i+m:n} - x_{i-m:n}) \right),$$

where $x_{i:n} = x_{1:n}$ for $i < 1$ and $x_{i:n} = x_{n:n}$ for $i > n$.

We can assume that the performance of $H_{mn}$ is not so good and that most of poor performance comes from tails since the underlying distribution, $f_{mn}$, has the problem of poor estimation of lower and upper quantiles. To verify this, we first calculate true amount of the entropy in each interval as

$$H_i(f) = - \int_{x_{i-1:n}}^{x_{i:n}} f(x) \log f(x) dx$$

and compare it with its estimate based on $f_{mn}$ as

$$H_i(f_{mn}) = \frac{1}{n} \log \left( \frac{n}{2m} (x_{i+m:n} - x_{i-m:n}) \right).$$

In Figure 2, we calculate $H_i(f)$ for the standard normal, exponential, uniform and Weibull (shape = 2) distributions and $H_i(f_{mn})$ based on a sample of size 20 from each distribution where the sample
is taken to be $\xi_{i/21}$, $i = 1, \ldots, 20$’s to avoid sample variation. We see that the sample entropy underestimates the true entropy in both tails and the difference increases in both tails. Hence, we can confirm that the poor performance of the sample entropy comes from the poor estimation of the entropy in both tails, that basically results from the poor estimation of lower and upper quantiles in $f_{mn}$.

### 3. Modified Density Estimators

$f_{mn}$ in the previous section is defined based on a bounded support, $(\hat{\xi}_0, \hat{\xi}_1)$. In modifying $f_{mn}$ to be defined on $(-\infty, \infty)$, we cannot establish the piecewise uniform distribution in $(-\infty, \hat{\xi}_{1/n})$ and $(\hat{\xi}_{(n-1)/n}, \infty)$. Hence, we will employ the exponential smoothing in those subintervals, which has been also discussed in Theil (1980). Then the density function in $(-\infty, \hat{\xi}_{1/n})$ can be taken to be two-parameter exponential distribution function, $n^{-1} \exp[(x - \hat{\xi}_{1/n})/\theta]/\theta$, instead of the uniform distribution. Determining the scale parameter is a remaining part but we determine here $\theta$ to be $\xi_{2/n} - \hat{\xi}_{1/n}$ so that the density function is continuous at $\hat{\xi}_{1/n}$. In a similar way, the distribution in $(\hat{\xi}_{(n-1)/n}, \infty)$ can be taken to be $n^{-1} \exp(-(x - \hat{\xi}_{(n-1)/n})/\theta)/\theta$ and $\theta$ can be determined to be $\xi_{(n-1)/n} - \hat{\xi}_{(n-2)/n}$.
Hence, the exponentially smoothed piecewise uniform density function of (2.1) is

\[
f_{sn}(x) = \begin{cases} 
\frac{1}{n(\xi_{i+1}^n - \xi_i^n)} \exp \left\{ \frac{x - \xi_i^n}{\xi_{i+1}^n - \xi_i^n} \right\}, & \text{if } x \leq \xi_i^n, \\
\frac{1}{n(\xi_{i-1}^n - \xi_i^n)} \exp \left\{ \frac{x - \xi_i^n}{\xi_{i-1}^n - \xi_i^n} \right\}, & \text{if } \xi_{i-1}^n < x \leq \xi_i^n, \quad i = 2, \ldots, n-1, \\
\frac{1}{n(\xi_{(m-1)n}^n - \xi_{(m-2)n}^n)} \exp \left\{ \frac{x - \xi_{(m-2)n}^n}{\xi_{(m-1)n}^n - \xi_{(m-2)n}^n} \right\}, & \text{if } x \geq \xi_{(m-1)n}^n.
\end{cases}
\]  

(3.1)

Next, we will provide two alternatives to the quantile estimator \( \hat{\xi}_{i/n} \) employed in \( f_{sn}(x) \). We see that \( \hat{\xi}_{i/n} \) is not an adequate estimator of \( \xi_{i/n} \) for large \( m \). The first alternative can be established by generating pseudo data points as

\[
x_{i,n} = x_{1,n} + (i - 1)d, \quad i < 1 \quad \text{and} \\
x_{i,n} = x_{n,n} + (i - n)d, \quad i > n,
\]

where \( d \) is the estimated gap of order statistics of order 1. The possible candidates of \( d \) are \( (x_{i+1,n} - x_{i,n})'s \) for \( i = 1, \ldots, m \) so that their average, \( \sum_{i=1}^{m} (x_{i+1,n} - x_{i,n})/m = (x_{m+1,n} - x_{1,n})/m \), can be a possible choice. Then we can replace \( x_{i,n} \)'s for \( i < 1 \) and \( i > n \) with the pseudo data points as follow:

\[
x_{i,n} = \begin{cases} 
x_{1,n} + (i - 1) \left( \frac{x_{m+1,n} - x_{1,n}}{m} \right), & \text{if } i < 1, \\
x_{n,n} + (i - n) \left( \frac{x_{n,n} - x_{n-m,n}}{m} \right), & \text{if } i > n.
\end{cases}
\]  

(3.2)

Hence, the nonparametric quantile estimator can be written as

\[
\hat{\xi}_{i/n} = \frac{x_{i-m+1,n} + \cdots + x_{i+m,n}}{2m},
\]

where \( x_{i,n} \)'s for \( i < 1 \) and \( i > n \) are defined in (3.2).

The proposed replacement procedure actually adds some pseudo data points beyond each end. Therefore, this modified density estimator may not be adequate for the bounded support case since some lower and upper quantile estimators are outside the support.

The second alternative can be established by estimating the quantiles with only available order statistics so that

\[
\hat{\xi}_a = \frac{x_{a+1,n} + \cdots + x_{i+m,n}}{2m},
\]

where \( a \) is chosen to be \( \min(m, i, n - i) \).

For example, \( \hat{\xi}_{1/n} \) is estimated with \( (x_{1,n} + x_{2,n})/2 \). We see that both \( \hat{\xi}_a \) and \( \hat{\xi}_b \) are within the given support even for a distribution function with bounded support. Hence, \( \hat{\xi}_0 \) and \( \hat{\xi}_1 \) can be determined to be \( a \) and \( b \), respectively, for a distribution function defined on \( (a, b) \).

Hence, the smoothed piecewise uniform density estimators can be established by replacing \( \hat{\xi}_{i/n} \)'s in \( f_{sn}(x) \) with \( \hat{\xi}_{i/n} \) or \( \hat{\xi}_{i/n} \) 's, which we will finally denote to be \( f_{sn}^a(x) \) and \( f_{sn}^b(x) \), respectively. A simulation study was conducted to compare the finite sample properties of \( f_{sn}^a(x) \) and \( f_{sn}^b(x) \). For
Table 1: Monte Carlo MISE of \( f_{mn}(x), f_{mn}^a(x), f_{mn}^b(x) \) and kernel density estimator \( \hat{f}_k(x) \) for a standard normal distribution.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( f_k )</th>
<th>( m )</th>
<th>( f_{mn} )</th>
<th>( m )</th>
<th>( f_{mn}^a )</th>
<th>( m )</th>
<th>( f_{mn}^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.0097</td>
<td>4</td>
<td>0.1423</td>
<td>4</td>
<td>0.0455</td>
<td>4</td>
<td>0.00502</td>
</tr>
<tr>
<td>20</td>
<td>0.0047</td>
<td>5</td>
<td>0.0428</td>
<td>7</td>
<td>0.0153</td>
<td>8</td>
<td>0.0196</td>
</tr>
<tr>
<td>30</td>
<td>0.0034</td>
<td>6</td>
<td>0.0230</td>
<td>8</td>
<td>0.0101</td>
<td>10</td>
<td>0.0092</td>
</tr>
<tr>
<td>40</td>
<td>0.0027</td>
<td>7</td>
<td>0.0157</td>
<td>9</td>
<td>0.0081</td>
<td>13</td>
<td>0.0068</td>
</tr>
<tr>
<td>50</td>
<td>0.0023</td>
<td>8</td>
<td>0.0124</td>
<td>11</td>
<td>0.0097</td>
<td>15</td>
<td>0.0055</td>
</tr>
</tbody>
</table>

comparison, we consider the nonparametric kernel density estimator based on the empirical density function, that can be written as

\[
f_k(x) = \frac{1}{nh} \sum_{i=1}^{n} K \left( \frac{x - X_i}{h} \right),
\]

where \( K \) is a kernel density function and \( h \) is a bandwidth.

We generated the random sample from the standard normal distribution with sample size of \( n = 10, 20, 30, 100 \).

Table 1 presents optimal \( m \) which produces the least MISE (Mean Integrated Squared Error) and the corresponding MISE of \( f_{mn}(x), f_{mn}^a(x), f_{mn}^b(x) \) for a standard normal distribution over 10,000 simulation samples. The MISE of kernel density estimator, \( f_k(x) \) is also presented. The gaussian function was used for the kernel function of \( f_k(x) \) and the bandwidth was selected by the rule of thumb which is known as a very efficient rule for the normal data.

Note that both \( f_{mn}(x) \) and \( f_{mn}^a(x) \) are superior to \( f_{mn}(x) \) for all cases. We can see that the performance of piecewise uniform density estimator improved substantially by modifying the tail parts. However, \( f_{mn}^a(x) \) and \( f_{mn}^b(x) \) are not as good as \( f_k(x) \). Therefore, it is not a good idea to use the suggested here density estimators instead of \( f_k(x) \) as a density estimator. However, we should remember that \( f_{mn}^a(x) \) and \( f_{mn}^b(x) \) have merits over \( f_k(x) \) that it is easier to evaluate its any functional form and this will be discuss in the next section.

4. Nonparametric Entropy Estimation

The improved entropy estimator can be obtained by taking the entropy of \( f_{mn}^a(x) \) defined in the previous section as follows:

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

where \( X_{i,n} \) for \( i < 1 \) and \( i > n \) are defined in (3.2).

We note that \( H_{mn}^a \) does not employ the boundary information since some lower and upper quantile estimates may be outside the support.

The entropy of \( f_{mn}^a(x) \) for the unbounded support case can be also obtained as

\[
H_{mn}^a = -\frac{1}{n} \left\{ \sum_{i=2}^{n-1} \log \left( n \left( \hat{\xi} - \hat{\xi}_{i,n} \right) \right) + \left( \log \left( n \left( \hat{\xi} - \hat{\xi}_{i,1} \right) \right) + 1 \right) + \left( \log \left( n \left( \hat{\xi} - \hat{\xi}_{n,1} \right) \right) + 1 \right) \right\},
\]
where

\[ \hat{\xi}_i = \frac{x_{i-a+1:n} + \cdots + x_{i+n:n}}{2a} \]

and \( a = \min(m, i, n - i) \).

However, \( H_{bsn}^b \) employs the boundary information by taking \( \hat{\xi}_0 \) and \( \hat{\xi}_1 \) to be lower and upper bounds, respectively. For example, \( \hat{\xi}_0 \) and \( \hat{\xi}_1 \) are taken to be \( a \) and \( b \) for a distribution defined on \((a, b)\), and \( H_{bsn}^b \) can be obtained as

\[ H_{bsn}^b = \frac{1}{n} \left\{ \sum_{i=1}^{n} \log \left( n \left( \frac{\hat{\xi}_i - \hat{\xi}_{i+1}}{c_i} \right) \right) \right\}. \]

Hence, the exponential smoothing in \( H_{bsn}^b \) can be done only in the unbounded direction. For the lifetime data defined on \((0, 1)\), \( \hat{\xi}_0 \) is taken to be 0 and the exponential smoothing can be done only on the right side.

We present the results of the simulation study for comparing the finite sample properties of two suggested entropy estimators. In comparison, we include two entropy estimators suggested by Ebrahimi et al. (1994) and Correa (1995), which are based on improved estimators of \( dF^{-1}(u)/du \). Ebrahimi et al. (1994) adjusted the expected value of \( X_i + m:n \) and suggested

\[ H_e(n, m) = \frac{1}{n} \sum_{i=1}^{n} \log \left( \frac{X_{i+m:n} - X_{i-m:n}}{c_i m/n} \right), \]

where

\[ c_i = \begin{cases} 
1 + \frac{i - 1}{m}, & \text{if } 1 \leq i \leq m, \\
2, & \text{if } m + 1 \leq i \leq n - m, \\
1 + \frac{n - i}{m}, & \text{if } n - m + 1 \leq i \leq n.
\end{cases} \]

Ebrahimi et al. (1994) also generated pseudo data points by giving lower limit \( a \) and upper limit \( b \) and provided an entropy estimator, which are useful in the case of the bounded support, as

\[ H_d(n, m) = \frac{1}{n} \sum_{i=1}^{n} \log \left( \frac{z_{i+m} - z_{i-m}}{d_i m/n} \right), \]

where the \( z_i \)’s are

\[ \begin{align*}
z_{i-m} &= a + \frac{i - 1}{m} (x_{1:n} - a), & \text{if } 1 \leq i \leq m, \\
z_i &= x_{i:n}, & \text{if } m + 1 \leq i \leq n - m - 1, \\
z_{i+m} &= b - \frac{n - i}{m} (b - x_{n:n}), & \text{if } n - m \leq i \leq n
\end{align*} \]

and

\[ d_i = \begin{cases} 
1 + \frac{i + 1}{m^2} - \frac{i}{m^2}, & \text{if } 1 \leq i \leq m, \\
2, & \text{if } m + 1 \leq i \leq n - m - 1, \\
1 + \frac{n - i}{m^2}, & \text{if } n - m \leq i \leq n.
\end{cases} \]
in Table 3. We note that modified entropy estimators, except the sample entropy, are comparable for both distributions, but taken to be 0 and the exponential smoothing is done only on (entropy estimators based on 10,000 simulations. In Table 2, we compare where $\bar{X}$ of $dF$ may be less than the lower limit $a$. Hence, we consider $H_d(n,m)$ only in the uniform distribution.

We also consider the entropy estimator (Correa, 1995), which is based on the regression estimate of $dF^{-1}(u)/du$.

$$H_{cor}(n,m) = -\frac{1}{n} \sum_{i=1}^{n} \log \frac{\sum_{j=i-m}^{i+m} (X_{j:n} - \hat{X}_i)(j/n - i/n)}{\sum_{j=i-m}^{i+m} (X_{j:n} - \hat{X}_i)^2},$$

where $\hat{X}_i = \sum_{j=1-m}^{j=m} X_{j:n}/(2m + 1)$.

In Tables 2 to 4, we determine the optimal $m$ to be one which produces the least mean squared error, and present the optimal $m$’s and the corresponding root of mean squared errors (RMSE) of entropy estimators based on 10,000 simulations. In Table 2, we compare $H_{uni}$, $H_{d}$, $H_{cor}$, $H_{smn}^a$ and $H_{smn}^b$ for normal distribution. The same simulation has been done for the exponential distribution in Table 3. We note that $H_{smn}^b$ for the exponential distribution is the entropy of $f_{uni}^b(x)$ where $\xi_0$ is taken to be 0 and the exponential smoothing is done only on $(\xi(n-1)/n, \infty)$. The performances of all modified entropy estimators, except the sample entropy, are comparable for both distributions, but $H_{smn}^b$ slightly outperforms others. We also consider the uniform distribution in Table 4 and include $H_d$ instead of $H_{uni}$ for comparison. In Table 4, $H_{smn}^b$ slightly outperforms; however, the performance of $H_{smn}^b$ is not so good though the latter utilized the boundary information. This is because, we think, the chosen lower and upper quantile estimators are not good for the uniform distribution; In $H_{smn}^b$, $\hat{\xi}_{1/2}$! n is $(x_{1:n} + x_{2:n})/2$ and its expected value is $1.5/(n + 1)$ which is quite far from $1/n$ for a small sample size.

| Table 2: Monte Carlo RMSE of $H_{uni}$, $H_{d}$, $H_{cor}$, $H_{smn}^a$ and $H_{smn}^b$ for normal distribution. |
|---|---|---|---|---|---|
| $n$ | $m$ | $H_{uni}$ | $m$ | $H_{d}$ | $m$ | $H_{cor}$ | $m$ | $H_{smn}^a$ | $m$ | $H_{smn}^b$ |
| 10 | 2 | 0.5990 | 4 | 0.3986 | 4 | 0.3483 | 4 | 0.2989 | 4 | 0.3085 |
| 20 | 3 | 0.3625 | 7 | 0.2363 | 4 | 0.2028 | 5 | 0.1939 | 9 | 0.1927 |
| 30 | 4 | 0.2759 | 10 | 0.1736 | 5 | 0.1528 | 6 | 0.1520 | 14 | 0.1493 |
| 40 | 4 | 0.2256 | 12 | 0.1402 | 5 | 0.1273 | 6 | 0.1290 | 11 | 0.1253 |
| 50 | 5 | 0.1964 | 13 | 0.1221 | 5 | 0.1135 | 7 | 0.1163 | 12 | 0.1121 |

| Table 3: Monte Carlo RMSE of $H_{uni}$, $H_{d}$, $H_{cor}$, $H_{smn}^a$ and $H_{smn}^b$ for exponential distribution. |
|---|---|---|---|---|---|
| $n$ | $m$ | $H_{uni}$ | $m$ | $H_{d}$ | $m$ | $H_{cor}$ | $m$ | $H_{smn}^a$ | $m$ | $H_{smn}^b$ |
| 10 | 2 | 0.5596 | 4 | 0.3820 | 3 | 0.3686 | 3 | 0.3596 | 4 | 0.3567 |
| 20 | 4 | 0.3524 | 6 | 0.2511 | 4 | 0.2477 | 4 | 0.2470 | 6 | 0.2428 |
| 30 | 5 | 0.2715 | 7 | 0.1982 | 4 | 0.1950 | 5 | 0.1962 | 7 | 0.1926 |
| 40 | 6 | 0.2278 | 6 | 0.1708 | 4 | 0.1684 | 5 | 0.1697 | 8 | 0.1664 |
| 50 | 7 | 0.1964 | 8 | 0.1514 | 4 | 0.1500 | 6 | 0.1509 | 8 | 0.1484 |

| Table 4: Monte Carlo RMSE of $H_{uni}$, $H_{d}$, $H_{cor}$, $H_{smn}^a$ and $H_{smn}^b$ for uniform distribution. |
|---|---|---|---|---|---|
| $n$ | $m$ | $H_{uni}$ | $m$ | $H_{d}$ | $m$ | $H_{cor}$ | $m$ | $H_{smn}^a$ | $m$ | $H_{smn}^b$ |
| 10 | 2 | 0.4529 | 4 | 0.1373 | 4 | 0.1784 | 4 | 0.1594 | 4 | 0.1483 |
| 20 | 3 | 0.2699 | 9 | 0.0799 | 7 | 0.0881 | 8 | 0.0795 | 8 | 0.0952 |
| 30 | 4 | 0.2037 | 12 | 0.0576 | 9 | 0.0581 | 11 | 0.0535 | 12 | 0.0713 |
| 40 | 4 | 0.1695 | 15 | 0.0461 | 10 | 0.0449 | 13 | 0.0417 | 16 | 0.0582 |
| 50 | 4 | 0.1470 | 18 | 0.0386 | 12 | 0.0368 | 15 | 0.0350 | 20 | 0.0496 |
5. Conclusion

We use the piecewise uniform density function, where the intervals are determined to be quantiles, and verify the poor performance of the sample entropy on tails. Then we propose some modified entropy estimators by simply improving the lower and upper quantile estimators. It is also shown that the performances of the proposed estimators are comparable with some competing estimators.

References


Received March 2011; Accepted May 2011