Journal of Mathematical Extension
Vol. XX, No. XX, (2014), pp-pp (Will be inserted by layout editor)
ISSN: 1735-8299
URL: http://www.ijmex.com

A refined upper bound for entropy of stochastic process

Yamin Sayyari*
Department of Mathematics, Sirjan University Of Technology, Sirjan, Iran

Abstract. Estimation of Shannon's entropies of stochastic process from numerical simulation of long orbits is difficult. Our aim within this paper is to present a strong upper bound for the Shannon's entropy of information sources.

AMS Subject Classification: 37B40; 94A17; 26A51.
Keywords and Phrases: Entropy, Shannon's entropy, Information source, Stochastic process, Random variable.

1 Introduction

Entropy and mutual information for discrete-valued random variables play important roles in information theory. The entropy actually measures the degree of irregularities of a dynamic system, and researchers have done so much to calculate this concept, which is often successful [1, 4], but numerical calculations of entropy are still difficult. Tapus and Popescu presented a strong upper bound for the classical Shannon entropy [3]. In [3, 6, 8], the authors presented a strong upper bound for the classical Shannon entropy. In [5], the authors presented the algebraic and Shannon entropies for hypergroupoids and commutative

Received: XXXX; Accepted: XXXX (Will be inserted by editor)
*Corresponding Author

hypergroups, respectively, and studies their fundamental properties. In [8], the author applying Jensens inequality in information theory and we obtain some results for the Shannons entropy of random variables and Shannons entropy of stochastic process. Our purpose within this work is to present a strong upper bound for the Shannon entropy of information sources, refining recent results from the literature.

Let $X \neq \emptyset$ be a set. Then $(X, \mathcal{F}, \mu)$ is called measure probability space if $\mathcal{F}$ is an σ-algebra of subsets of X, μ is a measure on X, and $\mu(X) = 1$. A finite set of measurable sets $\alpha = \{A_1, \ldots, A_n\}$ is called a finite partition if the following properties are fulfilled [7]:

$$\bigcup_{i=1}^{n} A_i = X, \quad \text{and} \quad A_i \cap A_j = \emptyset \quad \text{for every } i, j (1 \leq i \neq j \leq n).$$

For a partition $\alpha = \{A_1, \ldots, A_n\}$, the entropy of α is defined by

$$H_\mu(\alpha) := -\sum_{i=1}^{n} \mu(A_i) \log(\mu(A_i)).$$

Definition 1.1. [2] Let S be a random variable on X with discrete finite state space $A = \{a_1, ..., a_N\}$. We define $p : A \to [0, 1]$ by $p(s) = \mu\{\omega \in X : S(\omega) = s\}$. The Shannon's entropy of S is defined by

$$H_\mu(S) := -\sum_{s \in A,\ p(s) \neq 0} p(s) \log p(s).$$

A stochastic process $\mathbf{S}$ is a sequence $(S_n)_{n=1}^{\infty}$ of the random variables $S_n : X \longrightarrow A$, where $n \in \mathbb{N}$. For given $L \geq 1$ we define a mapping $p : A^L \to [0, 1]$ by $p(s_1^L) = \mu\{\omega \in X : S_1(\omega) = s_1, ..., S_L(\omega) = s_L\}$. The Shannon entropy of order L and the Shannon entropy of source $\mathbf{S}$ are respectively defined by

$$H_\mu(S_1^L) = -\frac{1}{L} \sum_{s_1^L \in A^L} p(s_1, ..., s_L) \log p(s_1, ..., s_L), \text{ and } h_\mu(\mathbf{S}) = \lim_{L \to \infty} H_\mu(S_1^L).$$

where the the summation is taken over the collection $\{s_1^L \in A^L : p(s_1^L) \neq 0\}$. In this paper we use the symbol s_1^L instead of notation $(s_1, ..., s_L)$ and Let $p(s_1^L) \neq 0$ for every $L \in \mathbb{N}$.

2 main results

In this section, we continue with a refinement of Theorem 2.1 from [8], as follows:

Theorem 2.1. *Let $I=[a,b]$ be an interval, $H:A^L \longrightarrow I$ be a function, and $f:I \longrightarrow \mathbb{R}$ be a convex function, then*

$$\sum_{s_1^L \in A^L} p(s_1^L) f(H(s_1^L)) - f(\sum_{s_1^L \in A^L} p(s_1^L) H(s_1^L))$$
$$\geq \max\{p(r_1^L) f(H(r_1^L)) + p(t_1^L) f(H(t_1^L)) + p(u_1^L) f(H(u_1^L))$$
$$-(p(r_1^L)+p(t_1^L)+p(u_1^L)) f(\frac{p(r_1^L)H(r_1^L)+p(t_1^L)H(t_1^L)+p(u_1^L)H(u_1^L)}{p(r_1^L)+p(t_1^L)+p(u_1^L)})\},$$
$$\geq \max\{p(r_1^L) f(H(r_1^L)) + p(t_1^L) f(H(t_1^L)) + p(u_1^L) f(H(u_1^L))\}$$
$$-(p(r_1^L)+p(t_1^L)+p(u_1^L)) f(\frac{p(r_1^L)H(r_1^L)+p(t_1^L)H(t_1^L)+p(u_1^L)H(u_1^L)}{p(r_1^L)+p(t_1^L)+p(u_1^L)})\},$$

where the maximum is taken over all distinct $r_1^L, t_1^L, u_1^L \in A^L$.

Proof. Choose arbitrary $t_1^L, r_1^L, u_1^L \in A^L$. So,

$$f(\sum_{s_1^L \in A^L} p(s_1^L) H(s_1^L)) = f(\sum_{s_1^L \neq r_1^L, t_1^L, u_1^L \in A^L} p(s_1^L) H(s_1^L))$$
$$+(p(r_1^L)+p(t_1^L)+p(u_1^L))(\frac{p(r_1^L)H(r_1^L)+p(t_1^L)H(t_1^L)+p(u_1^L)H(u_1^L)}{p(r_1^L)+p(t_1^L)+p(u_1^L)})$$
$$\leq \sum_{s_1^L \neq r_1^L, t_1^L . u_1^L \in A^L} p(s_1^L) f(H(s_1^L))$$
$$+(p(r_1^L)+p(t_1^L)+p(u_1^L)) f(\frac{p(r_1^L)H(r_1^L)+p(t_1^L)H(t_1^L)+p(u_1^L)H(u_1^L)}{p(r_1^L)+p(t_1^L)+p(u_1^L)}).$$

Therefore,

$$\sum_{s_1^L \in A^L} p(s_1^L) f(H(s_1^L)) - f(\sum_{s_1^L \in A^L} p(s_1^L) H(s_1^L))$$
$$\geq p(r_1^L) f(H(r_1^L)) + p(t_1^L) f(H(t_1^L)) + p(u_1^L) f(H(u_1^L))$$
$$-(p(r_1^L)+p(t_1^L)+p(u_1^L)) f(\frac{p(r_1^L)H(r_1^L)+p(t_1^L)H(t_1^L)+p(u_1^L)H(u_1^L)}{p(r_1^L)+p(t_1^L)+p(u_1^L)}).$$

Since $s_1^L, t_1^L \in A^L, u_L^1$ are arbitrary,

$$\begin{aligned}
&\sum_{s_1^L \in A^L} p(s_1^L) f(H(s_1^L)) - f(\sum_{s_1^L \in A^L} p(s_1^L) H(s_1^L)) \\
&\geq \max\{p(r_1^L) f(H(r_1^L)) + p(t_1^L) f(H(t_1^L)) + p(u_1^L) f(H(u_1^L))\} \\
&- (p(r_1^L) + p(t_1^L) + p(u_1^L)) f(\frac{p(r_1^L) H(r_1^L) + p(t_1^L) H(t_1^L) + p(u_1^L) H(u_1^L)}{p(r_1^L) + p(t_1^L) + p(u_1^L)})\}.
\end{aligned}$$

where the maximum is taken over all distinct $r_1^L, t_1^L, u_1^L \in A^L$. On the other hand,

$$\begin{aligned}
&f(\frac{p(r_1^L) H(r_1^L) + p(t_1^L) H(t_1^L) + p(u_1^L) H(u_1^L)}{p(r_1^L) + p(t_1^L) + p(u_1^L)}) \\
&= f(\frac{p(r_1^L) + p(t_1^L)}{p(r_1^L) + p(t_1^L) + p(u_1^L)} \frac{p(r_1^L) H(r_1^L) + p(t_1^L) H(t_1^L)}{p(r_1^L) + p(t_1^L)} + \frac{p(u_1^L) H(u_1^L)}{p(r_1^L) + p(t_1^L) + p(u_1^L)}) \\
&\leq \frac{p(r_1^L) + p(t_1^L)}{p(r_1^L) + p(t_1^L) + p(u_1^L)} f(\frac{p(r_1^L) H(r_1^L) + p(t_1^L) H(t_1^L)}{p(r_1^L) + p(t_1^L)}) \\
&+ \frac{p(u_1^L)}{p(r_1^L) + p(t_1^L) + p(u_1^L)} f(H(u_1^L)).
\end{aligned}$$

So,

$$\begin{aligned}
&(p(r_1^L) + p(t_1^L) + p(u_1^L)) f(\frac{p(r_1^L) H(r_1^L) + p(t_1^L) H(t_1^L) + p(u_1^L) H(u_1^L)}{p(r_1^L) + p(t_1^L) + p(u_1^L)}) \\
&\leq (p(r_1^L) + p(t_1^L)) f(\frac{p(r_1^L) H(r_1^L) + p(t_1^L) H(t_1^L)}{p(r_1^L) + p(t_1^L)}) + (p(u_1^L)) f(H(u_1^L)).
\end{aligned}$$

Thus,

$$\begin{aligned}
&p(r_1^L) f(H(r_1^L)) + p(t_1^L) f(H(t_1^L)) + p(u_1^L) f(H(u_1^L)) \\
&- (p(r_1^L) + p(t_1^L) + p(u_1^L)) f(\frac{p(r_1^L) H(r_1^L) + p(t_1^L) H(t_1^L) + p(u_1^L) H(u_1^L)}{p(r_1^L) + p(t_1^L) + p(u_1^L)}) \\
&\geq p(r_1^L) f(H(r_1^L)) + p(t_1^L) f(H(t_1^L))\} \\
&- (p(r_1^L) + p(t_1^L)) f(\frac{p(r_1^L) H(r_1^L) + p(t_1^L) H(t_1^L)}{p(r_1^L) + p(t_1^L)}),
\end{aligned}$$

which completes the proof. □

In order to present the generalization we define some notation, as follows:

$$T_k := \max\{\sum_{i=1}^{k} p({r_i}_1^L) f(H({r_i}_1^L)) - (\sum_{i=1}^{k} p({r_i}_1^L)) f(\frac{\sum_{i=1}^{k} p({r_i}_1^L) H({r_i}_1^L)}{\sum_{i=1}^{k} p({r_i}_1^L)})\}$$

where $2 \le k \le N^L - 1$, the maximum is taken over all distinct ${r_i}_1^L \in A^L$.

Theorem 2.2. *Let $I = [a, b]$ be an interval, $H : A^L \longrightarrow I$ be a function, $|A| = N$ and $f : I \longrightarrow \mathbb{R}$ be a convex function, then*

$$0 \le T_2 \le T_3 \le ... \le T_{N^L-1} \le \sum_{s_1^L \in A^L} p(s_1^L) f(H(s_1^L)) - f(\sum_{s_1^L \in A^L} p(s_1^L) H(s_1^L)).$$

Proof. The proof is similar to the proof of Theorem 2.1. □

3 The sources entropy upper bound

In this section we present a strong upper bound for the Shannon's entropy of information sources.

Theorem 3.1. $h_\mu(\mathbf{S}) \le \log N - \max_k \{\lim_{L\to\infty} \frac{1}{L} log[\{\frac{k}{\sum_{i=1}^{k} p({r_i}_1^L)}\}^{\sum_{i=1}^{k} p({r_i}_1^L)}] \times [\prod_{i=1}^{k} \{p({r_i}_1^L)\}^{p({r_i}_1^L)}]\}$

Proof. Since

$$- LH_\mu(S_1^L) + \log(N^L) \ge \max_k \{-\sum_{i=1}^{k} p({r_i}_1^L) \log(\frac{1}{p({r_i}_1^L)}) + (\sum_{i=1}^{k} p({r_i}_1^L))$$
$$\times \log(\frac{k}{\sum_{i=1}^{k} p({r_i}_1^L)})\} = \max_k \{\log(\prod_{i=1}^{k} \{p({r_i}_1^L)\}^{p({r_i}_1^L)})$$
$$+ \log[\{\frac{k}{\sum_{i=1}^{k} p({r_i}_1^L)}\}^{\sum_{i=1}^{k} p({r_i}_1^L)}]\},$$

$$\log N - H_\mu(S_1^L) \ge \max\{\frac{1}{L} log[\{\frac{k}{\sum_{i=1}^{k} p({r_i}_1^L)}\}^{\sum_{i=1}^{k} p({r_i}_1^L)}][\prod_{i=1}^{k} \{p({r_i}_1^L)\}^{p({r_i}_1^L)}]\},$$

and

$$H_\mu(S_1^L) \leq \log N - \max_k \{\frac{1}{L} log[\{\frac{k}{\sum_{i=1}^k p(r_{i_1}^L)}\}^{\sum_{i=1}^k p(r_{i_1}^L)}][\prod_{i=1}^k \{p(r_{i_1}^L)\}^{p(r_{i_1}^L)}]\}.$$

Therefore,

$$h_\mu(\mathbf{S})$$
$$\leq \log N - \lim_{L\to\infty} \max_k \{\frac{1}{L} log[\{\frac{k}{\sum_{i=1}^k p(r_{i_1}^L)}\}^{\sum_{i=1}^k p(r_{i_1}^L)}][\prod_{i=1}^k \{p(r_{i_1}^L)\}^{p(r_{i_1}^L)}]\}$$
$$\leq \log N - \max_{2\leq k\leq N^L-1} \{\lim_{L\to\infty} \frac{1}{L} log[\{\frac{k}{\sum_{i=1}^k p(r_{i_1}^L)}\}^{\sum_{i=1}^k p(r_{i_1}^L)}][\prod_{i=1}^k \{p(r_{i_1}^L)\}^{p(r_{i_1}^L)}]\}.$$

□

References

[1] J.M. Amigo, M.B. Kennel, Topological permutation entropy. Physica D. 231

[2] J.M. Amigo, Permutation Complexity in Dynamical Systems "Ordinal Patterns, Permutation Entropy, and All That", Springer-Verlag, Berlin, 2010.

[3] N.Tapus, P.G. Popescu, A new entropy upper bound, Appl. Math. Lett. 25 (11) 2012, 1887-1890.

[4] Ch. Corda, M. FatehiNia, M.R. Molaei, Y. Sayyari, Entropy of Iterated Function Systems and Their Relations with Black Holes and Bohr-Like Black Holes Entropies, Entropy. 20, 56, 2018.

[5] A. Mehrpooya, Y. Sayyari, M.R. Molaei, Algebraic and Shannon entropies of commutative hypergroups and their connection with information and permutation entropies and with calculation of entropy for chemical algebras, Soft Computing 23 (24) 2019, 13035-13053.

[6] S. SIMIC, Jensens inequality and new entropy bounds, Appl. Math. Lett. 22 (8) 2009, 1262-1265.

[7] P. Walters, An introduction to ergodic theory (Springer Verlag, New York, (2000).

[8] Y. Sayyari, New bounds for entropy of information sources, Wavelets and Linear Algebra, 7(2) 2020, 1-9.

[9] Y. Sayyari, New entropy bounds via uniformly convex functions, Chaos, Solitons and Fractals 141 (2020) 110360 (Doi.org/10.1016/j.chaos.2020.110360).

Yamin Sayyari
Department of Mathematics
Assistant Professor of Mathematics
Sirjan University Of Technology
Sirjan, Iran
E-mail: ysayyari@gmail.com