RELATIVE ENTROPY AND RELATIVE CONDITIONAL ENTROPY WITH INFINITE PARTITIONS

Mohamad Hosein Asadiyan

Department of Mathematics
Bandar Abbas Branch
Islamic Azad University
Bandar Abbas
Iran
e-mail: m.asadian@iauba.ac.ir

Abolfazl Ebrahimzadeh

Department of Mathematics
Zahedan Branch
Islamic Azad University
Zahedan
Iran
e-mail: Abolfazl35@yahoo.com

Abstract. In this paper, we introduce the notions of relative entropy and relative conditional entropy for infinite partitions on a relative probability measure space. We present some examples and prove some theorems about relative conditional entropy.

Keywords: observer; entropy; relative entropy; relative conditional entropy.

AMS Subject Classification: 28D20, 37A35, 54C70.

1. Introduction

The study of concept entropy is very important in the current sciences. Entropy plays an important role in a variety of problem areas, including physics, computer science, general systems theory, information theory, statistics, biology, chemistry, sociology and others. In 1958, Kolmogorov introduced the concept of entropy in the ergodic theory. Entropy is defined on several structures [2]-[4]. Molaie in [9] has studied the notion of one dimensional observer. This notion has been applied in dynamical systems [7], [12], topology [4], [6], [8], geometry [10], and mathematical physics [10]. Let X be a non-empty set, then the function $\mu : X \to [0, 1]$ is called a one-dimensional observer of X. In this paper we assume that $f : X \to X$ is a mapping and $\mu : X \to [0, 1]$ is an observer of X and E is an arbitrary subset of X. The relative probability measure of E with respect to an observer μ is the function $m_\mu^f(E) : X \to [0,1]$ is defined in [5] by

$$m_\mu^f(E)(x) = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_E(f^i(x))\mu(f^i(x)).$$

The notation $m_\mu^f(E)(x)$ is the measure of E according to an observer viewpoint when it look at x.

Example 1.1. Suppose $X = \Re\backslash\{0,-1\}$ and also let $f : X \to X$ be defined by $x \mapsto \frac{1}{x}$. Also let $E = [0,1]$ and $\mu : X \to [0,1]$ be defined by $x \mapsto \frac{x}{x+1}$. Then

$$\begin{aligned}
m_\mu^f(E)(3) &= \limsup \frac{1}{n} \sum_{i=0}^{n-1} \chi_E(f^i(3)) \cdot \mu(f^i(3)) \\
&= \limsup \frac{1}{n} \left(\chi_E(3) \cdot \frac{3}{3+1} + \chi_E(\frac{1}{3}) \cdot \frac{\frac{1}{3}}{\frac{1}{3}+1} + \chi_E(3) \cdot \frac{3}{3+1} + ... \right) \\
&= \limsup \frac{1}{n} \left(0 + \frac{1}{4} + 0 + \frac{1}{4} + ... \right) \\
&= \limsup \frac{1}{n} \left(\frac{n}{8} \right) \\
&= \frac{1}{8}.
\end{aligned}$$

In Section 2, we introduce the notion of relative entropy for infinite partitions.

2. Relative entropy for partitions

We assume that (X, m_μ^f) is a relative probability space.

A partition of X is a disjoin collection of elements of $P(X)$ whose union is X, where $P(X)$ is the power set of X.

If $\xi = \{A_i : i \in \mathbb{N}\}$, and $\eta = \{C_j : j \in \mathbb{N}\}$ are two countable partitions of X, then their join is the partition :

$$\xi \vee \eta = \{A_i \cap C_j : m_\mu^f(A_i \cap C_j)(x) \geq m_\mu^f(A_i)(x) m_\mu^f(C_j)(x), \forall x \in X\}.$$

If $T : X \to X$ is a mapping, then T is called relative probability measure-preserving if $m_\mu^f(T^{-1}E)(x) = m_\mu^f(E)(x)$, for all $x \in X$.

Suppose $\mathcal{C}, \mathcal{D}$ are two subsets of $P(X)$ then we say $\mathcal{C} \subset_\mu^f \mathcal{D}$ if for each $C \in \mathcal{C}$ there exist $D \in \mathcal{D}$ such that $m_\mu^f(D\Delta C)(x) = 0$.

We say $\mathcal{C} =_\mu^f \mathcal{D}$ if $\mathcal{C} \subset_\mu^f \mathcal{D}$ and $\mathcal{D} \subset_\mu^f \mathcal{C}$. Also we say $\mathcal{D}$ is a refinement of $\mathcal{C}$, and write $\mathcal{C} \preccurlyeq_\mu^f \mathcal{D}$, when we can write each element of $\mathcal{C}$, as union some elements of $\mathcal{D}$.

Definition 2.1. Let $\mathcal{A} = \{A_i : i \in \mathbb{N}\}$ be a countable partition of X, then the relative entropy of $\mathcal{A}$ is the mapping $H_\mu^f(\mathcal{A}) : X \to \mathbb{R}$ defined by

$$H_\mu^f(\mathcal{A})(x) = -\log \sup_{i\in\mathbb{N}} m_\mu^f(A_i)(x).$$

Example 2.2. Let $X=(0,1]$. If $A=\{A_i : i\in \mathbf{N}\}$ such that for each $i\in\mathbf{N}$,

$$A_i=\left(\frac{1}{i+1},\frac{1}{i}\right].$$

A is a countable partition of X, because $\bigcup_{i=1}^{\infty} A_i = X$ and elements of $\{A_i : i\in\mathbf{N}\}$ are disjoint.

Now, let $f: X\to X$ be defined by $x\mapsto \frac{1}{k}$, $k\geq 1$. Let $\mu : X\to[0,1]$ be defined by $x\mapsto x$. Then

$$\begin{aligned} m_\mu^f(A_j)(x) &= \limsup \frac{1}{n}\sum_{i=0}^{n-1}\chi_{A_j}(f^i(x))\mu_(f^i(x)) \\ &= \limsup \frac{1}{n}\left[\chi_{A_j}(x)\cdot x+\sum_{i=1}^{n-1}\chi_{A_j}\left(\frac{1}{k}\right)\cdot\left(\frac{1}{k}\right)\right]. \end{aligned}$$

So $m_\mu^f(A_j)(x)=0$ for $j\neq k$ and $m_\mu^f(A_j)(x)=\frac{1}{k}$ for $j=k$. Thus

$$H_\mu^f(\mathcal{A})(x) = -\log\sup_{i\in\mathbb{N}} m_\mu^f(A_i)(x) = -\log\left(\frac{1}{k}\right)=\log(k).$$

Remark 2.3. If $T: X\to X$ is a relative probability measure preserving map then $H_\mu^f(T^{-1}\mathcal{A})(x)=H_\mu^f(\mathcal{A})(x)$, for all $x\in X$.

Lemma 2.4. *If $\mathcal{A}=^f_\mu \mathcal{C}$, then $H_\mu^f(\mathcal{A})(x)=H_\mu^f(\mathcal{C})(x)$ for all $x\in X$.*

Proof. If $\mathcal{A}\subseteq^f_\mu\mathcal{C}$, then for each $A_i\in\mathcal{A}$ there exists $C_j\in\mathcal{C}$ such that $m_\mu^f(A_i\Delta C_j)(x)=0$ for all $x\in X$.

We have $A_i=(A_i\bigcap C_j)\bigcup(A_i-C_j)$. So

$$\begin{aligned} A_i\cup(C_j-A_i) &= (A_i\cap C_j)\cup(C_j-A_i)\cup(A_i-C_j) \\ &= (A_i\cap C_j)\cup(A_i\Delta C_j). \end{aligned}$$

Hence

$$m_\mu^f(A_i)(x)+m_\mu^f(C_j-A_i)(x)=m_\mu^f(A_i\cap C_j)(x).$$

Thus

$$m_\mu^f(A_i)(x)=m_\mu^f(A_i\cap C_j)(x)\leq m_\mu^f(C_j)(x),$$

because $0\leq m_\mu^f(C_j-A_i)(x)\leq m_\mu^f(A_i\Delta C_j)(x)=0$. So for each $x\in X$,

$$\{m_\mu^f(A_i)(x) : i\in\mathbb{N}\}\subseteq\{m_\mu^f(C_j)(x) : j\in\mathbb{N}\}.$$

Hence

$$\sup_{i\in\mathbb{N}} m_\mu^f(A_i)(x)\leq\sup_{j\in\mathbb{N}} m_\mu^f(C_j)(x).$$

Thus $H_\mu^f(\mathcal{A})(x)\geq H_\mu^f(\mathcal{C})(x)$. Now, if $\mathcal{C}\subseteq^f_\mu\mathcal{A}$, then by changing the role of A and C we imply $H_\mu^f(\mathcal{A})(x)\leq H_\mu^f(\mathcal{C})(x)$. So $H_\mu^f(\mathcal{A})(x)=H_\mu^f(\mathcal{C})(x)$. ∎

3. Relative conditional entropy

In this section, the relative conditional entropy for infinite partitions and its properties are studied. Suppose that $\mathcal{A}$ and $\mathcal{C}$ are two infinite partitions of X.

Definition 3.1. The relative entropy of $\mathcal{A}$ given $\mathcal{C}$ is the number:

$$H_\mu^f(\mathcal{A}/\mathcal{C})(x) = -\log \frac{\sup\limits_{i,j\in\mathbb{N}} m_\mu^f(A_i \cap C_j)(x)}{\sup\limits_{j\in\mathbb{N}} m_\mu^f(C_j)(x)}.$$

$H_\mu^f(\mathcal{A}/\mathcal{C})(x)$ measure the uncertainty of the outcome of $\mathcal{A}$ when the observer who looks to the orbit of x, knows the outcome of $\mathcal{C}$.

Theorem 3.2. *Let $\mathcal{D} = \{\emptyset, X\}$. Then*

$$H_\mu^f(\mathcal{A}/\mathcal{D})(x) = H_\mu^f(\mathcal{A})(x) - H_\mu^f(X)(x).$$

Proof.

$$\begin{aligned} H_\mu^f(\mathcal{A}/\mathcal{D})(x) &= -\log \frac{\sup\limits_{i\in\mathbb{N}} m_\mu^f(A_i \cap X)(x)}{m_\mu^f(X)(x)} \\ &= -\log \frac{\sup\limits_{i\in\mathbb{N}} m_\mu^f(A_i)(x)}{m_\mu^f(X)(x)} - \log \sup_{i\in\mathbb{N}} m_\mu^f(A_i)(x) + \log m_\mu^f(X)(x) \\ &= H_\mu^f(\mathcal{A})(x) - H_\mu^f(X)(x). \end{aligned}$$

■

Corollary 3.3. *If $\mathcal{D} = \{\emptyset, X\}$ and $m_\mu^f(X)(x) = 1$, then*

$$H_\mu^f(\mathcal{A}/\mathcal{D})(x) = H_\mu^f(\mathcal{A})(x).$$

Theorem 3.4.

1. *If $\mathcal{A} =_\mu^f \mathcal{D}$ then $H_\mu^f(\mathcal{A}/\mathcal{C})(x) = H_\mu^f(\mathcal{D}/\mathcal{C})(x)$.*
2. *If $\mathcal{C} =_\mu^f \mathcal{D}$ then $H_\mu^f(\mathcal{A}/\mathcal{C})(x) = H_\mu^f(\mathcal{A}/\mathcal{D})(x)$.*

Proof. 1. $\mathcal{A} =_\mu^f \mathcal{D}$ implies that $\mathcal{A} \vee \mathcal{C} =_\mu^f \mathcal{D} \vee \mathcal{C}$. So

$$\begin{aligned} H_\mu^f(\mathcal{A}/\mathcal{C})(x) &= H_\mu^f(\mathcal{A} \vee \mathcal{C})(x) - H_\mu^f(\mathcal{C})(x) \\ &= H_\mu^f(\mathcal{D} \vee \mathcal{C})(x) - H_\mu^f(\mathcal{C})(x) \\ &= H_\mu^f(\mathcal{D}/\mathcal{C})(x). \end{aligned}$$

2. $\mathcal{C} =_\mu^f \mathcal{D}$ implies that $\mathcal{A} \vee \mathcal{C} =_\mu^f \mathcal{A} \vee \mathcal{D}$. So

$$\begin{aligned} H_\mu^f(\mathcal{A}/\mathcal{C})(x) &= H_\mu^f(\mathcal{A} \vee \mathcal{C})(x) - H_\mu^f(\mathcal{C})(x) \\ &= H_\mu^f(\mathcal{A} \vee \mathcal{D})(x) - H_\mu^f(\mathcal{D})(x) \\ &= H_\mu^f(\mathcal{A}/\mathcal{D})(x). \end{aligned}$$

■

Theorem 3.5. *If (X, m_μ^f) is a relative probability space, and $\mathcal{A}, \mathcal{C}, \mathcal{D}$ are countable partitions of X, then*

$$H_\mu^f(\mathcal{A} \vee \mathcal{C}/\mathcal{D})(x) = H_\mu^f(\mathcal{A}/\mathcal{D})(x) + H_\mu^f(\mathcal{C}/\mathcal{A} \vee \mathcal{D})(x).$$

Proof. Let $\mathcal{A} = \{A_i\}, \mathcal{C} = \{C_j\}$,and $\mathcal{D} = \{D_k\}$, then

$$H_\mu^f(\mathcal{A} \vee \mathcal{C}/\mathcal{D})(x) = -\log \frac{\sup\limits_{i,j,k\in\mathbb{N}} m_\mu^f(A_i \cap C_j \cap D_k)(x)}{\sup\limits_{k\in\mathbb{N}} m_\mu^f(D_k)(x)}$$

$$= -\log \left(\frac{\sup\limits_{i,k\in\mathbb{N}} m_\mu^f(A_i \cap D_k)(x)}{\sup\limits_{k\in\mathbb{N}} m_\mu^f(D_k)(x)} \times \frac{\sup\limits_{i,j,k\in\mathbb{N}} m_\mu^f(A_i \cap C_j \cap D_k)(x)}{\sup\limits_{i,k\in\mathbb{N}} m_\mu^f(A_i \cap D_k)(x)} \right)$$

$$= H_\mu^f(\mathcal{A}/\mathcal{D})(x) + H_\mu^f(\mathcal{C}/\mathcal{A} \vee \mathcal{D})(x).$$ ■

Corollary 3.6. *Suppose (X, m_μ^f) is a relative probability space. If $\mathcal{A}, \mathcal{C}$ are infinite partitions of X then*

$$H_\mu^f(\mathcal{A} \vee \mathcal{C})(x) = H_\mu^f(\mathcal{A})(x) + H_\mu^f(\mathcal{C}/\mathcal{A})(x).$$

Proof. Put $\mathcal{D} = \{\emptyset, X\}$ in Theorem 3.5 and apply Theorem 3.2. ■

Theorem 3.7. *If $\mathcal{A} \preccurlyeq_\mu^f \mathcal{C}$ then $H_\mu^f(\mathcal{A})(x) \leq H_\mu^f(\mathcal{C})(x)$.*

Proof. Let $\mathcal{A} = \{A_i\}$,and $\mathcal{C} = \{C_j\}$. Since $\mathcal{A} \preccurlyeq_\mu^f \mathcal{C}$, we can write each element of $\mathcal{A}$, as union some elements of $\mathcal{C}$. So $\sup\limits_{i\in\mathbb{N}} m_\mu^f(A_i)(x) \leq \sup\limits_{j\in\mathbb{N}} m_\mu^f(C_j)(x)$. Thus $H_\mu^f(\mathcal{A})(x) \leq H_\mu^f(\mathcal{C})(x)$. ■

Corollary 3.8. *Let (X, m_μ^f) be a relative probability space,and let $\mathcal{A}, \mathcal{C}, \mathcal{D}$ be countable partitions of X. Moreover let $\mathcal{A} \preccurlyeq_\mu^f \mathcal{C}$. Then*

$$H_\mu^f(\mathcal{A}/\mathcal{D})(x) \leq H_\mu^f(\mathcal{C}/\mathcal{D})(x).$$

Proof. Since $\mathcal{A} \preccurlyeq_\mu^f \mathcal{C}$, we have $\mathcal{A} \vee \mathcal{D} \preccurlyeq_\mu^f \mathcal{C} \vee \mathcal{D}$. By corollary 3.6, we can write

$$\begin{aligned} H_\mu^f(\mathcal{C}/\mathcal{D})(x) &= H_\mu^f(\mathcal{C} \vee \mathcal{D})(x) - H_\mu^f(\mathcal{D})(x) \\ &\geq H_\mu^f(\mathcal{A} \vee \mathcal{D})(x) - H_\mu^f(\mathcal{D})(x) \\ &= H_\mu^f(\mathcal{A}/\mathcal{D})(x). \end{aligned}$$ ■

Corollary 3.9. *If (X, m_μ^f) is a relative probability space, and $\mathcal{A}, \mathcal{C}$ are countable partitions of X, then $H_\mu^f(\mathcal{A} \vee \mathcal{C})(x) \leq H_\mu^f(\mathcal{A})(x) + H_\mu^f(\mathcal{C})(x)$.*

Proof. By definition, we have

$$\begin{aligned} H_\mu^f(\mathcal{A} \vee \mathcal{C})(x) &= -\log \sup_{i,j\in\mathbb{N}} m_\mu^f(A_i \cap C_j)(x) \\ &\leq -\log \sup_{i\in\mathbb{N}} m_\mu^f(A_i)(x) - \log \sup_{j\in\mathbb{N}} m_\mu^f(C_j)(x) \\ &= H_\mu^f(\mathcal{A})(x) + H_\mu^f(\mathcal{C})(x). \end{aligned}$$

■

Corollary 3.10. *If (X, m_μ^f) is a relative probability space, T is a relative probability measure-preserving map, and $\mathcal{A}, \mathcal{C}$ are countable partitions of X then*

$$H_\mu^f(T^{-1}(\mathcal{A})/T^{-1}(\mathcal{C}))(x) = H_\mu^f(\mathcal{A}/\mathcal{D})(x).$$

Theorem 3.11. *Suppose $\mathcal{A}, \mathcal{C}$ are countable partitions of X and $\mathcal{C} \subset_\mu^f \mathcal{A}$, then for all $x \in X, H_\mu^f(\mathcal{A}/\mathcal{C})(x) = 0$.*

Proof. Since $\mathcal{C} \subset_\mu^f \mathcal{A}$, for each $j \in \mathbb{N}$ there exist $i \in \mathbb{N}$ such that

$$m_\mu^f(A_i \cap C_j)(x) = m_\mu^f(C_j)(x).$$

Hence

$$\{m_\mu^f(C_j)(x) : j \in \mathbb{N}\} \subseteq \{m_\mu^f(A_i \cap C_j)(x) : i, j \in \mathbb{N}\}.$$

So for each $x \in X$,

$$\sup_{j\in\mathbb{N}} m_\mu^f(C_j)(x) \leq \sup_{i,j\in\mathbb{N}} m_\mu^f(A_i \cap C_j)(x).$$

On the other hand,

$$\sup_{j\in\mathbb{N}} m_\mu^f(C_j)(x) \geq \sup_{i,j\in\mathbb{N}} m_\mu^f(A_i \cap C_j)(x).$$

Thus

$$H_\mu^f(\mathcal{A}/\mathcal{C})(x) = -\log \frac{\sup_{i,j\in\mathbb{N}} m_\mu^f(A_i \cap C_j)(x)}{\sup_{j\in\mathbb{N}} m_\mu^f(C_j)(x)} = 0. \qquad \blacksquare$$

Two countable partitions $\mathcal{A}$ and $\mathcal{C}$ are called independent if

$$m_\mu^f(A \cap C)(x) = m_\mu^f(A)(x) m_\mu^f(C)(x)$$

for all $A \in \mathcal{A}, C \in \mathcal{C}$, and $x \in X$.

Theorem 3.12. *Let $\mathcal{A}, \mathcal{C}$ be two countable partitions of (X, m_μ^f), and let $\mathcal{A}, \mathcal{C}$ be independent. Then*

$$H_\mu^f(\mathcal{A}/\mathcal{C})(x) = H_\mu^f(\mathcal{A})(x).$$

Proof. If $\mathcal{A}$ and $\mathcal{C}$ are independent then

$$\begin{aligned} H_\mu^f(\mathcal{A}/\mathcal{C})(x) &= -\log \frac{\sup_{i,j\in\mathbb{N}} m_\mu^f(A_i \bigcap C_j)(x)}{\sup_{j\in\mathbb{N}} m_\mu^f(C_j)(x)} \\ &= -\log \frac{\sup_{i,j\in\mathbb{N}} m_\mu^f(A_i)(x) m_\mu^f(C_j)(x)}{\sup_{j\in\mathbb{N}} m_\mu^f(C_j)(x)} \\ &= H_\mu^f(\mathcal{A})(x). \end{aligned}$$

$\blacksquare$

4. Conclusion

We extended the notions of relative entropy and relative conditional entropy to infinite partitions. Some properties about the notions with infinite partitions studied. Relative entropy of dynamical systems with infinite partitions can be a topic for further research.

References

[1] Asadian, M.H., Molaei, M.R., *Entropy from an observer viewpoint*, Sylwan, 158 (8) (2014), 298-306.

[2] Ebrahimzadeh, A., Ebrahimi, M., *The entropy of countable dynamical systems*, U.P.B. Sci. Bull., Series A, 76 (4) (2014), 107-114.

[3] Ebrahimzadeh, A., Ebrahimi, M., *Entropy of dynamical systems on weights of a graph*, Journal of Mahani Mathematical Research Center, 2 (2013), 53-63.

[4] Malziri, M, Molaei, M.R., *Relative metric space*, Hacettepe Journal of Mathematics, and Statistics, 40 (5) (2011), 703-709.

[5] Molaei, M.R., Ghazanfar, B., *Relative probability measure*, Fuzzy Sets, Rough Sets, Multivalued Operations and Applications, vol. 1, no. 1 (2008), 89-97.

[6] Molaei, M.R., *Observatinal modeling of topological spaces*, Chaos, Solitons and Fractals, 42 (2009), 615-619.

[7] Molaei, M.R., Hoseini Anvari, M.R., Haqiri, T., *On relative semi-dynamical systems*, Intelligent Automation and Soft Computing, vol. 13, no. 4 (2007), 405-413.

[8] Molaei, M.R., Molaei, H., *Relative topological pressure*, Mathematical Reports, vol. 12 (62), no. 1 (2010), 31-36.

[9] Molaei, M.R., *Relative semi-dynamical*, International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, vol. 12, No. 2 (2004), 237-243.

[10] Molaei, M.R., *Selective manifold*, Journal of Advanced Research in Applied Mathematices, 3 (2) (2011) 1-7.

[11] Molaei, M.R., Salili, SH., *A new approach to fuzzy topological spaces*, Hadronic Journal, 25 , (2002)81-90.

[12] Molaei, M.R., *The concept of synchronization from the observers viewpoint*, Cankaya University Journal of Science and Engineering, 8 (2) (2011), 255-262.

[13] Walters, P., *An Introduction to Ergodic Theory*, Springer-Verlag, 1982.

Accepted: 10.05.2015