数理解析研究所講究録
第 1561 巻 2007 年 152-165

Some results on Tsallis entropies in classical system*

Shigeru Furuichi†
Department of Electronics and Computer Science,
Tokyo University of Science, Onoda City, Yamaguchi, 756-0884, Japan

Abstract. In this survey, we review some theorems and properties of Tsallis entropies in classical system without proofs. See our previous papers [11, 8, 9, 10] for the proofs and details.

Keywords : Tsallis entropy, Tsallis relative entropy, uniqueness theorem, information theory, maximum Tsallis entropy principle, q-Fisher information and q-Cramér-Rao inequality

1 Tsallis entropies in classical system

First of all, we define the Tsallis entropy and the Tsallis relative entropy. We denote the q-logarithmic function by

$$\ln_q x \equiv \frac{x^{1-q}-1}{1-q} \qquad (q \in \mathbb{R}, q \neq 1, x > 0)$$

and the q-exponential function by

$$\exp_q(x) \equiv \begin{cases} (1+(1-q)x)^{\frac{1}{1-q}}, & if\ 1+(1-q)x > 0, \\ 0 & otherwise \end{cases} \qquad (q \in \mathbb{R}, q \neq 1, x \in \mathbb{R}).$$

For these functions, we have the following relations:

$$\ln_q(xy) = \ln_q x + \ln_q y + (1-q)\ln_q x \ln_q y, \quad \exp_q(x+y+(1-q)xy) = \exp_q(x)\exp_q(y)$$

and

$$\lim_{q\to 1} \ln_q x = \log x, \quad \lim_{q\to 1} \exp_q(x) = \exp(x).$$

By the use of q-logarithmic function, we define Tsallis entropy [27] by

$$S_q(A) = -\sum_{j=1}^{n} a_j^q \ln_q a_j, \quad (q \neq 1),$$

for a probability distribution $A = \{a_j\}$. After about one decade of discover of the Tsallis entropy, the Tsallis relative entropy was independently introduced in the following [28, 21, 19].

$$D_q(A|B) \equiv -\sum_{j=1}^{n} a_j \ln_q \frac{b_j}{a_j}, \quad (q \neq 1),$$

for two probability distributions $A = \{a_j\}$ and $B = \{b_j\}$.

Note that the Tsallis entropies are one parameter extensions of the Shannon entropy $S_1(A)$ and the relative entropy $D_1(A|B)$ [17, 16] respectively, in the sense that :

$$\lim_{q\to 1} S_q(A) = S_1(A) \equiv -\sum_{j=1}^{n} a_j \log a_j, \tag{1}$$

$$\lim_{q\to 1} D_q(A|B) = D_1(A|B) \equiv \sum_{j=1}^{n} a_j \log \frac{a_j}{b_j}. \tag{2}$$

*This work was partially supported by the Japanese Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Encouragement of Young scientists (B), 17740068.
†E-mail:furuichi@ed.yama.tus.ac.jp

In addition, the Tsallis entropies for $q \neq 1$ are non-additive entropies in the sense that :

$$S_q(A \times B) = S_q(A) + S_q(B) + (1-q)S_q(A)S_q(B) \tag{3}$$

$$D_q\left(A^{(1)} \times A^{(2)} \middle| B^{(1)} \times B^{(2)}\right) = D_q\left(A^{(1)} \middle| B^{(1)}\right) + D_q\left(A^{(2)} \middle| B^{(2)}\right) + (q-1)\, D_q\left(A^{(1)} \middle| B^{(1)}\right) D_q\left(A^{(2)} \middle| B^{(2)}\right), \tag{4}$$

where

$$A^{(1)} \times A^{(2)} = \left\{ a_j^{(1)} a_j^{(2)} \middle| a_j^{(1)} \in A^{(1)}, a_j^{(2)} \in A^{(2)} \right\}, B^{(1)} \times B^{(2)} = \left\{ b_j^{(1)} b_j^{(2)} \middle| b_j^{(1)} \in B^{(1)}, b_j^{(2)} \in B^{(2)} \right\}.$$

2 A uniqueness theorem of Tsallis relative entropy

A uniqueness theorem for Shannon entropy is fundamental theorem in information theory [20, 14, 15]. In this section, we review the uniqueness theorem of Tsallis relative entropy [8] which was derived by combining the Hobson's axiom [13] and Suyari's one [22].

Theorem 2.1 ([13]) We suppose the function $D_1(A|B)$ is defined for any pair of two probability distributions $A = \{a_j\}$ and $B = \{b_j\}$ for $j = 1, \cdots, n$. If $D_1(A|B)$ satisfies the following conditions, then it is necessary given by the form $k \sum_{j=1}^{n} a_j \log \frac{a_j}{b_j}$ with a positive constant k.

(H1) *Continuity*: $D_1(A|B)$ is a continuous function of its $2n$ variables.

(H2) *Symmetry* :

$$D_1(a_1, \cdots, a_j, \cdots, a_k, \cdots, a_n | b_1, \cdots, b_j, \cdots, b_k, \cdots, b_n) = D_1(a_1, \cdots, a_k, \cdots, a_j, \cdots, a_n | b_1, \cdots, b_k, \cdots, b_j, \cdots, b_n) \tag{5}$$

(H3) *Grouping axiom* :

$$D_1\left(a_{1,1}, \cdots, a_{1,m}, a_{2,1}, \cdots, a_{2,m} \middle| b_{1,1}, \cdots, b_{1,m}, b_{2,1}, \cdots, b_{2,m}\right) = D_1\left(c_1, c_2 \middle| d_1, d_2\right) + c_1 D_1\left(\frac{a_{1,1}}{c_1}, \cdots, \frac{a_{1,m}}{c_1} \middle| \frac{b_{1,1}}{d_1}, \cdots, \frac{b_{1,m}}{d_1}\right) + c_2 D_1\left(\frac{a_{2,1}}{c_2}, \cdots, \frac{a_{2,m}}{c_2} \middle| \frac{b_{2,1}}{d_2}, \cdots, \frac{b_{2,m}}{d_2}\right)$$

where $c_i = \sum_{j=1}^{m} a_{i,j}$ and $d_i = \sum_{j=1}^{m} b_{i,j}$.

(H4) $D_1(A|B) = 0$ if $a_j = b_j$ for all j.

(H5) $D_1(\frac{1}{n}, \cdots, \frac{1}{n}, 0, \cdots, 0 | \frac{1}{n_0}, \cdots, \frac{1}{n_0})$ is an increasing function of n_0 and a decreasing function of n, for any integers n, n_0 such that $n_0 \geq n$.

For the Tsallis relative entropy, it is known that the several fundamental properties, which are summarized in the below, hold as parametrically extensions of the relative entropy. For example, see [11].

Proposition 2.2 ([11])

(1) (Nonnegativity) $D_q(A|B) \geq 0$.

(2) (Symmetry) $D_q\left(a_{\pi(1)}, \cdots, a_{\pi(n)} \middle| b_{\pi(1)}, \cdots, b_{\pi(n)}\right) = D_q\left(a_1, \cdots, a_n \middle| b_1, \cdots, b_n\right)$.

(3) (Possibility of extention) $D_q\left(a_1, \cdots, a_n, 0 \middle| b_1, \cdots, b_n, 0\right) = D_q\left(a_1, \cdots, a_n \middle| b_1, \cdots, b_n\right)$.

(4) (Non-additivity) Eq.(4) holds.

(5) (Joint convexity) For $0 \leq \lambda \leq 1$, any $q \geq 0$ and the probability distributions $A^{(i)} = \left\{a_j^{(i)}\right\}$, $B^{(i)} = \left\{b_j^{(i)}\right\}$, $(i = 1, 2)$, we have

$$D_q\left(\lambda A^{(1)} + (1-\lambda) A^{(2)} | \lambda B^{(1)} + (1-\lambda) B^{(2)}\right) \leq \lambda D_q\left(A^{(1)} | B^{(1)}\right) + (1-\lambda) D_q\left(A^{(2)} | B^{(2)}\right).$$

(6) (Strong additivity)

$$\begin{aligned} &D_q\left(a_1,\cdots,a_{i-1},a_{i_1},a_{i_2},a_{i+1},\cdots,a_n\,|b_1,\cdots,b_{i-1},b_{i_1},b_{i_2},b_{i+1},\cdots,b_n\right) \\ &= D_q\left(a_1,\cdots,a_n\,|b_1,\cdots,b_n\right) + b_i^{1-q}a_i^q D_q\left(\tfrac{a_{i_1}}{a_i},\tfrac{a_{i_2}}{a_i}\,\middle|\,\tfrac{b_{i_1}}{b_i},\tfrac{b_{i_2}}{b_i}\right) \end{aligned}$$

where $a_i = a_{i_1} + a_{i_2}, b_i = b_{i_1} + b_{i_2}$.

(7) (Monotonicity) For the transition probability matrix W, we have

$$D_q(WA\,|WB) \le D_q(A\,|B).$$

Conversely, we axiomatically characterized the Tsallis relative entropy by some of these properties.

Theorem 2.3 ([8]) If the function $D_q(A|B)$, defined for any pairs of the probability distributions $A = \{a_i\}$ and $B = \{b_i\}$ on a finite probability space, satisfies the conditions (A1)-(A3) in the below, then $D_q(A|B)$ is necessary given by the form

$$D_q(A|B) = \frac{\sum_{j=1}^n \left(a_j - a_j^q b_j^{1-q}\right)}{\phi(q)} \tag{6}$$

with a certain function $\phi(q)$.

(A1) *Continuity* : $D_q(a_1,\cdots,a_n|b_1,\cdots,b_n)$ is a continuous function for $2n$ variables.

(A2) *Symmetry* :

$$\begin{aligned} &D_q(a_1,\cdots,a_j,\cdots,a_k,\cdots,a_n|b_1,\cdots,b_j,\cdots,b_k,\cdots,b_n) \\ &= D_q(a_1,\cdots,a_k,\cdots,a_j,\cdots,a_n|b_1,\cdots,b_k,\cdots,b_j,\cdots,b_n) \end{aligned} \tag{7}$$

(A3) *Generalized additivity* :

$$\begin{aligned} &D_q(a_{1,1},\cdots,a_{1,m},\cdots,a_{n,1},\cdots,a_{n,m}|b_{1,1},\cdots,b_{1,m},\cdots,b_{n,1},\cdots,b_{n,m}) \\ &= D_q(c_1,\cdots,c_n|d_1\cdots,d_n) + \sum_{i=1}^n c_i^q d_i^{1-q} D_q\left(\frac{a_{i,1}}{c_i},\dots,\frac{a_{i,m}}{c_i}|\frac{b_{i,1}}{d_i},\dots,\frac{b_{i,m}}{d_i}\right), \end{aligned} \tag{8}$$

where $c_i = \sum_{j=1}^m a_{i,j}$ and $d_i = \sum_{j=1}^m b_{i,j}$.

The function $\phi(q)$ was characterized in the following.

Proposition 2.4 ([8]) The property that Tsallis relative entropy is one parameter extension of relative entropy:

$$\lim_{q\to 1} D_q(A|B) = k\sum_{j=1}^n a_j \log\frac{a_j}{b_j} \tag{9}$$

characterize the function $\phi(q)$ such as

(c1) $\lim_{q\to 1}\phi(q) = 0$.

(c2) There exists an interval (a,b) such that $a < 1 < b$ and $\phi(q)$ is differentiable on the interval $(a,1) \subset (1,b)$.

(c3) There exists positive number k such taht $\lim_{q\to 1}\frac{d\phi(q)}{dq} = -\frac{1}{k}$.

Proposition 2.5 ([8]) The condition that

(A5) $D_q(A|U)$ takes the minimum value for fixed posterior probability distribution as uniform distribution $U = \{\frac{1}{n},\cdots,\frac{1}{n}\}$:

$$D_q(a_1,\cdots,a_n|\frac{1}{n},\cdots,\frac{1}{n}) \ge D_q(\frac{1}{n},\cdots,\frac{1}{n}|\frac{1}{n},\cdots,\frac{1}{n}),$$

implies

(c4) $\phi(q)(1-q) > 0$ for $q \neq 1$.

As a simple example of $\phi(q)$ to satisfy the above four conditions from (c1) to (c4), we may take $\phi(q) = 1 - q$ and $k = 1$. Then we can obtain the Tsallis relative entropy.

Finally, we give a few remarks on the conditions of our axiom in the following two propositions.

Proposition 2.6 ([8]) The following conditions (A3') and (A4) imply the condition (A3) in Theorem 2.3.

(A3') *Generalized grouping axiom:* The following additivity holds.

$$D_q\left(a_{1,1},\cdots,a_{1,m},a_{2,1},\cdots,a_{2,m}\,|b_{1,1},\cdots,b_{1,m},b_{2,1},\cdots,b_{2,m}\right) = D_q\left(c_1,c_2\,|d_1,d_2\right)$$
$$+c_1^q d_1^{1-q} D_q\left(\frac{a_{1,1}}{c_1},\cdots,\frac{a_{1,m}}{c_1}\,\middle|\,\frac{b_{1,1}}{d_1},\cdots,\frac{b_{1,m}}{d_1}\right) + c_2^q d_2^{1-q} D_q\left(\frac{a_{2,1}}{c_2},\cdots,\frac{a_{2,m}}{c_2}\,\middle|\,\frac{b_{2,1}}{d_2},\cdots,\frac{b_{2,m}}{d_2}\right)$$

where $c_i = \sum_{j=1}^m a_{i,j}$ and $d_i = \sum_{j=1}^m b_{i,j}$.

(A4) $D_q(A|B) = 0$ if $a_j = b_j$ for all j.

Proposition 2.7 ([8]) The conditions (A3') in the above Proposition 2.6 and the following condition (A4') imply the condition (A3) Theorem 2.3.

(A4') *Expandability* :

$$D_q(a_1,\cdots,a_n,0|b_1,\cdots,b_n,0) = D_q(a_1,\cdots,a_n|b_1,\cdots,b_n) \tag{10}$$

Proposition 2.6 and Proposition 2.7 tell us that we may use the axiom composed from the set of [**(A1)**,**(A2)**,**(A3')** and **(A4)**] or [**(A1)**,**(A2)**,**(A3')** and **(A4')**] instead of the set of [**(A1)**,**(A2)** and **(A3)**] in Theorem 2.3.

3 A uniqueness theorem of Tsallis entropy

In this section, we review the uniqueness theorem of Tsallis entropy. We proved that the uniqueness theorem for the Tsallis entropy by introducing the generalized Faddeev's axiom is proven [8].

We suppose that the function $S_q(x_1,\cdots,x_n)$ is defined for the n-tuple $(x_1,\cdots,x_n)$ belonging to $\Delta_n \equiv \{(p_1,\cdots,p_n)|\sum_{i=1}^n p_i = 1, p_i \geq 0\ (i=1,2,\cdots,n)\}$ and takes values in $\mathbb{R}^+ \equiv [0,\infty)$. In order to characterize the function $S_q(x_1,\cdots,x_n)$, we introduce the following axiom which is a slight generalization of Faddeev's axiom.

Axiom 3.1 (Generalized Faddeev's axiom :[8])

(GF1) *Continuity:* The function $f_q(x) \equiv S_q(x, 1-x)$ with a parameter $q \geq 0$ is continuous on the closed interval $[0,1]$ and $f_q(x_0) > 0$ for some $x_0 \in [0,1]$.

(GF2) *Symmetry* : For arbitrary permutation $\{x'_k\} \in \Delta_n$ of $\{x_k\} \in \Delta_n$,

$$S_q(x_1,\cdots,x_n) = S_q(x'_1,\cdots,x'_n). \tag{11}$$

(GF3) *Generalized additivity* : For $x_n = y + z$, $y \geq 0$ and $z > 0$,

$$S_q(x_1,\cdots,x_{n-1},y,z) = S_q(x_1,\cdots,x_n) + x_n^q S_q\left(\frac{y}{x_n},\frac{z}{x_n}\right). \tag{12}$$

The conditions (GF1) and (GF2) are just same with the original Faddeev's conditions except for the addition of the parameter q. The condition (GF3) is a generalization of the original Faddeev's additivity condition in the sense that our condition (GF3) uses the x_n^q as the factor of the second term in the right hand side, while original condition uses x_n itself as the factor of that. It is notable that our condition (GF3) is a simplification of the condition [GSK3] in the paper [22], since our condition (GF3) does not have to take the summation on i from 1 to n. Moreover our axiom does not need the maximality condition [GSK2] in [22]. In such viewpoints, our axiom improves the generalized Shannon-Khinchin's axiom in [22]. For the above generalized Faddeev's axiom, we have the following uniqueness theorem for Tsallis entropy.

Theorem 3.2 ([8]) Three conditions (GF1),(GF2) and (GF3) uniquely give the form of the function $S_q : \Delta_n \to \mathbb{R}^+$ such that

$$S_q(x_1, \cdots, x_n) = -\lambda_q \sum_{i=1}^{n} x_i^q \ln_q x_i, \tag{13}$$

where λ_q is a positive constant number depending on the parameter $q \geq 0$.

In the rest of this subsection, we study the relation between the generalized Shannon-Khinchin's axiom introduced in [22] and the generalized Faddeev's axiom presented in the previous section. To do so, we review the generalized Shannon-Khinchin's axiom in the following.

Axiom 3.3 (Generalized Shannon-Khinchin's axiom : [22])

(GSK1) *Continuity*: The function $S_q : \Delta_n \to \mathbb{R}^+$ is continuous.

(GSK2) *Maximality*: $S_q(\frac{1}{n}, \cdots, \frac{1}{n}) = max\{S_q(X) : x_i \in \Delta_n\} > 0$.

(GSK3) *Generalized Shannon additivity*: For $x_{ij} \geq 0$, $x_i = \sum_{j=1}^{m_i} x_{ij}$, $(i = 1, \cdots, n; j = 1, \cdots, m_i)$,

$$S_q(x_{11}, \cdots, x_{nm_n}) = S_q(x_1, \cdots, x_n) + \sum_{i=1}^{n} x_i^q S_q\left(\frac{x_{i1}}{x_i}, \cdots, \frac{x_{im_i}}{x_i}\right).$$

(GSK4) *Expandability*: $S_q(x_1, \cdots, x_n, 0) = S_q(x_1, \cdots, x_n)$.

We should note that the above condition (GSK4) is slightly changed from [GSK4] of the original axiom in [22]. Then we have the following proposition.

Proposition 3.4 ([8]) Axiom 3.3 implies Axiom 3.1.

We also have the following proposition.

Proposition 3.5 ([8]) $S_q(X) = -\lambda_q \sum_{i=1}^{n} x_i^q \ln_q x_i$ satisfies Axiom 3.3.

From Theorem 3.2, Proposition 3.4 and Proposition 3.5, we have the following equivalent relation among Axiom 3.1, Axiom 3.3 and the Tsallis entropy.

Theorem 3.6 ([8]) The following three statements are equivalent to one another.

(1) $S_q : \Delta_n \to \mathbb{R}^+$ satisfies Axiom 3.3

(2) $S_q : \Delta_n \to \mathbb{R}^+$ satisfies Axiom 3.1

(3) For $(x_1, \cdots, x_n) \in \Delta_n$, there exists $\lambda_q > 0$ such that

$$S_q(x_1, \cdots, x_n) = -\lambda_q \sum_{i=1}^{n} x_i^q \ln_q x_i.$$

4 Some properties of Tsallis entropies

In this section, we review some information-theoretical properties on the Tsallis entropies. We define the Tsallis conditional entropy and the Tsallis joint entropy in the following.

Definition 4.1 ([9]) For the conditional probability $p(x_i|y_j) \equiv p(X = x_i|Y = y_j)$ and the joint probability $p(x_i, y_j) \equiv p(X = x_i, Y = y_j)$, we define Tsallis conditional entropy and Tsallis joint entropy by

$$S_q(X|Y) \equiv -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)^q \ln_q p(x_i|y_j), \quad (q \neq 1), \tag{14}$$

and

$$S_q(X,Y) \equiv -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)^q \ln_q p(x_i, y_j), \quad (q \neq 1). \tag{15}$$

We note that the above definitions were essentially introduced in [5, 3] by

$$H_\beta(X,Y) \equiv \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\left(p(x_i,y_j)^\beta - p(x_i,y_j)\right)}{2^{1-\beta}-1}, \quad (\beta > 0, \beta \neq 1)$$
$$H_\beta(X|Y) \equiv \sum_{i=1}^{n}\sum_{j=1}^{m} p(y_j)^\beta H_\beta(X|y_j), \quad (\beta > 0, \beta \neq 1)$$

except for the difference of the multiplicative function. And then a chain rule and a subadditivity:

$$H_\beta(X,Y) = H_\beta(X) + H_\beta(Y|X),$$
$$H_\beta(Y|X) \leq H_\beta(Y), \quad \beta > 1,$$

were shown in Theorem 8 of [5].

It is important to study so-called a chain rule which gives the relation between a conditional entropy and a joint entropy in not only information theory [4] but also statistical physics. For these Tsallis entropies, the following chain rule holds as similar as the chain rule holds for the joint entropy of type β and the conditional entropy of type β.

Proposition 4.2 ([5])

$$S_q(X,Y) = S_q(X) + S_q(Y|X). \tag{16}$$

(Therefore immediately $S_q(X) \leq S_q(X,Y)$.)

As a corollary of the above Proposition 4.2, we have the following lemma.

Lemma 4.3 The following chain rules hold.

(1) $S_q(X,Y,Z) = S_q(X,Y|Z) + S_q(Z)$.

(2) $S_q(X,Y|Z) = S_q(X|Z) + S_q(Y|X,Z)$.

From the non-additivity Eq.(3), for $q \geq 1$ and two independent random variables X and Y, the subadditivity holds:

$$S_q(X \times Y) \leq S_q(X) + S_q(Y).$$

It is known that the subadditivity for general random variables X and Y holds in the case of $q \geq 1$, thanks to the following proposition.

Proposition 4.4 ([5]) The following inequality holds for two random variables X and Y, and $q \geq 1$,

$$S_q(X|Y) \leq S_q(X), \tag{17}$$

with equality if and only if $q = 1$ and $p(x_i|y_j) = p(x_i)$ for all $i = 1, \cdots, n$ and $j = 1, \cdots, m$.

Eq.(17) and Eq.(16) imply the subadditivity of Tsallis entropies.

Theorem 4.5 ([5]) For $q \geq 1$, we have

$$S_q(X,Y) \leq S_q(X) + S_q(Y). \tag{18}$$

On the other hand, we easily find that for two independent random variables X and Y, and $0 \leq q < 1$, the superadditivity holds:

$$S_q(X \times Y) \geq S_q(X) + S_q(Y).$$

However, in general the superadditivity for two correlated random variables X and Y, and $0 \leq q < 1$ does not hold. Because we can show many counterexamples. For example, we consider the following joint distribution of X and Y.

$$p(x_1,y_1) = p(1-x), p(x_1,y_2) = (1-p)y, p(x_2,y_1) = px, p(x_2,y_2) = (1-p)(1-y), \tag{19}$$

where $0 \leq p, x, y \leq 1$. Then each marginal distribution can be computed by

$$p(x_1) = p(1-x) + (1-p)y, p(x_2) = px + (1-p)(1-y), p(y_1) = p, p(y_2) = 1-p. \tag{20}$$

In general, we clearly see X and Y are not independent each other for the above example. Then the value of $\Delta \equiv S_q(X,Y) - S_q(X) - S_q(Y)$ takes both positive and negative so that there does not exist the complete ordering between $S_q(X,Y)$ and $S_q(X) + S_q(Y)$ for correlated X and Y in the case of $0 \leq q < 1$. Indeed, $\Delta = -0.287089$ when $q = 0.8, p = 0.6, x = 0.1, y = 0.1$, while $\Delta = 0.0562961$ when $q = 0.8, p = 0.6, x = 0.1, y = 0.9$.

We also have the strong subadditivity holds in the case of $q \geq 1$.

Theorem 4.6 ([9]) For $q \geq 1$, the strong subadditivity

$$S_q(X,Y,Z) + S_q(Z) \leq S_q(X,Z) + S_q(Y,Z) \tag{21}$$

holds with equality if and only if $q = 1$ and, random variables X and Y are independent for a given random variable Z.

Theorem 4.7 ([9]) Let $X_1, \cdots, X_{n+1}$ be the random variables. For $q > 1$, we have

$$S_q(X_{n+1}|X_1, \cdots, X_n) \leq S_q(X_{n+1}|X_2, \cdots, X_n). \tag{22}$$

The subadditivity for Tsallis entropies conditioned by Z holds.

Proposition 4.8 ([9]) For $q \geq 1$, we have

$$S_q(X,Y|Z) \leq S_q(X|Z) + S_q(Y|Z). \tag{23}$$

Proposition 4.8 can be generalized in the following.

Theorem 4.9 ([9]) For $q \geq 1$, we have

$$S_q(X_1, \cdots, X_n|Z) \leq S_q(X_1|Z) + \cdots + S_q(X_n|Z). \tag{24}$$

In addition, we have the following propositions.

Proposition 4.10 ([9]) For $q \geq 1$, we have

$$2S_q(X,Y,Z) \leq S_q(X,Y) + S_q(Y,Z) + S_q(Z,X).$$

Proposition 4.11 ([9]) For $q > 1$, we have

$$S_q(X_n|X_1) \leq S_q(X_2|X_1) + \cdots + S_q(X_n|X_{n-1}).$$

For normalized Tsallis entropies, the mutual information was defined in [31] with the assumption of its non-negativity. We define the Tsallis mutual entropy in terms of the original (*not* normalized) Tsallis type entropies. The inequality Eq.(17) naturally leads us to define Tsallis mutual entropy without the assumption of its non-negativity.

Definition 4.12 ([9]) For two random variables X and Y, and $q > 1$, we define the Tsallis mutual entropy as the difference between Tsallis entropy and Tsallis conditional entropy such that

$$I_q(X;Y) \equiv S_q(X) - S_q(X|Y). \tag{25}$$

Note that we never use the term *mutual information* but use mutual entropy through this paper, since a proper evidence of channel coding theorem for information transmission has not ever been shown in the context of Tsallis statistics. From Eq.(16), Eq.(18) and Eq.(17), we easily find that $I_q(X;Y)$ has the following fundamental properties.

Proposition 4.13 ([9])

(1) $0 \leq I_q(X;Y) \leq min\{S_q(X), S_q(Y)\}$.

(2) $I_q(X;Y) = S_q(X) + S_q(Y) - S_q(X,Y) = I_q(Y;X)$.

Note that we have

$$S_q(X) \leq S_q(Y) \Longleftrightarrow S_q(X|Y) \leq S_q(Y|X) \tag{26}$$

from the symmetry of Tsallis mutual entropy. We also define the Tsallis conditional mutual entropy

$$I_q(X;Y|Z) \equiv S_q(X|Z) - S_q(X|Y,Z) \tag{27}$$

for three random variables X, Y and Z, and $q > 1$. In addition, $I_q(X;Y|Z)$ is nonnegative. For these quantities, we have the following chain rules.

Theorem 4.14 ([9])

(1) For three random variables X, Y and Z, and $q > 1$, the chain rule holds:

$$I_q(X;Y,Z) = I_q(X;Z) + I_q(X;Y|Z). \tag{28}$$

(2) For random variables $X_1, \cdots, X_n$ and Y, the chain rule holds:

$$I_q(X_1,\cdots,X_n;Y) = \sum_{i=1}^{n} I_q(X_i;Y|X_1,\cdots,X_{i-1}). \tag{29}$$

We have the following inequality for Tsallis mutual entropies by the strong subadditivity.

Proposition 4.15 ([9]) For $q > 1$, we have

$$I_q(X;Z) \leq I_q(X,Y;Z).$$

5 Maximum Tsallis entropy principle

Here we discuss the maximum entropy principle which is one of most important theorem in entropy theory and statistical physics. We give a new proof of the theorems on the maximum entropy principle in Tsallis statistics. That is, we show that the q-canonical distribution attains the maximum value of the Tsallis entropy, subject to the constraint on the q-expectation value and the q-Gaussian distribution attains the maximum value of the Tsallis entropy, subject to the constraint on the q-variance, as applications of the non-negativity of the Tsallis relative entropy, without using the Lagrange multipliers method.

The set of all probability density function on $\mathbb{R}$ is represented by

$$D_{cl} \equiv \left\{ f : \mathbb{R} \to \mathbb{R} : f(x) \geq 0, \int_{-\infty}^{\infty} f(x)dx = 1 \right\}.$$

In the classical continuous system, Tsallis entropy [27] is then defined by

$$H_q(\phi(x)) \equiv -\int_{-\infty}^{\infty} \phi(x)^q \ln_q \phi(x) dx \tag{30}$$

for any nonnegative real number q and a probability distribution function $\phi(x) \in D_{cl}$. In addition, the Tsallis relative entropy is defined by

$$D_q(\phi(x)|\psi(x)) \equiv \int_{-\infty}^{\infty} \phi(x)^q (\ln_q \phi(x) - \ln_q \psi(x))dx \tag{31}$$

for any nonnegative real number q and two probability distribution functions $\phi(x) \in D_{cl}$ and $\psi(x) \in D_{cl}$. Taking the limit as $q \to 1$, the Tsallis entropy and the Tsallis relative entropy converge to the Shannon entropy $H_1(\phi(x)) \equiv -\int_{-\infty}^{\infty} \phi(x) \log \phi(x)$ and the Kullback-Leibler divergence $D_1(\phi(x)|\psi(x)) \equiv \int_{-\infty}^{\infty} \phi(x)(\log \phi(x) - \log \psi(x))dx$.

We define two sets involving the constraints on the q-expectation and the q-variance:

$$C_q^{(c)} \equiv \left\{ f \in D_{cl} : \frac{1}{c_q} \int_{-\infty}^{\infty} x f(x)^q dx = \mu_q \right\}$$

and

$$C_q^{(g)} \equiv \left\{ f \in C_q^{(c)} : \frac{1}{c_q} \int_{-\infty}^{\infty} (x - \mu_q)^2 f(x)^q dx = \sigma_q^2 \right\}.$$

Then the q-canonical distribution $\phi_q^{(c)}(x) \in D_{cl}$ and the q-Gaussian distribution $\phi_q^{(g)}(x) \in D_{cl}$ were formulated [18, 30, 2, 1, 22, 25, 29] by

$$\phi_q^{(c)}(x) \equiv \frac{1}{Z_q^{(c)}} \exp_q \left\{ -\beta_q^{(c)}(x - \mu_q) \right\}, \quad Z_q^{(c)} \equiv \int_{-\infty}^{\infty} \exp_q \left\{ -\beta_q^{(c)}(x - \mu_q) \right\}$$

and

$$\phi_q^{(g)}(x) \equiv \frac{1}{Z_q^{(g)}} \exp_q \left\{ -\frac{\beta_q^{(g)}(x - \mu_q)^2}{\sigma_q^2} \right\}, \quad Z_q^{(g)} \equiv \int_{-\infty}^{\infty} \exp_q \left\{ -\frac{\beta_q^{(g)}(x - \mu_q)^2}{\sigma_q^2} \right\},$$

respectively.

Here, we revisit the maximum entropy principle in non-additive statistical physics. The maximum entropy principles in Tsallis statistics have been studied and modified in many literatures [18, 30, 2, 1, 23]. Here we prove two theorems that maximize the Tsallis entropy under two different constraints by the use of the non-negativity of the Tsallis relative entropy instead of the use of the Lagrange multipliers method.

Lemma 5.1 For $q \neq 1$, we have

$$D_q(\phi(x)|\psi(x)) \geq 0,$$

with equality if and only if $\phi(x) = \psi(x)$ for all x.

Theorem 5.2 ([10]) If $\phi \in C_q^{(c)}$, then

$$H_q(\phi(x)) \leq -c_q \ln_q \frac{1}{Z_q^{(c)}},$$

with equality if and only if

$$\phi(x) = \frac{1}{Z_q^{(c)}} \exp_q \left\{ -\beta_q^{(c)}(x - \mu_q) \right\},$$

where $Z_q^{(c)} \equiv \int_{-\infty}^{\infty} \exp_q \left\{ -\beta_q^{(c)}(x - \mu_q) \right\} dx$ and $c_q \equiv \int_{-\infty}^{\infty} \phi(x)^q dx$.

Corollary 5.3 If $\phi \in C_q^{(c)}$, then $H_q(\phi(x)) \leq \log Z_1^{(c)}$ with equality if and only if

$$\phi(x) = \frac{1}{Z_1^{(c)}} \exp \left\{ -\beta_1^{(c)}(x - \mu) \right\}.$$

By the condition on the existence of q-variance σ_q (i.e., the convergence condition of the integral $\int x^2 \exp_q(-x^2)dx$), we consider q such that $0 < q < 3, q \neq 1$.

Theorem 5.4 ([10]) If $\phi \in C_q^{(g)}$ for q such that $0 < q < 3, q \neq 1$, then

$$H_q(\phi(x)) \leq -c_q \ln_q \frac{1}{Z_q^{(g)}} + \beta_q^{(g)} c_q Z_q^{(g)^{q-1}},$$

with equality if and only if

$$\phi(x) = \frac{1}{Z_q^{(g)}} \exp_q \left\{ -\beta_q^{(g)} (x - \mu_q)^2 / \sigma_q^2 \right\},$$

where $Z_q^{(g)} \equiv \int_{-\infty}^{\infty} \exp_q \left\{ -\beta_q^{(g)} (x - \mu_q)^2 / \sigma_q^2 \right\} dx$ with $\beta_q^{(g)} = 1/(3-q)$.

Corollary 5.5 If $\phi \in C_q^{(g)}$, then $H_1(\phi(x)) \leq \log \sqrt{2\pi e}\sigma$ with equality if and only if

$$\phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}.$$

The previous theorem and the fact that the Gaussian distribution minimizes the Fisher information leads us to study the Tsallis distribution (q-Gaussian distribution) minimizes the q-Fisher information as a parametric extension. To this end, we prepare some definitions. That is, we define the q-Fisher information and then prove the q-Cramér-Rao inequality which implies the q-Gaussian distribution with special q-variances attains the minimum value of the q-Fisher information.

In what follows, we abbreviate β_q and Z_q instead of $\beta_q^{(g)}$ and $Z_q^{(g)}$, respectively.

Definition 5.6 ([10]) For the random variable X with the probability density function $f(x)$, we define the q-score function $s_q(x)$ and q-Fisher information $J_q(X)$ by

$$s_q(x) \equiv \frac{d \ln_q f(x)}{dx}, \tag{32}$$

$$J_q(X) \equiv E_q \left[s_q(x)^2 \right], \tag{33}$$

where q-expectation E_q is defined by $E_q(X) \equiv \frac{\int x f(x)^q dx}{\int f(x)^q dx}$.

Example 5.7 For the random variable G obeying to q-Gaussian distribution

$$p_{q-G}(x) \equiv \frac{1}{Z_q} \exp_q \left\{ -\frac{\beta_q (x-\mu_q)^2}{\sigma_q^2} \right\},$$

where $\beta_q \equiv \frac{1}{3-q}$ and q-partition function $Z_q \equiv \int \exp_q \left\{ -\frac{\beta_q (x-\mu_q)^2}{\sigma_q^2} \right\} dx$, q-score function is calculated as

$$s_q(x) = -\frac{2\beta_q Z_q^{q-1}}{\sigma_q^2} (x - \mu_q).$$

Thus we can calculate q-Fisher information as

$$J_q(G) = \frac{4\beta_q^2 Z_q^{2q-2}}{\sigma_q^2}. \tag{34}$$

Note that

$$\lim_{q \to 1} J_q(G) = \frac{1}{\sigma_1^2}. \tag{35}$$

Theorem 5.8 ([9]) For any $q \in [0,1) \cup (1,3]$, we have the following statement.

(I) Given the random variable X with the probability density function $p(x)$, the q-expectation value $\mu_q \equiv E_q[X]$ and q-variance $\sigma_q^2 \equiv E_q\left[(X-\mu_q)^2\right]$, we have the inequality :

$$J_q(X) \geq \frac{1}{\sigma_q^2} \left(\frac{2}{\int p(x)^q dx} - 1 \right). \tag{36}$$

(II) We have the inequality

$$\frac{4\beta_q^2 Z_q^{2q-2}}{\sigma_q^2} \geq \frac{1}{\sigma_q^2}\left(\frac{2}{\int p_{q-G}(x)^q dx} - 1\right), \tag{37}$$

with equality if

$$\sigma_q = \frac{2^{\frac{q}{1-q}} (3-q)^{\frac{q+1}{2(q-1)}} (1-q)^{\frac{1}{2}}}{B\left(\frac{1}{2}, \frac{1}{1-q}\right)}, \quad (0 < q < 1) \tag{38}$$

or

$$\sigma_q = \frac{2^{\frac{1}{1-q}} (3-q)^{\frac{3-q}{2(q-1)}} (q-1)^{\frac{1}{2}}}{B\left(\frac{1}{q-1} - \frac{1}{2}, \frac{1}{2}\right)}, \quad (1 < q < 3) \tag{39}$$

6 Conclusion, remarks and discussions

As we have seen, we have characterized the Tsallis relative entropy by the parametrically extended conditions of the axiom formulated by A.Hobson [13]. This means that our theorem is a generalization of Hobson's one. Our result also includes the uniqueness theorem proven by H.Suyari [22] as a special case, in the sense that the choice of a trivial distribution for $B = \{b_j\}$ of the Tsallis relative entropy produces the essential form of the Tsallis entropy. However we should give a comment that our theorem require the symmetry (A2), although Suyari's one not so.

In addition, the Tsallis entropy was characterized by the generalized Faddeev's axiom which is a simplification of the generalized Shannon-Khinchin's axiom introduced in [22]. And then we slightly improved the uniqueness theorem proved in [22], by introducing the generalized Faddeev's axiom. Simultaneously, our result gives a generalization of the uniqueness theorem for Shannon entropy by means of Faddeev's axiom [7, 26].

Furthermore, we have proved the chain rules and the subadditivity for Tsallis entropies. Thus we could give important results for the Tsallis entropies in the case of $q \geq 1$ from the information theoretical point of view.

Finally, we derived the maximum entropy principle for the Tsallis entropy by applying the non-negativity of the Tsallis relative entropy. Also we introduced the q-Fisher information and then derived q-Cramér-Rao inequality.

In the following subsections, we give some remarks and discussions on the Tsallis entropies and related topics.

6.1 Inequalities on non-additivity

The non-additivity Eq.(3) for independent random variables X and Y gives rise to the mathematical interest whether we have the complete ordering between the left hand side and the right hand side in Eq.(3) for two general random variables X and Y. Such a kind of problem was taken in the paper [6] for the normalized Tsallis type entropies which are different from the definitions of the Tsallis type entropies in the present paper. However, its inequality appeared in (3.5) of the paper [6] was not true as we found the counter example in [24].

Unfortunately, in the present case, we also find the counter example for the inequalities between $S_q(X,Y)$ and $S_q(X) + S_q(Y) + (1-q)S_q(X)S_q(Y)$. In the same setting of Eq.(19) and Eq.(20), $\delta \equiv S_q(X,Y) - \{S_q(X) + S_q(Y) + (1-q)S_q(X)S_q(Y)\}$ takes both positive and negative values for both cases $0 \leq q < 1$ and $q > 1$. Indeed, $\delta = 0.00846651$ when $q = 1.8, p = 0.1, x = 0.1, y = 0.8$, while $\delta = -0.0118812$ when $q = 1.8, p = 0.1, x = 0.8, y = 0.1$. Also, $\delta = 0.00399069$ when $q = 0.8, p = 0.1, x = 0.8, y = 0.1$, while $\delta = -0.0128179$ when $q = 0.8, p = 0.1, x = 0.1, y = 0.8$.

Therefore there does not exist the complete ordering between $S_q(X,Y)$ and $S_q(X) + S_q(Y) + (1-q)S_q(X)S_q(Y)$ for both cases $0 \leq q < 1$ and $q > 1$.

6.2 A remarkable inequality derived from subadditivity for Tsallis entropies

From Eq.(18), we have the following inequality

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{m}p(x_i,y_j)\right)^r+\sum_{j=1}^{m}\left(\sum_{i=1}^{n}p(x_i,y_j)\right)^r\leq\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_i,y_j)^r+\left(\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_i,y_j)\right)^r \tag{40}$$

for $r\geq 1$ and $p(x_i,y_j)$ satisfying $0\leq p(x_i,y_j)\leq 1$ and $\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_i,y_j)=1$. By putting $p(x_i,y_j)=\frac{a_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}}$ in Eq.(40), we have the following inequality as a corollary of Theorem 4.5.

Corollary 6.1 For $r\geq 1$ and $a_{ij}\geq 0$,

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{m}a_{ij}\right)^r+\sum_{j=1}^{m}\left(\sum_{i=1}^{n}a_{ij}\right)^r\leq\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^r+\left(\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}\right)^r. \tag{41}$$

It is remarkable that the following inequality holds [12]

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{m}a_{ij}\right)^r\geq\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^r \tag{42}$$

for $r\geq 1$ and $a_{ij}\geq 0$.

6.3 Difference between Tsallis entropy and Shannon entropy

We point out on the difference between Tsallis entropy and Shannon entropy from the viewpoint of mutual entropy. In the case of $q=1$, the relative entropy between the joint probability $p(x_i,y_j)$ and the direct probability $p(x_i)p(y_j)$ is equal to the mutual entropy:

$$D_1((X,Y)|X\times Y)=S_1(X)-S_1(X|Y).$$

However, in the general case ($q\neq 1$), there exists the following relation:

$$\begin{aligned}&D_q((X,Y)|X\times Y)=S_q(X)-S_q(X|Y)\\&+\sum_{i,j}p(x_i,y_j)\left\{p(x_i)^{q-1}\ln_q p(x_i)+p(y_j)^{q-1}\ln_q p(y_j)-p(x_i,y_j)^{q-1}\ln_q p(x_i)p(y_j)\right\},\end{aligned} \tag{43}$$

which gives the crucial difference between the special case ($q=1$) and the general case ($q\neq 1$). The third term of the right hand side in the above equation Eq.(43) vanishes if $q=1$. The existence of the third term of Eq.(43) means that we have two possibilities of the definition of Tsallis mutual entropy, that is, $I_q(X;Y)\equiv S_q(X)-S_q(X|Y)$ or $I_q(X;Y)\equiv D_q((X,Y)|X\times Y)$. We have adopted the former definition in the present paper, along with the definition of the capacity in the origin of information theory by Shannon [20].

6.4 Another candidate of Tsallis conditional entropy

It is remarkable that Tsallis entropy $S_q(X)$ can be regarded as the expected value of $\ln_q\frac{1}{p(x_i)}$, that is, since $\ln_q(x)=-x^{1-q}\ln_q(1/x)$, it is expressed by

$$S_q(X)=\sum_{i=1}^{n}p(x_i)\ln_q\frac{1}{p(x_i)},\quad(q\neq 1), \tag{44}$$

where the convention $0\ln_q(\cdot)=0$ is set. Along with the view of Eq.(44), we may define Tsallis conditional entropy and Tsallis joint entropy in the following.

Definition 6.2 For the conditional probability $p(x_i|y_j)$ and the joint probability $p(x_i, y_j)$, we define Tsallis conditional entropy and Tsallis joint entropy by

$$\hat{S}_q(X|Y) \equiv \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j) \ln_q \frac{1}{p(x_i|y_j)}, \quad (q \neq 1), \tag{45}$$

and

$$S_q(X,Y) \equiv \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j) \ln_q \frac{1}{p(x_i, y_j)}, \quad (q \neq 1). \tag{46}$$

We should note that Tsallis conditional entropy defined in Eq.(45) is not equal to that defined in Eq.(14), while Tsallis joint entropy defined in Eq.(46) is equal to that defined in Eq.(15). If we adopt the above definitions Eq.(45) instead of Eq.(14), we have the following inequality.

Proposition 6.3 For $q > 1$, we have

$$S_q(X,Y) \leq S_q(X) + \hat{S}_q(Y|X).$$

For $0 \leq q < 1$, we have

$$S_q(X,Y) \geq S_q(X) + \hat{S}_q(Y|X).$$

Therefore we do not have the chain rule for $\hat{S}_q(Y|X)$ in general, namely we are not able to construct a parametrically extended information theory under Definition 6.2.

References

[1] S.Abe, Heat and entropy in nonextensive thermodynamics: transmutation from Tsallis theory to Rényi-entropy-based theory,Physica A,Vol.300,pp.417-423(2001).

[2] S.Abe, S. Martínez, F. Pennini, and A. Plastino, Nonextensive thermodynamic relations, Phys.Lett.A,Vol.281,pp.126-130(2001).

[3] J.Aczél and Z.Daróczy, On measures of information and their characterizations, Academic Press, 1975.

[4] T.M.Cover and J.A.Thomas, Elements of Information Theory, John Wiley and Sons, 1991 (Second Edition 2006).

[5] Z.Daróczy, General information functions, Information and Control, Vol.16,pp.36-51(1970).

[6] K.S.Fa, Note on generalization of Shannon theorem and inequality, J.Phys.A; Math. and Gen.,Vol.31,pp.8159-8164(1998).

[7] A.D.Faddeev, On the notion of entropy of a finite probability space, Uspekhi Mat. Nauk, Vol.11,pp.227-231(1956), (in Russian).

[8] S.Furuichi, On uniqueness theorems for Tsallis entropy and Tsallis relative entropy, IEEE Trans. Information Theory, Vol.51,pp.3638-3645(2005).

[9] S.Furuichi, Information theoretical properties of Tsallis entropies, J. Math. Phys., Vol.47,023302(2006).

[10] S.Furuichi, On the maximum entropy principle in Tsallis statistics, preprint.

[11] S.Furuichi, K.Yanagi and K.Kuriyama, Fundamental properties of Tsallis relative entropy, J.Math.Phys, Vol.45, No.12, pp. 4868-4877 (2004).

[12] G.H.Hardy, J.E.Littlewood and G.Pólya, Inequality, Cambridge University Press, 1952.

[13] A.Hobson, A new theorem of information theory, J.Stat.Phys.,Vol.1,pp.383-391(1969).

[14] A.I.Khinchin, The concept of entropy in probability theory, Uspekhi Mat. Nauk, Vo.8,pp.3-20(1953), (in Russian).

[15] A.I.Khinchin, Mathematical foundations of information theory, Dover Publications,Inc., New York,1957.

[16] S.Kullback, Information theory and statistics, Dover Publications,Inc., New York,1967.

[17] S.Kullback and R.A.Leibler, On information and sufficiency, Ann.Math.Statist., Vol.22,pp.79-86(1951).

[18] S. Martinez, F. Nicolás, F. Penninia and A. Plastinoa, Tsallis' entropy maximization procedure revisited, Physica A,Vol.286, pp.489-502(2000).

[19] A.K.Rajagopal and S.Abe,Implications of form invariance to the structure of nonextensive entropies, Phys. Rev. Lett., Vol.83, pp.1711-1714(1999).

[20] C.E.Shannon, A mathematical theory of communication, Bell Syst.Tech.J.,Vol.27(1948),pp.379-423,pp.623-656.

[21] M.Shiino, H-theorem with generalized relative entropies and the Tsallis statistics, J.Phys.Soc.Japan,Vol.67,pp.3658-3660(1998).

[22] H.Suyari, Generalization of Shannon-Khinchin axioms to nonextensive systems and the uniqueness theorem for the nonextensive entropy, IEEE Trans. Information Theory, Vol.50, pp.1783-1787(2004).

[23] H.Suyari, The unique non self-referential q-canonical distribution and the physical temperature derived from the maximum entropy principle in Tsallis statistics, cond-mat/0502298.

[24] H.Suyari and S.Furuichi, Comment on "Note on generalization of Shannon theorem and inequality", J.Phys.A: Math. and Gen., Vol.36,pp.5157-5158(2003).

[25] H.Suyari and M.Tsukada, Law of error in Tsallis statistics, IEEE Trans. Information Theory, Vol.51,pp.753-757(2005).

[26] H.Tveberg, A new derivation of the information function, Math. Scand., Vol.6, pp.297-298(1958).

[27] C.Tsallis, Possible generalization of Bolzmann-Gibbs statistics, J.Stat.Phys.,Vol.52,pp.479-487(1988).

[28] C.Tsallis, Generalized entropy-based criterion for consistent testing, Phys.Rev.E,Vol.58,pp.1442-1445(1998).

[29] C. Tsallis et al., Nonextensive Statistical Mechanics and Its Applications, edited by S. Abe and Y. Okamoto, Springer-Verlag, Heidelberg,2001; see also the comprehensive list of references at http://tsallis.cat.cbpf.br/biblio.htm.

[30] C.Tsallis, R.S.Mendesc and A.R.Plastino, The role of constraints within generalized nonextensive statistics, Physica A,Vol.261,pp.534-554(1998).

[31] T.Yamano, Information theory based on nonadditive information content, Phys.Rev.E,Vol.63,046105(2001).