A generalized Faddeev’s axiom and the uniqueness theorem for Tsallis entropy

Shigeru Furuichi

1Department of Electronics and Computer Science,
Tokyo University of Science, Yamaguchi, 756-0884, Japan

Abstract. The uniqueness theorem for the Tsallis entropy by introducing the generalized Faddeev’s axiom is proven. Our result improves the recent result, the uniqueness theorem for Tsallis entropy by the generalized Shannon-Khinchin’s axiom in [7], in the sense that our axiom is simpler than his one, as similar that Faddeev’s axiom is simpler than Shannon-Khinchin’s one.

Keywords: generalized Faddeev’s axiom, uniqueness theorem and Tsallis entropy

2000 Mathematics Subject Classification: 94A17, 46N55

1 Introduction

As a typical feature of Shannon entropy

\[ S_1(X) \equiv - \sum_{i=1}^{n} x_i \log x_i \]  

(1)

defined for the probability distribution \( p_i \equiv p(X = x_i) \) of the random variable \( X \), the additivity

\[ S_1(X \times Y) = S_1(X) + S_1(Y) \]  

(2)

for two independent random variables \( X \) and \( Y \) is known. The additivity also holds for Rényi entropy [1, 2] which is famous as a generalization of Shannon entropy.

As an another generalization of Shannon entropy, Tsallis entropy

\[ S_q(X) \equiv - \sum_{i=1}^{n} x_i^q \ln_q x_i, \]  

(3)

where \( q \)-logarithm function \( \ln_q \) is defined by \( \ln_q(x) \equiv \frac{x^{1-q}}{1-q} \) for \( q \geq 0 \) and \( x \geq 0 \), was introduced in [3] with different mathematical feature from Rényi entropy, since acutually Tsallis entropy does not have the additivity as it will be noted in the below. Since \( q \)-logarithm function \( \ln_q(x) \) uniformly converges to \( \log x \) as \( q \to 1 \) for \( x \geq 0 \) by Dini’s theorem, Tsallis entropy converges to Shannon entropy as \( q \to 1 \), which means Tsallis entropy is one parameter extension of Shannon entropy. Also since \( q \)-logarithm function \( \ln_q(x) \) has the pseudoadditivity :

\[ \ln_q(xy) = \ln_q(x) + \ln_q(y) + (1 - q) \ln_q(x) \ln_q(y), \quad (q \neq 1), \]  

(4)
Tsallis entropy has the pseudoadditivity:

\[ S_q(X \times Y) = S_q(X) + S_q(Y) + (1 - q)S_q(X)S_q(Y), \quad (q \neq 1). \] (5)

As a similar generalization of Shannon entropy, the structural \(a\)-entropy \([4]\) or called the entropy of type \(\beta\) \([4]\) is traditionally known \([6]\). These entropies are classified into the nonextensive (nonadditive) entropies since they do not have the additivity for two independent random variables \(X\) and \(Y\), while Shannon entropy and Rényi entropy are classified into the extensive (additive) entropies.

Recently, the nonextensive entropies including the Tsallis entropy was characterized by H.Suyari in terms of the generalized Shannon-Khinchin’s axiom in \([7]\). See also \([4]\) for the uniqueness theorem by a generalization of the Shannon-Khinchin’s axiom for the structural \(a\)-entropy which is one of the nonextensive entropies, as a first appearance of such a generalized result.

In the previous paper, we developed these works to the characterization of the Tsallis relative entropy \([8]\). Historically, the Shannon-Khinchin axiom \([9]\) was improved by A.D.Faddeev \([10]\) in the sense that Faddeev’s axiom is simpler than Shannon-Khinchin’s one. The proof of the uniqueness theorem for Shannon entropy by means of the weaker condition than the original condition of Faddeev’s axiom was completed by H.Tveberg in \([11]\). See also \([12, 13, 14]\) for the details. Inspired by this fact and the recent fine result \([7]\), in this short paper, we simplify the generalized Shannon-Khinchin’s axiom \([7]\) as the generalization of the Faddeev’s axiom, and then prove the uniqueness theorem for Tsallis entropy.

2 A generalized axiom and the uniqueness theorem

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 \(R^+ \equiv [0, \infty)\). 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 2.1 (Generalized Faddeev’s axiom)

(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 \(\{x_k\} \in \Delta_n\),

\[ S_q(x_1, \cdots, x_n) = S_q(x_1', \cdots, x_n'). \] (6)

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

\[ S_q(x_1, \cdots, x_{n-1}, y, z) = S_q(x_1, \cdots, x_n) + x_n^qS_q \left( \frac{y}{x_n}, \frac{z}{x_n} \right). \] (7)

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 \([7]\), since our condition (GF3) does not have to take the summation on \(i\) from 1 to \(n\).

For the above generalized Faddeev’s axiom, we have the following uniqueness theorem for Tsallis entropy.

2
Theorem 2.2 Three conditions (GF1),(GF2) and (GF3) uniquely give the form of the function $S_q : \Delta_n \to \mathbb{R}^+$ such that

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

(8)

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

(Proof) In the special case of $q = 1$, the theorem follows by \[1\]. Thus we suppose $q \neq 1$ in the sequel. We prove the theorem as similar way of the proof by H.Tveberg \[11\]. From (GF2) and (GF3), for any $x, y, z$ satisfying $x, y \geq 0$, $z > 0$ and $x + y + z = 1$, we expand $S_q(x, y, z)$ into the different equations,

$$S_q(x, y, z) = S_q(x, y + z) + (y + z)^q S_q \left( \frac{y}{y + z}, \frac{z}{y + z} \right)$$

$$= S_q(y, x + z) + (x + z)^q S_q \left( \frac{x}{x + z}, \frac{z}{x + z} \right).$$

Therefore we have

$$f_q(x) + (1 - x)^q f_q \left( \frac{y}{1 - x} \right) = f_q(y) + (1 - y)^q f_q \left( \frac{x}{1 - y} \right)$$

(9)

Since Eq.(9) is defined for any $0 \leq x < 1$ and $0 \leq y < 1$, by setting $x = 0$ and $y > 0$, we have

$$f_q(0) + f_q(y) = f_q(y) + (1 - y)^q f_q(0).$$

Thus we have

$$S_q(0, 1) = f_q(0) = 0.$$  

(10)

Integrating both sides in Eq.(10) with respect to $y$ from $0$ to $1 - x$, we have

$$\int_0^{1-x} f_q(x) dy + (1 - x)^q \int_0^{1-x} f_q \left( \frac{y}{1 - x} \right) dy = \int_0^{1-x} f_q(y) dy + \int_0^{1-x} (1 - y)^q f_q \left( \frac{x}{1 - y} \right) dy,$$

which can be deformed as follows

$$(1 - x) f_q(x) + (1 - x)^{q+1} \int_0^1 f_q(t) dt = \int_0^{1-x} f_q(t) dt + x^{q+1} \int_x^1 t^{-q-2} f_q(t) dt.$$  

(11)

Since the function $f_q(x)$ is continuous on the closed interval $[0, 1]$ due to (GF1), it is differentiable on the open interval $x \in (0, 1)$. By differentiating both sides of Eq.(11) and applying the relation

$$f_q(x) = f_q(1 - x)$$

(12)

due to (GF2), we have

$$(1 - x) f_q'(x) = (q + 1) (1 - x)^q \int_0^1 f_q(t) dt + (q + 1) x^q \int_x^1 t^{-q-2} f_q(t) dt - \frac{f_q(x)}{x}. $$

(13)

Again differentiating both sides in Eq.(13), we have

$$(1 - x) f_q''(x) = -q (q + 1) (1 - x)^{q-1} \int_0^1 f_q(t) dt + q (q + 1) x^{q-1} \int_x^1 t^{-q-2} f_q(t) dt$$

$$- \frac{q f_q(x)}{x^2} - \frac{f_q'(x)}{x} + f_q''(x)$$

(14)
Multiplying $x$ to both sides in Eq. (14), we have

$$x (1-x) f''_q (x) = -q (q+1) x (1-x)^{q-1} \left( \int_0^1 f_q (t) \, dt + q (q+1) x^q \int_x^1 t^{-q-2} f_q (t) \, dt \right)$$

$$-q f'_q (x) + (x-1) f'_q (x).$$

(15)

Also multiplying $q$ to both sides in Eq. (13), we have

$$q (1-x) f'_q (x) = q (q+1) (1-x)^q \left( \int_0^1 f_q (t) \, dt + q (q+1) x^q \int_x^1 t^{-q-2} f_q (t) \, dt - \frac{q f'_q (x)}{x} \right)$$

(16)

From Eq. (15) and Eq. (16), we have the following differential equation:

$$x f''_q (x) = (q-1) f'_q (x) - q (q+1) \mu_q (1-x)^q,$$

(17)

where we set $\mu_q \equiv \int_0^1 f_q (t) \, dt$. This differential equation has the following general solution with the constant numbers $c_1$ and $c_2$:

$$f_q (x) = c_1 + c_2 \frac{x^q}{q} + \frac{(q+1) \mu_q (1-x)^q}{1-q}.$$  

(18)

The initial condition Eq. (10) implies $c_1 = \frac{(q+1) \mu_q}{q-1}$. Also the condition Eq. (12) implies $c_2 = \frac{q(q+1) \mu_q}{1-q}$. Substituting $c_1$ and $c_2$ into Eq. (18), we have

$$f_q (x) = -\lambda_q \left\{ x^q \ln_x x + (1-x)^q \ln_q (1-x) \right\}.$$  

(19)

after the calculations, where we again set $\lambda_q \equiv (q+1) \mu_q$. Since there exists some $t_0 \in [0, 1]$ such that $f_q (t_0) > 0$ due to (GF1) and the range of $S_q$ is $\mathbb{R}_+$, we have $\mu_q > 0$ and then we have $\lambda_q > 0$ for any $q \geq 0$. Therefore we could prove Eq. (8) for $n=2$. Finally we prove Eq. (8) for the general $n \geq 3$ by induction on $n$. On the assumption that Eq. (8) is true for any $n$, the following calculations directly follow.

$$S_q (x_1, \ldots, x_n, x_{n+1}) = S_q (x_1, \ldots, x_n + x_{n+1}) + (x_n + x_{n+1})^q S_q \left( \frac{x_n}{x_n + x_{n+1}}, \frac{x_{n+1}}{x_n + x_{n+1}} \right)$$

$$= -\lambda_q \sum_{i=1}^{n-1} x_i^q \ln x_i - \lambda_q (x_n + x_{n+1})^q \ln_q (x_n + x_{n+1})$$

$$-\lambda_q (x_n + x_{n+1})^q \left\{ \left( \frac{x_n}{x_n + x_{n+1}} \right)^q \ln_q \left( \frac{x_n}{x_n + x_{n+1}} \right) + \left( \frac{x_{n+1}}{x_n + x_{n+1}} \right)^q \ln_q \left( \frac{x_{n+1}}{x_n + x_{n+1}} \right) \right\}$$

$$= -\lambda_q \sum_{i=1}^{n-1} x_i^q \ln x_i - \lambda_q (x_n + x_{n+1})^q \ln_q (x_n + x_{n+1})$$

$$-\lambda_q x_n^q \ln_q \left( \frac{x_n}{x_n + x_{n+1}} \right) - \lambda_q x_{n+1}^q \ln_q \left( \frac{x_{n+1}}{x_n + x_{n+1}} \right)$$

$$= -\lambda_q \sum_{i=1}^{n-1} x_i^q \ln x_i + \lambda_q (x_n + x_{n+1}) \ln_q \frac{1}{x_n + x_{n+1}}$$

$$-\lambda_q x_n^q \left( \ln_q x_n + x_n^{1-q} \ln_q \frac{1}{x_n + x_{n+1}} \right) - \lambda_q x_{n+1}^q \left( \ln_q x_{n+1} + x_{n+1}^{1-q} \ln_q \frac{1}{x_n + x_{n+1}} \right)$$

$$= -\lambda_q \sum_{i=1}^{n+1} x_i^q \ln x_i.$$  

This shows that Eq. (8) is also true for $n+1$. Thus the proof of this theorem completed.
3 A relation to the generalized Shannon-Khinchin’s axiom

In this section, we study the relation between the generalized Shannon-Khinchin’s axiom introduced in \[7\] 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 3.1 (Generalized Shannon-Khinchin’s axiom)**

\((\text{GSK1})\) Continuity: The function \(S_q : \Delta_n \rightarrow \mathbb{R}^+\) is continuous.

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

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

\[
S_q(x_{11}, \cdots, x_{nm_n}) = S_q(x_1, \cdots, x_n) + \sum_{i=1}^{n} x_i^q \sum_{j=1}^{m_i} \left( S_q\left( x_{1i}, \cdots, x_{mi} \right) \right).
\]

\((\text{GSK4})\) 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 \[7\]. Then we have the following proposition.

**Proposition 3.2** Axiom 3.1 implies Axiom 2.1

**(Proof)** It is trivial that (GSK1) and (GSK2) imply (GF1). We show that (GSK1) and (GSK3) imply (GF2). If all \(x_i, (i = 1, \cdots, n)\) are positive rational numbers, each \(x_i\) can be represented by \(\frac{l_i}{m}, (2 \leq l_i \leq m, l_i, m \in \mathbb{Z})\). Applying (GSK3), since \(x_i = \frac{l_i}{m} = \sum_{j=1}^{l_i} \frac{1}{m}\), we have

\[
S_q(x_1, \cdots, x_n) = S_q\left( \frac{l_1}{m}, \cdots, \frac{l_n}{m} \right)
= S_q\left( \frac{1}{m}, \frac{1}{m}, \cdots, \frac{1}{m} \right) - \sum_{i=1}^{m} x_i^q S_q\left( \frac{1}{l_i}, \cdots, \frac{1}{l_i} \right)
\]

The first term of the right hand side in the above equation does not depend on the order of \((l_1, \cdots, l_n)\). Also the way to take the summation in the second term of the right hand side in the above equation is arbitrary so that the above equation is equal to

\[
S_q\left( \frac{1}{m}, \frac{1}{m}, \cdots, \frac{1}{m} \right) - \sum_{i=1}^{m} x_i^q S_q\left( \frac{1}{l'_1}, \cdots, \frac{1}{l'_n} \right)
\]

for the permutation \(\{x'_i\}\) from \(\{x_i\}\) where \(x'_i = \frac{l_i'}{m}, (2 \leq l_i' \leq m, l_i', m \in \mathbb{Z})\). That is, (GF2) holds for any rational numbers \(x_i\). If \(x_i\) is not the rational number, then we use the continuity of (GSK1) after the approximation of \(x_i\) by the rational number, and then we have (GF2). Finally we show that (GSK3) and (GSK4) imply (GF3). From (GSK3), (GSK4) and (GF2), we have

\[
S_q\left( \frac{1}{2}, \frac{1}{2} \right) = S_q\left( \frac{1}{2}, \frac{1}{2}, 0, 0 \right) = S_q\left( \frac{1}{2}, 0, \frac{1}{2}, 0 \right) = S_q\left( \frac{1}{2}, \frac{1}{2}, 1 \right) = \frac{1}{2} S_q(1,0) + \frac{1}{2} S_q(1,0).
\]
Therefore we have \( S_q(1, 0) = 0 \). Thus we have

\[
S_q(x_1, \ldots, x_{n-1}, y, z) = S_q(x_1, 0, x_2, 0, \ldots, x_{n-1}, 0, y, z)
\]

\[
= S_q(x_1, \ldots, x_n) + \sum_{i=1}^{n-1} x_i^qS_q(1, 0) + x_n^qS_q\left(\frac{y}{x_n}, \frac{z}{x_n}\right)
\]

\[
= S_q(x_1, \ldots, x_n) + \sum_{i=1}^{n-1} x_i^qS_q\left(\frac{y}{x_n}, \frac{z}{x_n}\right)
\]

which implies (GF3).

We also have the following proposition.

**Proposition 3.3** \( S_q(X) = -\lambda_q \sum_{i=1}^{n} x_i^q \ln_q x_i \) satisfies Axiom 3.1.

**(Proof)** (GSK1) and (GSK4) are trivial. We prove (GSK2) by the use of the non-negativity of the Tsallis relative entropy:

\[
D_q(X|Y) \equiv -\sum_{i=1}^{n} x_i \ln_q \frac{y_i}{x_i}
\]

for two random variables \( X \) and \( Y \), where \( \{x_i\} \) and \( \{y_i\}, (i = 1, \ldots, n) \) are probability distributions of \( X \) and \( Y \), respectively. See [15] for the mathematical properties of the Tsallis relative entropy. Its non-negativity can be easily proven by the convexity of \( -\ln_q(x) \). The non-negativity implies \( S_q(X) \leq \ln_q n \) by setting the random variable \( U = \{ \frac{1}{n}, \ldots, \frac{1}{n} \} \) having the uniform distribution instead of \( Y \). We easily find that the maximum value is attained when \( X = \{ \frac{1}{n}, \ldots, \frac{1}{n} \} \). Note that \( \lambda_q \) does not depend on the way to take the maximum of \( S_q(X) \). Thus (GSK2) is proven. Finally (GSK3) is proven by the direct calculations.

\[
S_q(x_1, \ldots, x_n) + \sum_{i=1}^{n} x_i^qS_q\left(\frac{x_i}{x_i}, \frac{x_{im1}}{x_i}\right)
\]

\[
= -\lambda_q \sum_{i=1}^{n} x_i^q \ln_q x_i + \lambda_q \sum_{i=1}^{n} x_i^q \ln_q \frac{x_{im1}}{x_i} + \cdots + \lambda_q \sum_{i=1}^{n} x_n^q \ln_q \frac{x_{nmn}}{x_n}
\]

\[
= -\lambda_q \sum_{i=1}^{n} x_i^q \ln_q x_i - \lambda_q \left( x_i^q \ln_q x_{im1} + x_{im1}^q \ln_q x_{im1} + x_{nm1}^q \ln_q x_{im1} \ln_q \frac{1}{x_1} \right)
\]

\[
= \lambda_q \left( x_{im1}^q \ln_q x_{im1} + x_{nm1}^q \ln_q x_{nm1} + x_{nmn}^q \ln_q x_{nmn} \right)
\]

\[
= S_q\left(x_{11} + \cdots + x_{im1} \ln_q \frac{1}{x_1} + \cdots + x_{nmn} \ln_q \frac{1}{x_n}\right)
\]

From Theorem 2.2, Proposition 3.2 and Proposition 3.3, we have the following equivalent relation among Axioms 2.1, Axiom 3.1 and the Tsallis entropy.
Theorem 3.4 The following three statements are equivalent to one another.

1. \( S_q : \Delta_n \rightarrow \mathbb{R}^+ \) satisfies Axiom 3.1
2. \( S_q : \Delta_n \rightarrow \mathbb{R}^+ \) satisfies Axiom 2.1
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 Conclusion

As we have seen, Tsallis entropy was characterized by the generalized Faddeev’s axiom which is a simplification of the generalized Shannon-Khinchin’s axiom introduced in [7]. And then we slightly improved the uniqueness theorem proved in [7], 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 [10] [11].

References

[1] A.Rényi, On measures of entropy and information, Proc. Forth Berkeley Symp. Math. Stat. Prob., Vol. 1, pp. 547-561, Univ. California Press, Berkeley, 1961.

[2] A.Rényi, Wahrscheinlichkeitsrechnung. Mit einem Anhang über Informationstheorie., VEB Deutscher Verlag der Wissenschaften, Berlin, 1962.

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

[4] J.Havrda and F.Charvát, Quantification method of classification processes, Concept of structural a-entropy, Kybernetika (Prague), Vol. 3, pp. 30-35 (1967).

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

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

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

[8] S.Furuichi, A characterization of the Tsallis relative entropy by the generalized properties, preprint.

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

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

[11] H.Tveberg, A new derivation of the information function, Math. Scand., Vol. 6, pp. 297-298 (1958).
[12] A.I.Khinchin, Mathematical foundations of information theory, Dover Publications, Inc., New York, 1957.

[13] A. Feinstein, Foundations of information theory, McGraw-Hill Book Co., New York, 1958.

[14] Y. Kakihara, Abstract methods in information theory, World Scientific, 1999.

[15] S. Furuichi, K. Yanagi and K. Kuriyama, Fundamental properties of Tsallis relative entropy, to appear in J. Math. Phys, Vol. 45, No. 12, 2004.