<table>
<thead>
<tr>
<th>Title</th>
<th>エントロピーから導かれる数理構造 情報科学と数解析の接点 [これまでとこれから]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>須鎗 弘樹</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 [これまでとこれから]</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2004-10</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/25981">http://hdl.handle.net/2433/25981</a></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

Kyoto University
Tsallis エントロピーから導かれる数理構造

Hiroki Suyari
Department of Information and Image Sciences, Chiba University, 263-8522, Japan
e-mail: suyari@ieee.org, suyari@faculty.chiba-u.jp

1 Introduction

Tsallis entropy [1][2] in discrete system and continuous system are respectively defined as follows:

\[ S_q^{(d)} := \frac{1 - \sum_{i=1}^{n} p_i^q}{q-1} \quad (q \in \mathbb{R}^+) \quad (1) \]

\[ S_q^{(c)} := \frac{1 - \int f(x)^q dx}{q-1} \quad (q \in \mathbb{R}^+) \quad (2) \]

where \( \{p_i\}_{i=1}^{n} \) is a probability distribution and \( f \) is a probability density function. Tsallis entropy recovers Shannon entropy when \( q \rightarrow 1 \).

\[ \lim_{q \rightarrow 1} S_q^{(d)} = -\sum_{i=1}^{n} p_i \ln p_i, \quad \lim_{q \rightarrow 1} S_q^{(c)} = -\int f(x) \ln f(x) dx \quad (3) \]

The maximum entropy principle for Tsallis entropy \( S_q^{(c)} \) under the constraints:

\[ \int f(x) dx = 1, \quad \frac{\int x^2 f(x)^q dx}{\int f(y)^q dy} = \sigma^2 \quad (4) \]

yields the so-called \( q \)-Gaussian probability density function:

\[ f(x) = \frac{\exp_q (-\beta_q x^2)}{\int \exp_q (-\beta_q y^2) dy} \propto [1 + (1-q) (-\beta_q x^2)]^{\frac{1}{1-q}}, \quad (5) \]

*This work was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Encouragement of Young Scientists(B), 14780259, 2004.
where \( \exp_q(x) \) is the \( q \)-exponential function defined by

\[
\exp_q(x) := \begin{cases} 
[1 + (1 - q)x]^{\frac{1}{1-q}} & \text{if } 1 + (1 - q)x > 0, \\
0 & \text{otherwise}
\end{cases} \quad (x \in \mathbb{R})
\]  

(6)

and \( \beta_q \) is a positive constant related to \( \sigma \) and \( q \) [3][4]. For \( q \to 1 \), \( q \)-Gaussian distribution (5) recovers a usual Gaussian distribution. The power form in \( q \)-Gaussian (5) has been found to be fairly fitted to many physical systems which cannot be systematically studied in the usual Boltzmann-Gibbs statistical mechanics [5][6].

![Graph showing q-Gaussian probability density function](image)

**图 1: q-Gaussian probability density function (q = 0.2, 0.6, 1.0, 1.4, 1.8)**

The mathematical basis for Tsallis statistics comes from the deformed expressions for the logarithm and the exponential functions which are the \( q \)-logarithm function:

\[
\ln_q x := \frac{x^{1-q} - 1}{1-q} \quad (x \geq 0, q \in \mathbb{R})
\]  

(7)

and its inverse function, the \( q \)-exponential function (6). Using the \( q \)-logarithm function (7), Tsallis entropy (2) can be written as

\[
S^c_q = - \int f(x)^q \ln_q f(x) \, dx,
\]  

(8)

which is easily found to recover Shannon entropy when \( q \to 1 \).

The many successful applications of Tsallis statistics stimulate us to try to find the new mathematical structure behind Tsallis statistics [7][8][9]. Recently, a new multiplication operation determined by the \( q \)-logarithm and the \( q \)-exponential function which naturally emerge from Tsallis entropy is presented in [10] and [11]. Using this algebra, we can obtain the beautiful mathematical structure behind Tsallis statistics. In this summary, we briefly show our results without proofs. Each proof can be found in my references. The contents consist of the 7 sections, section I: introduction (this section), section II: the new multiplication operation, section III: law of error in Tsallis statistics, section IV: \( q \)-Stirling’s formula, section V: \( q \)-multinomial coefficient and symmetry in Tsallis statistics, section VI: \( q \)-central limit theorem in Tsallis statistics (numerical evidence only) and section VII: conclusion.
2 The new multiplication operation determined by $q$–logarithm function and $q$–exponential function

The new multiplication operation $\otimes_q$ is introduced in [10] and [11] for satisfying the following equations:

\[
\ln_q (x \otimes_q y) = \ln_q x + \ln_q y, \quad (9)
\]

\[
\exp_q (x) \otimes_q \exp_q (y) = \exp_q (x + y). \quad (10)
\]

These lead us to the definition of $\otimes_q$ between two positive numbers

\[
x \otimes_q y := \begin{cases} 
[x^{1-q} + y^{1-q} - 1]^{\frac{1}{1-q}}, & \text{if } x > 0, y > 0, \\
0, & \text{otherwise}
\end{cases} \quad (11)
\]

which is called $q$–product in [11]. The $q$–product recovers the usual product such that $\lim_{q \rightarrow 1} (x \otimes_q y) = xy$. The fundamental properties of the $q$–product $\otimes_q$ are almost the same as the usual product, but the distributive law does not hold in general.

\[
a (x \otimes_q y) \neq ax \otimes_q y \quad (a, x, y \in \mathbb{R}) \quad (12)
\]

The properties of the $q$–product can be found in [10] and [11].

In order to see one of the validities of the $q$–product in Tsallis statistics, we recall the well known expression of the exponential function $\exp (x)$ given by

\[
\exp (x) = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n. \quad (13)
\]

Replacing the power on the right side of (13) by the $n$ times of the $q$-product $\otimes_q^n$:

\[
x^{\otimes_q^n} := x \otimes_q \cdots \otimes_q x, \quad \text{n times} \quad (14)
\]

$\exp_q (x)$ is obtained. In other words, $\lim_{n \rightarrow \infty} (1 + \frac{x}{n})^{\otimes_q^n}$ coincides with $\exp_q (x)$.

\[
\exp_q (x) = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^{\otimes_q^n} \quad (15)
\]

The proof of (15) is given in [12]. This coincidence indicates a validity of the $q$-product in Tsallis statistics. In fact, the present results in the following sections reinforce it.

3 Law of error in Tsallis statistics

Consider the following situation: we get $n$ observed values:

\[
x_1, x_2, \ldots, x_n \in \mathbb{R} \quad (16)
\]

as results of mutually independent $n$ measurements for some observation. Each observed value $x_i$ $(i = 1, \ldots, n)$ is each value of independent, identically distributed (i.i.d. for short) random variable $X_i$ $(i = 1, \ldots, n)$, respectively. There exist a true value $x$ satisfying the additive relation:

\[
x_i = x + e_i \quad (i = 1, \ldots, n), \quad (17)
\]
where each of \( e_i \) is an error in each observation of a true value \( x \). Thus, for each \( X_i \), there exists a random variable \( E_i \) such that \( X_i = x + E_i \) \((i = 1, \ldots, n)\). Every \( E_i \) has the same probability density function \( f \) which is differentiable, because \( X_1, \ldots, X_n \) are i.i.d. (i.e., \( E_1, \ldots, E_n \) are i.i.d.). Let \( L(\theta) \) be a function of a variable \( \theta \), defined by

\[
L(\theta) := f(x_1 - \theta) f(x_2 - \theta) \cdots f(x_n - \theta).
\]  

(18)

**Theorem 1** If the function \( L(\theta) \) of \( \theta \) for any fixed \( x_1, x_2, \ldots, x_n \) takes the maximum at

\[
\theta = \theta^* := \frac{x_1 + x_2 + \cdots + x_n}{n},
\]

then the probability density function \( f \) must be a Gaussian probability density function:

\[
f(e) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{e^2}{2\sigma^2} \right\}.
\]  

(20)

The above result goes by the name of “Gauss’ law of error” [13]. Gauss’ law of error tells us that in measurements it is the most probable to assume a Gaussian probability distribution for additive noise, which is often used as an assumption in many scientific fields. On the basis of Gauss’ law of error, some functions such as error function are often used to estimate error rate in measurements.

Gauss’ law of error is generalized to Tsallis statistics, which results in \( q \)-Gaussian as a generalization of a usual Gaussian distribution.

Consider almost the same setting as Gauss’ law of error: we get \( n \) observed values (16) as results of \( n \) measurements for some observation. Each observed value \( x_i \) \((i = 1, \ldots, n)\) is each value of identically distributed random variable \( X_i \) \((i = 1, \ldots, n)\), respectively. There exist a true value \( x \) satisfying the additive relation (17) where each of \( e_i \) is an error in each observation of a true value \( x \). Thus, for each \( X_i \), there exists a random variable \( E_i \) such that \( X_i = x + E_i \) \((i = 1, \ldots, n)\). Every \( E_i \) has the same probability density function \( f \) which is differentiable, because \( X_1, \ldots, X_n \) are identically distributed random variables (i.e., \( E_1, \ldots, E_n \) are also so). Let \( L_q(\theta) \) be a function of a variable \( \theta \), defined by

\[
L_q(\theta) := f(x_1 - \theta) \otimes_q f(x_2 - \theta) \otimes_q \cdots \otimes_q f(x_n - \theta).
\]  

(21)

**Theorem 2** If the function \( L_q(\theta) \) of \( \theta \) for any fixed \( x_1, x_2, \ldots, x_n \) takes the maximum at (19), then the probability density function \( f \) must be a \( q \)-Gaussian:

\[
f(e) = \frac{\exp_q\left(-\beta_q e^2\right)}{\int \exp_q\left(-\beta_q e^2\right) de}
\]  

(22)

where \( \beta_q \) is a \( q \)-dependent positive constant.

See [14] for the proof. Our result \( f(e) \) in (22) coincides with (5) by applying the \( q \)-product to MLP instead of MEP.

### 4 \( q \)-Stirling’s formula

By means of the \( q \)-product (11), the \( q \)-factorial \( n!_q \) is naturally defined by

\[
n!_q := 1 \otimes_q \cdots \otimes_q n
\]  

(23)

for \( n \in \mathbb{N} \) and \( q > 0 \). Using the definition (11), \( \ln_q(n!_q) \) is explicitly expressed by

\[
\ln_q(n!_q) = \sum_{k=1}^{n} \frac{k^{1-q} - n}{1 - q}.
\]  

(24)
If an approximation of $\ln_q(n!_q)$ is not needed, the above explicit form (24) should be directly used for its computation. However, in order to clarify the correspondence between the studies $q = 1$ and $q \neq 1$, the approximation of $\ln_q(n!_q)$ is useful. In fact, using the present $q$-Stirling’s formula, we obtain the surprising mathematical structure in Tsallis statistics [15].

The tight $q$-Stirling’s formula is derived as follows:

$$
\ln_q(n!_q) = \begin{cases} 
(n + \frac{1}{2}) \ln n + (1 - \delta_1) & \text{if } q = 1, \\
- \frac{n - \frac{1}{2}}{2n} \ln n - \frac{1}{2} + \theta_{n,2} - \delta_2 & \text{if } q = 2, \\
\left(\frac{n}{2-q} + \frac{1}{2}\right)^{\frac{1-q}{1-q}} + \theta_{n,q} + \left(\frac{1-n}{2-q} - \delta_q\right) & \text{if } q > 0 \text{ and } q \neq 1,2 
\end{cases}
$$

(25)

where

$$
\lim_{n \to \infty} \theta_{n,q} = 0, \quad 0 < \theta_{n,1} < \frac{1}{12n}, \quad e^{1-\delta_1} = \sqrt{2\pi}.
$$

(26)

On the other hand, the rough $q$-Stirling’s formula is reduced from (25) to

$$
\ln_q(n!_q) = \begin{cases} 
\frac{n}{2-q} \ln n - \frac{n}{2-q} + O(\ln n) & \text{if } q \neq 2, \\
n \ln n + O(1) & \text{if } q = 2,
\end{cases}
$$

(27)

which is more useful in applications in Tsallis statistics. See section V, section VI and [15] for its applications. The derivation of the above formulas is given in [12].

5 $q$–multinomial coefficient and symmetry in Tsallis statistics

We define the $q$–binomial coefficient $\left[\begin{array}{c} n \\ k \end{array}\right]_q$ by

$$
\left[\begin{array}{c} n \\ k \end{array}\right]_q := (n!_q) \otimes_q [(k!_q) \otimes_q ((n-k)!_q)] \quad (n, k (\leq n) \in \mathbb{N})
$$

(28)

where $\otimes_q$ is the inverse operation to $\emptyset_q$, which is defined by

$$
x \otimes_q y := \left\{ \begin{array}{ll}
[x^{1-q} - y^{1-q} + 1]^{\frac{1}{1-q}}, & x > 0, y > 0, \\
0, & x^{1-q} + y^{1-q} - 1 > 0,
\end{array} \right.
$$

(29)

$\emptyset_q$ is also introduced by the following satisfactions as similarly as $\otimes_q$ [10][11].

$$
\ln_q x \otimes_q y = \ln_q x - \ln_q y, \quad \exp_q (x) \otimes_q \exp_q (y) = \exp_q (x - y).
$$

(30)

(31)

Applying the definitions of $\otimes_q$, $\emptyset_q$ and $n!_q$ to (28), the $q$–binomial coefficient is explicitly written as

$$
\left[\begin{array}{c} n \\ k \end{array}\right]_q = \left[\begin{array}{c} n \\ k \end{array}\right] \left[\begin{array}{c} \sum_{i=1}^{n} i^{1-q} - \sum_{i=1}^{k} i^{1-q} - \sum_{j=1}^{n-k} j^{1-q} + 1 \end{array}\right]^{\frac{1}{1-q}}.
$$

(32)

From the definition (28), it is clear that

$$
\lim_{q \to 1} \left[\begin{array}{c} n \\ k \end{array}\right]_q = \left[\begin{array}{c} n \\ k \end{array}\right] = \frac{n!}{k!(n-k)!}.
$$

(33)
In general, when $q < 1$, $\sum_{i=1}^{n} i^{1-q} - \sum_{j=1}^{n-k} j^{1-q} + 1 > 0$. On the other hand, when $\sum_{i=1}^{n} i^{1-q} - \sum_{j=1}^{n-k} j^{1-q} + 1 < 0$, \[ \begin{bmatrix} n \\ k \end{bmatrix} \] takes complex numbers in general, which divides the formulations and discussions of the $q-$binomial coefficient into two cases: it takes a real number or a complex number. In order to avoid such separate formulations and discussions, we consider the $q$-logarithm of the $q$-binomial coefficient:

\[
\ln_q \begin{bmatrix} n \\ k \end{bmatrix} = \ln_q (n!_q) - \ln_q (k!_q) - \ln_q ((n-k)!_q). \tag{34}
\]

The above definition (28) is artificial because it is defined from the analogy with the usual binomial coefficient $\left[ \begin{array}{c} n \\ k \end{array} \right]$. However, when $n$ goes infinity, the $q-$binomial coefficient (28) has a surprising relation to Tsallis entropy as follows:

\[
\ln_q \begin{bmatrix} n \\ k \end{bmatrix} \simeq \left\{ \begin{array}{ll}
\frac{n^{1-q}}{2-q} \cdot S_{2-q}() & \text{if } q > 0, q \neq 2, \\
-S_1(n) + S_1(k) + S_1(n-k) & \text{if } q = 2
\end{array} \right.
\tag{35}
\]

where $S_q$ is Tsallis entropy (1) and $S_1(n)$ is Boltzmann entropy:

\[ S_1(n) := \ln n. \tag{36} \]

Applying the rough expression of the $q-$Stirling's formula (27), the above relations (35) are easily proved [15]. The above correspondence (35) between the $q-$binomial coefficient (28) and Tsallis entropy (1) convinces us of the fact the $q-$binomial coefficient (28) is well-defined in Tsallis statistics. Therefore, we can construct Pascal's triangle in Tsallis statistics [15].

The above relation (35) is easily generalized to the case of the $q-$multinomial coefficient. The $q-$multinomial coefficient in Tsallis statistics is defined in a similar way as that of the the $q-$binomial coefficient (28).

\[
\begin{bmatrix} n \\ n_1 \cdots n_k \end{bmatrix}_q := (n!_q) \otimes_q [(n_1!_q) \otimes_q \cdots \otimes_q (n_k!_q)] \tag{37}
\]

where

\[ n = \sum_{i=1}^{k} n_i, \quad n_i \in \mathbb{N} \ (i = 1, \cdots, k). \tag{38} \]

Applying the definitions of $\otimes_q$ and $\emptyset_q$ to (37), the $q-$multinomial coefficient is explicitly written as

\[
\begin{bmatrix} n \\ n_1 \cdots n_k \end{bmatrix}_q = \left[ \sum_{\ell=1}^{n} \ell^{1-q} - \sum_{i_1=1}^{n_1} i_1^{1-q} - \cdots - \sum_{i_k=1}^{n_k} i_k^{1-q} + 1 \right]^{rac{1}{1-q}}. \tag{39}
\]

Along the same reason as stated above in case of the $q-$binomial coefficient, we consider the $q$-logarithm of the $q-$multinomial coefficient given by

\[
\ln_q \begin{bmatrix} n \\ n_1 \cdots n_k \end{bmatrix}_q = \ln_q (n!_q) - \ln_q (n_1!_q) - \cdots - \ln_q (n_k!_q). \tag{40}
\]

From the definition (37), it is clear that

\[
\lim_{q \to 1} \begin{bmatrix} n \\ n_1 \cdots n_k \end{bmatrix}_q = \begin{bmatrix} n \\ n_1 \cdots n_k \end{bmatrix} = \frac{n!}{n_1! \cdots n_k!}. \tag{41}
\]
When $n$ goes infinity, the $q$-multinomial coefficient (37) has the similar relation to Tsallis entropy (1) as (35):

$$\ln_q \left[ \begin{array}{c} n \\ n_1 \cdots n_k \end{array} \right] \simeq \frac{n^{2-q}}{2-q} \cdot S_{2-q} \left( \frac{n_1}{n}, \cdots, \frac{n_k}{n} \right) \text{ if } q > 0, q \neq 2 \tag{42}$$

This is a natural generalization of (35). In the same way as the case of the $q$-binomial coefficient, the above relation (42) is easily proved [15].

When $q \to 1$, (42) recovers the well known result:

$$\ln \left[ \begin{array}{c} n \\ n_1 \cdots n_k \end{array} \right] \simeq nS_1 \left( \frac{n_1}{n}, \cdots, \frac{n_k}{n} \right) \tag{43}$$

where $S_1$ is Shannon entropy.

The relation (42) reveals a surprising symmetry: (42) is equivalent to

$$\ln_{1-(1-q)} \left[ \begin{array}{c} n \\ n_1 \cdots n_k \end{array} \right] \simeq \frac{n^{1+(1-q)}}{1+(1-q)} \cdot S_{1+(1-q)} \left( \frac{n_1}{n}, \cdots, \frac{n_k}{n} \right) \text{ if } q > 0, q \neq 2. \tag{44}$$

This expression represents that behind Tsallis statistics there exists a symmetry with a factor $1 - q$ around $q = 1$. Substitution of some concrete values of $q$ into (42) or (44) helps us understand the symmetry mentioned above.

### 6 $q$-central limit theorem in Tsallis statistics (numerical evidence only)

It is well known that any binomial distribution converge to a Gaussian distribution when $n$ goes infinity. This is a typical example of the central limit theorem in the usual probability theory. By analogy with this famous result, each set of normalized $q$-binomial coefficients is expected to converge to each $q$-Gaussian distribution with the same $q$ when $n$ goes infinity. As shown in this section, the present numerical results come up to our expectations.

In Fig.2, each set of bars and solid line represent each set of normalized $q$-binomial coefficients and $q$-Gaussian distribution with normalized $q$-mean 0 and normalized $q$-variance 1 for each $n$ when $q = 0.1$, respectively. Each of the three graphs on the first row of Fig.2 represents two kinds of probability distributions stated above, and the three graphs on the second row of Fig.2 represent the corresponding cumulative probability distributions, respectively. From Fig.2, we find the convergence of a set of normalized $q$-binomial coefficients to a $q$-Gaussian distribution when $n$ goes infinity. Other cases with different $q$ represent the similar convergences as the case of $q = 0.1$.

Note that we never use any curve-fitting in these numerical computations. Every bar and solid line is computed and plotted independently each other.

In order to confirm these convergences more precisely, we compute the maximal difference $\Delta_{q,n}$ among the values of two cumulative probabilities (a set of normalized $q$-binomial coefficients and $q$-Gaussian distribution) for each $q = 0.1, 0.2, \cdots, 0.9$ and $n$. $\Delta_{q,n}$ is defined by

$$\Delta_{q,n} := \max_{i=0,\cdots,n} \left| F_{q\text{-bin}} (i) - F_{q\text{-Gauss}} (i) \right| \tag{45}$$
where $F_{q\text{-bino}} (i)$ and $F_{q\text{-Gauss}} (i)$ are cumulative probability distributions of a set of normalized $q$-binomial coefficients $p_{q\text{-bino}} (k)$ and its corresponding $q$-Gaussian distribution $f_{q\text{-Gauss}} (x)$, respectively.

$$F_{q\text{-bino}} (i) := \sum_{k=0}^{i} p_{q\text{-bino}} (k), \quad F_{q\text{-Gauss}} (i) := \int_{-\infty}^{i} f_{q\text{-Gauss}} (x) \, dx \quad (46)$$

Fig.3 results in convergences of $\Delta_{q,n}$ to 0 when $n \to \infty$ for $q = 0.1, 0.2, \cdots, 0.9$. This result indicates that the limit of every convergence is a $q$-Gaussian distribution with the same $q \in (0,1]$ as that of a given set of normalized $q$-binomial coefficients.

The present convergences reveal a possibility of the existence of the central limit theorem in Tsallis statistics, in which any distributions converge to a $q$-Gaussian distribution as nonextensive generalization of a Gaussian distribution. The central limit theorem in Tsallis statistics provides not only a mathematical result in Tsallis statistics but also the physical reason why there exist universally power-law behaviors in many physical systems. In other words, the central limit theorem in Tsallis statistics mathematically explains the reason of ubiquitous existence of power-law behaviors in nature.

7 Conclusion

We discovered the conclusive and beautiful mathematical structure behind Tsallis statistics: law of error [14], $q$-Stirling’s formula [12], $q$-multinomial coefficient [15], $q$-central limit theorem (numerical evidences only) [15]. The one-to-one correspondence (42) between the $q$-multinomial coefficient and Tsallis entropy provides us with the significant mathematical structure such as symmetry (44) and others [15]. Moreover, the convergence of each set of $q$-binomial coefficients to each $q$-Gaussian with the same $q$ when $n$ increases represents the existence of the central limit theorem in Tsallis statistics.

In all of our recently presented results such as law of error [14], $q$-Stirling’s formula [12], symmetry (44) and the present numerical evidences for the central limit theorem in Tsallis statistics [15], the $q$-product is found to play crucial roles in these successful applications. The $q$-product is uniquely determined by Tsallis entropy. Therefore, the introduction of Tsallis entropy provides the wealthy foundations in mathematics and physics as well-organized generalization of the Boltzmann-Gibbs-Shannon theory.
図3: プロバビリティ分布（第一行）とその対応する累積分布関数（第二行）において、標準化$q$-二項係数$(q=0.5)$とTsallis分布の$n=5,10,20$

図4: プロバビリティ分布（第一行）とその対応する累積分布関数（第二行）において、標準化$q$-二項係数$(q=0.9)$とTsallis分布の$n=5,10,20$

参考文献


5: transition of maximal difference among the values of two cumulative probabilities (normalized $q$-binomial coefficient and $q$-Gaussian distribution) for each $q = 0.1, 0.2, \cdots, 0.9$ when $n \to \infty$


