Conceptual difficulties with the $q$-averages in non-extensive statistical mechanics

Sumiyoshi Abe
Department of Physical Engineering, Mie University, Mie 514-8507, Japan
suabe@sf6.so-net.ne.jp

Abstract. The $q$-average formalism of nonextensive statistical mechanics proposed in the literature is critically examined by considerations of several pedagogical examples. It is shown that there exist a number of difficulties with the concept of $q$-averages.

1. Introduction

In the last 1.5 decades, a lot of works have been done on so-called nonextensive statistical mechanics [1], which had been expected to be a generalization of Boltzmann-Gibbs statistical mechanics for describing certain complex systems. This theory is primarily based on a kind of the maximum entropy method, where the form of entropy and the definition of averages are changed respectively as follows [1,2]:

\[ S_q[p] = \frac{1}{1-q} \left[ \sum_{n=1}^{W} (p_n)^q - 1 \right], \quad (1) \]
\[ \langle Q \rangle_q = \sum_{n=1}^{W} Q_n P_n^{(q)}, \quad (2) \]

where \( \{P_n^{(q)}\}_{n=1,2,...,W} \) is the escort distribution defined in terms of the original distribution \( \{p_n\}_{n=1,2,...,W} \) as follows:

\[ P_n^{(q)} = \frac{(p_n)^q}{\sum_{m=1}^{W} (p_m)^q}. \quad (3) \]

Here, \( q \) is a positive constant, \( W \) the number of accessible states and \( \{Q_n\}_{n=1,2,...,W} \) a certain quantity of interest. Equation (2) is termed the $q$-average. \( S_q \) in equation (1) is referred to as the $q$-entropy. Together with the normalization constraint, \( \sum_{n=1}^{W} p_n = 1 \), the maximum-$S_q$-distribution is found to be given as follows:

\[ p_n = \frac{1}{Z_q(\beta)} e_q\left(-\left(\beta / c_q\right)(Q_n - \tilde{Q}_q)\right), \quad (4) \]

where \( \beta \) is a Lagrange multiplier for the constraint on the $q$-average in equation (2), \( e_q(x) = [1 + (1 - q)x]^{(1-q)} \) with the notation \( [a]_+ = \max\{0,a\} \).
\(Z_q(\beta) = \sum_{n=1}^{W} e_q\left(-\frac{\beta}{c_q} (Q_n - \tilde{Q}_q)\right)\), and \(c_q = \sum_{n=1}^{W} (p_n)^q\). Equation (4) is termed the \(q\)-exponential distribution. \(\tilde{Q}_q\) has to be calculated in terms of \(p_n\) in equation (4) itself in a self-referential manner. A basic requirement is that this scheme tends to the ordinary maximum entropy method [3] in the limit \(q \to 1\). In such a limit, equations (1) and (2) in fact converge to the Boltzmann-Gibbs-Shannon entropy, \(S[p] = -\sum_{n=1}^{W} p_n \ln p_n\) (with the Boltzmann constant being set equal to unity), and the ordinary average, \(\langle Q \rangle = \sum_{n=1}^{W} Q_n p_n\), respectively. Accordingly, equation (4) tends to the exponential distribution familiar in equilibrium statistical mechanics.

There is a naive question why \(q\) in equations (2) and (3) is the same as that in Eq. (1). In other words, why not

\[
\langle Q \rangle_{f(q)} = \frac{\sum_{n=1}^{W} Q_n (p_n)^{f(q)}}{\sum_{n=1}^{W} (p_m)^{f(q)}}
\]

with any \(f(q)\) satisfying \(f(q) \to 1\) \((q \to 1)\)?

In this article, we focus our attention on the notion of \(q\)-averages. We show that it does suffer from a series of conceptual difficulties and therefore cannot be employed in nonextensive statistical mechanics.

2. Lottery
This example can be seen as the simplest/trivial one, which shows that it is unreasonable to employ the \(q\)-average for a \(q\)-exponential distribution.

Suppose you are going to sell 869 lots. And among these 869, there are 100 first prizes with prize money 1000 Yen each, 144 second prizes with 500 Yen, 225 third prizes with 100 Yen and finally 400 fourth prizes with 0 Yen. The total amount of money you have to prepare is

\[
0 \times 400 + 100 \times 225 + 500 \times 144 + 1000 \times 100 = 197500\text{ Yen.}
\]

Therefore, if you sell each piece of lot more expensive than 197500 / 869 = 227.272727... Yen, then your pocket will be in safety.

The corresponding probability distribution is given by

\[
p_4 = N / 9 \text{ (first prize)}, \quad p_3 = 4N / 25 \text{ (second prize)},
p_2 = N / 4 \text{ (third prize)}, \quad p_1 = 4N / 9 \text{ (fourth prize)}.
\]

where \(N = 900 / 869\) is the normalization constant. You see that using this probability distribution you obtain the following normal average number of the prize money, \(\langle M \rangle\):

\[
\langle M \rangle = 0 \times p_1 + 100 \times p_2 + 500 \times p_3 + 1000 \times p_4,
\]

which is precisely your previously obtained average value, 227.272727... Yen.

Now, note that the above probability distribution has the form

\[
p_n = \frac{N}{(1 + n / 2)^2} \quad (n = 1, 2, 3, 4),
\]

which is simply rewritten in the form of the \(q\)-exponential distribution

\[
p_n = N e_q(-n)
\]
with $q = 3/2$. If you use the $q$-average with this $q$-exponential distribution as in [2], then you have

$$
\langle M \rangle_{q=3/2} = \frac{1}{\sum_{n=1}^{4} (p_n)^{3/2}} \left[ 0 \times \left( \frac{400}{869} \right)^{3/2} + 100 \times \left( \frac{225}{869} \right)^{3/2} + 500 \times \left( \frac{144}{869} \right)^{3/2} + 1000 \times \left( \frac{100}{869} \right)^{3/2} \right],
$$

which is $\langle M \rangle_{3/2} = 156.1015387...$ Yen.

A question is if you sell the lots with reference to $\langle M \rangle_{3/2} = 156.1015387...$ Yen. If you do so, then your pocket will seriously be damaged.

So far, you considered only the average. There is another important quantity, which is the variance or its square root, i.e., the standard deviation. Using the normal average, you obtain for the standard deviation, $\sigma = \sqrt{\langle M^2 \rangle - \langle M \rangle^2} = 327.7773887...$ Yen, whereas using the $q$-average, $\sigma_{q=3/2} = \sqrt{\langle M^2 \rangle_{3/2} - \langle M \rangle_{3/2}^2} = 282.0710854...$ Yen. Clearly, the latter is smaller than the former.

This is natural, since it is nothing but a general feature of equation (5): the larger the value of $f(q) (>1)$ is, the smaller the generalized standard deviation, $\sigma_{f(q)}$, is. That is, the $q$-average makes fluctuation small.

The above toy example indicates that the use of the $q$-average for the $q$-exponential distribution is in doubt.

### 3. Experiment

Imagine an experimentalist, who does not know/need nonextensive statistical mechanics. He/She does measurements of a certain physical quantity of a statistical-mechanical system and obtains a probability distribution of a positive random variable, $x$, that turns out to be well fitted by the Zipf-Mandelbrot distribution

$$
f(x) = \frac{A}{(1 + \lambda x)^\alpha}
$$

over a finite range, say $[a, b]$. Here, $\alpha$ and $\lambda$ are observed positive constants, and $A$ is a normalization factor. He/She may naively calculate the average value of a physical quantity, $Q(x)$, using the normal definition

$$
\langle Q \rangle = \int_a^b dx \ Q(x) \ f(x).
$$

There is nothing wrong here.

On the other hand, the distribution in equation (6) can be rewritten as a $q$-exponential distribution with the following identifications:

$$
\alpha = \frac{1}{q-1}, \quad \lambda = (q-1) \beta.
$$
In fact, with these identifications, equation (6) is expressed as follows:

$$ f(x) = A e_q(-\beta x). $$

(9)

Then, according to [2], one has to use, instead of equation (7), the following \(q\)-average:

$$ \langle Q \rangle_q = \frac{\int_a^b dx Q(x) [f(x)]^q}{\int_a^b dx' [f(x')]^q}. $$

(10)

The experimentalist will never understand why equation (7) cannot be employed, and the expression

$$ \langle Q \rangle_q = \frac{\int_a^b dx Q(x) [f(x)]^{1+1/\alpha}}{\int_a^b dx' [f(x')]^{1+1/\alpha}}. $$

(11)

has to be used.

Furthermore, suppose that the experimentalist does another experiment and obtains this time the following distribution of a positive random variable, \(x\):

$$ F(x) = B e^{-x_0/x \gamma}, $$

where \(x_0\), \(\kappa\), and \(\gamma\) are observed positive constants, and \(B\) is a normalization factor. Equation (12) asymptotically tends to the \(q\)-exponential distribution for large \(x\) but contains the extra exponential factor, here. Should the \(q\)-average with \(q = 1 + 1/\gamma\) be used following [2]? One may see no reasonable answers to this question for supporting the \(q\)-average.

4. Engineering

Suppose that an engineer, who does not know anything about nonextensive statistical mechanics, is analyzing diffusion in a structured medium. He/She is trying to model it by using the following nonlinear Fokker-Planck equation for a physical distribution \(p(x, t)\) of diffusing particles:

$$ \frac{\partial p(x, t)}{\partial t} = - \frac{\partial}{\partial x} \left[k x p(x, t)\right] + \frac{D}{2} \frac{\partial^2 p(x, t)}{\partial x^2}, $$

(13)

where \(k\) and \(D\) are positive constant and a generalized diffusion coefficient, respectively. He/She may find an analytic solution of this equation, which has the form

$$ p(x, t) = \frac{1}{Z(t)} \left[1 - (\nu - 1) B(t) (x - \xi(t))^2\right]^{1/(\nu-1)}, $$

(14)

where \(Z(t)\), \(B(t)\), and \(\xi(t)\) are somewhat involved. He/She is interested especially in the case when \(\nu < 1\), because equation (14) is an asymptotically power-law distribution in this case. Then, he/she may try to evaluate the shifted variance as follows:
\[ \langle (x - \xi)^2 \rangle(t) = \int_{-\infty}^{\infty} dx (x - \xi)^2 p(x, t). \quad (15) \]

The engineer will soon recognize that the variance in equation (15) is finite if and only if \( \nu > 1/3 \). But, after this recognition, he/she may be informed that the solution in equation (14) has already been obtained in [4], in which the following identification is made:

\[ \nu = 2 - q. \quad (16) \]

Then, the solution is written as

\[ p(x, t) = \frac{1}{Z(t)} e^q \left(-B(t)(x - \xi(t))^2\right). \quad (17) \]

Now, imagine that the engineer may look at the literature [1,2] and find that, in the field of nonextensive statistical mechanics, the distribution in equation (17) is termed the \( q \)-Gaussian distribution, which may/might maximize the \( q \)-entropy

\[ S_q[p] = \frac{1}{1-q} \left[ \int_{-\infty}^{\infty} dx [p(x, t)]^q \right]^q - 1. \quad (18) \]

under the constraints on the normalization condition and the shifted \( q \)-variance

\[ \langle (x - \xi )^2 \rangle_q(t) = \frac{\int_{-\infty}^{\infty} dx (x - \xi)^2 [p(x, t)]^q}{\int_{-\infty}^{\infty} dx' [p(x', t)]^q}, \quad (19) \]

where the time-dependence is assumed to be mild.

Of course, the engineer understands that the range of \( \nu \) for finite \( \langle (x - \xi)^2 \rangle_q \) becomes widened from \( 1/3 < \nu < 1 \) to \(-1 < \nu < 1\), if equation (19) is used. However, he/she will never understand why he/she cannot use equation (15) but must calculate

\[ \frac{\int_{-\infty}^{\infty} dx (x - \xi)^2 [p(x, t)]^{2-\nu}}{\int_{-\infty}^{\infty} dx' [p(x', t)]^{2-\nu}} \]

from his/her data, \( p(x, t) \), because basically his/her work is perfectly fine with the nonlinear Fokker-Planck equation (13) and does not need any entropic approach.

5. Quantum theory

Consider an isolated quantum system. Its state is represented by \( |\psi\rangle \), which can be expanded in terms of the complete set of eigenstates \( \{ |u_n\rangle \}_n \) of some observable as \( |\psi\rangle = \sum_n a_n |u_n \rangle \), where the expansion coefficients \( \{a_n \}_n \) satisfy the normalization condition: \( \sum_n a_n^* a_n = 1 \). A state of the system is described by the expansion coefficients, largely. For example, consider two states, \( |\psi\rangle \) and
\[ \left| \psi' \right\rangle . \text{ They have the following expansions: } \left| \psi \right\rangle = \sum_n \alpha_n \left| u_n \right\rangle \text{ and } \left| \psi' \right\rangle = \sum_n \alpha'_n \left| u_n \right\rangle. \] The difference between \( \left| \psi \right\rangle \) and \( \left| \psi' \right\rangle \) is reflected in the difference between the expansion coefficients.

The quantum-mechanical average of some observable, \( Q \), in the state \( \left| \psi \right\rangle \) is given by
\[ Q = \text{Tr} \left( Q \rho_0 \right) \]
\[ = \sum_{m,n} \alpha'_n \alpha_m \left( u_n \right) \left( u_m \right) \sum \left( u_n \right) \left( u_m \right). \]
This bilinear structure with respect to the expansion coefficients is radical for the probabilistic interpretation of quantum theory, i.e., the Born rule. One may also consider a pure-state density matrix,
\[ \rho_0 = \left| \psi \right\rangle \left( \psi \right). \]
Using this, one has
\[ Q_0 = \text{Tr} \left( Q \rho_0 \right). \]
Now, a crucial question is the following. Is the \( q \)-average,
\[ Q_q = \text{Tr} \left( Q \rho^q \right) / \text{Tr} \left( \rho^q \right), \]
consistent with quantum theory?

The point is that in the probabilistic interpretation of quantum theory the bilinear form \( a_n^* a_m \) is essential as mentioned above, and this bilinear structure must be respected in any quantum-mechanical calculation. Let us construct a mixed state using the most general linear race-preserving positive quantum operation
\[ \rho_0 = \left| \psi \right\rangle \left( \psi \right) \rightarrow \rho = \Phi \left( \rho_0 \right) = \sum_k V_k \rho_0 V_k^* + \sum_{n,m} \alpha_n^* \alpha_m \sum_k V_k \left( u_n \right) \left( u_m \right) V_k^*, \]
where the condition, \( \sum_k V_k^* V_k = I \), is imposed. The quantum-mechanical average is still in the bilinear form:
\[ \langle Q \rangle = \text{Tr} \left( Q \rho_0 \right) \rightarrow \text{Tr} \left( Q \rho \right) = \sum_{m,n} \alpha_n^* \alpha_m \left( u_n \right) \left( u_m \right). \]

On the other hand, \( \langle Q \rangle_q \) is obviously not in a bilinear form unless \( q = 1 \) and therefore is inconsistent with the Born rule in quantum mechanics.

### 6. Comment on the escort-distribution representation

The above stories certainly cause a serious problem for the researchers working on nonextensive statistical mechanics, since they want to establish a statistical-mechanical bridge between \( q \)-exponential distributions and the maximum \( q \)-entropy principle.

If the \( q \)-average is still to be employed, then the one and only way of resolving the difficulties may be to identify the observed \( q \)-exponential distribution with the escort distribution.

However, the situation turns out to remain hopeless. If an observed distribution is the escort distribution, \( P_n^{(q)} = \left( p_n \right)^q / \sum_{m=1}^{W} \left( p_m \right)^q \), then the value of the \( q \)-entropy should be calculated using it. The result for this is
\[ S_q = \frac{1}{1-q} \left[ \left( \sum_{n=1}^{W} \left( p_n^{(q)} \right)^{1/q} \right)^{-1/q} \right]. \]
Unfortunately, this is not an entropic quantity any more, since it fails to be concave with respect to \( \left\{ P_n^{(q)} \right\}_{n=1,2,...,W} \).
7. Farewell to the \( q \)-averages

These considerations lead to the conclusion that the \( q \)-average formalism suffers from fundamental conceptual difficulties and is not acceptable. One need employ the normal definition for averages and identify an observed physical distribution with the original distribution, \( \{ p_n \} \), not its associated escort distribution.

This conclusion is actually supported strongly by another recent physical discussion based on the generalized \( H \)-theorem [6].

8. Concluding remarks

We have shown the concept of the \( q \)-averages has fundamental difficulties. Once it has been discussed [7] that the Shore-Johnson theorem [8-10] leads to necessity of using the \( q \)-averages in nonextensive statistical mechanics. A point, which became clear today, is that the requirement of “subset independence” in the theorem is not realized in nonextensive statistical mechanics. Therefore, the basic premise in [7] is actually not fulfilled. Thus, the \( q \)-average formalism has lost its basis.

In addition, it has recently been shown [11-13] that the \( q \)-average is unstable (i.e., not uniformly continuous). In other words, its value can drastically changes even for very small changes of an observed distribution. Thus, it fails to possess experimental robustness.

Acknowledgement

The author was supported in part by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science.

References

[1] Tsallis C 2010 Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World (New York: Springer)
[2] Tsallis C, Mendes R S and Plastino A R 1998 Physica A 261 534
[3] Rosenkrantz R D (ed.) 1989 E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics (Dortrecht: Kluwer)
[4] Tsallis C and Bukman D J 1996 Phys. Rev. E 54 R2197
[5] Abe S 2000 Phys. Lett. A 275 250
[6] Abe S 2009 Phys. Rev. E 79 041116
[7] Abe S and Bagci G B 2005 Phys. Rev. E 71 016139
[8] Shore J E and Johnson R W 1980 IEEE Trans. Inf. Theory IT-26 26
[9] Shore J E and Johnson R W 1981 IEEE Trans. Inf. Theory IT-27 472
[10] Shore J E and Johnson R W 1983 IEEE Trans. Inf. Theory IT-29 942
[11] Abe S 2008 EPL 84, 60006
[12] Abe S 2009 J. Stat. Mech. P07027
[13] El Kaabouchi A and Abe S 2011 J. Appl. Math. 804836