Note on generalization of Shannon theorem and inequality

1998 J. Phys. A: Math. Gen. 31 8159

(http://iopscience.iop.org/0305-4470/31/40/010)

J. Phys. A: Math. Gen. **31** (1998) 8159–8164. Printed in the UK PII: S0305-4470(98)94646-7

Note on generalization of Shannon theorem and inequality

Kwok Sau Fa

Departamento de Física, Universidade Estadual de Maringá, Campus Universitário, Av. Colombo, 5790, Maringá, PR, Brazil

Received 29 May 1998

Abstract. Recently, Santos obtained a generalized entropy using four assumptions which stated that an entropy must: (i) be a continuous function of the probabilities $\{p_i\}$; (ii) be a monotonic increasing function of the number of states W, in the case of equiprobability; (iii) satisfy $S_q^T(A+B)/k = S_q^T(A)/k + S_q^T(B)/k + (1-q)S_q^T(A)S_q^T(B)/k^2$ (where A and B are two independent systems) and (iv) satisfy the relation $S_q^T(\{p_i\}) = S_q^T(\{p_L, p_M\}) + p_L^q S_q^T(\{p_i/p_L\}) + p_M^q S_q^T(\{p_i/p_M\})$, where $p_L + p_M = 1$ ($p_L = \sum_{i=1}^{W_L} p_i$ and $p_M = \sum_{i=W_L}^{W} p_i$). Santos showed that the only function which satisfies all of these properties is the generalized Tsallis entropy. In this paper we perform a similar analysis and we obtain a family of entropies which are equivalent to the Tsallis entropy. We also discuss the Shannon inequality in the context of the generalized Tsallis entropy.

1. Introduction

Recently, Santos [1] obtained a generalized entropy using four assumptions stating that an entropy must (i) be a continuous function of the probabilities $\{p_i\}$; (ii) be a monotonic increasing function of the number of states W, in the case of equiprobability; (iii) satisfy $S_q^T(A+B)/k = S_q^T(A)/k + S_q^T(B)/k + (1-q)S_q^T(A)S_q^T(B)/k^2$ (A and B being two independent systems) and (iv) satisfy the relation $S_q^T(\{p_i\}) = S_q^T(\{p_L, p_M\}) + p_L^q S_q^T(\{p_i/p_L\}) + p_M^q S_q^T(\{p_i/p_M\})$, where $p_L + p_M = 1$ ($p_L = \sum_{i=1}^{W_L} p_i$ and $p_M = \sum_{i=W_L}^{W} p_i$). Santos showed that the only function satisfying all of these properties is the generalized Tsallis entropy [2]

$$S_q^T = k\frac{1-\sum p_i^q}{q-1} \tag{1.1}$$

where k is a positive constant and q is a real number. For $q \to 1$ we recover Shannon entropy

$$S_{q\to 1}^T = -k\sum_i p_i \ln p_i. \tag{1.2}$$

Our purpose here is to perform a similar analysis, however, we employ a general parameter a in the non-extensive entropy property ($S(AB) = S(A) + S(B) + aS(A)S(B)$, where A and B are two distinct systems). We show that by taking a as a function of q, it is possible to obtain a family of entropies, equivalent to Tsallis entropy. We also discuss the general behaviour of S_q^T for two interacting systems. More precisely, we obtain the generalized Shannon inequality and we discuss its physical implications.

2. Generalized entropy

To begin the analysis, we first list the assumptions as given in Landsberg [3] (except that we modify the third one).

(I) The accessible quantum states of an isolated equilibrium system are equally probable, i.e. the probabilities $\{p_i\}$ have the same values $\{1/n\}$, where n is the total number of accessible states of the system. The entropy of the system approaches the equilibrium state, which cannot decrease

$$S(p_1, \ldots, p_n) \leqslant S(1/n, \ldots, 1/n). \tag{2.1}$$

(II) The addition of an inaccessible state ($p_i = 0$) cannot affect the entropy

$$S(p_1, \ldots, p_n, 0) = S(p_1, \ldots, p_n). \tag{2.2}$$

(III.I) Consider two arbitrary non-interacting systems A and B with $p_{ij}^{A\cup B} = p_i^A p_j^B$. The entropy of a composite system satisfies $S(AB) = S(A) + S(B) + aS(A)S(B)$, where a is a real number. If A and B are identical, then $S(AB) = 2S(A) + a[S(A)]^2$.

(III.II) In the case of interacting systems A and B (with the probabilities $p_1, \ldots, p_n$ for A and the probabilities $q_{i1}, q_{i2}, \ldots$ for B), the composite system satisfies $S(AB) = S(A) + S_A(B) + aS(A)S_A(B)$, where $S_A(B)$ is the normalized mean conditional entropy given by

$$S_A(B) = \frac{\sum_i p_i^d S_i(B)}{\sum_j p_j^d} \tag{2.3}$$

and d is a real number.

If the systems A and B do not interact, $q_{i1}, q_{i2}, \ldots$ are independent of i, hence $S_A(B) = \frac{\sum_i p_i^d S_i(B)}{\sum_j p_j^d} = \frac{S(B)\sum_i p_i^d}{\sum_j p_j^d} = S(B)$.

These last two assumptions deserve some comments. In (III.I) we have introduced $S(AB)$ with a general parameter a and we also assume that it does not depend on the systems A and B. The last term of this relation represents a breakdown of the extensive property of standard theory, except for $a \to 0$ (we will show that we can recover the form of standard entropy for the composite system). Moreover, the use of this general parameter opens the possibility of studying different forms of generalized entropies. In (III.II) we have the generalization of normalized mean conditional entropy with the parameter d. Certainly, $S_A(B)$ depends on both the probabilities and the parameter d. We will show that the parameter a along with d are important to obtain new forms of generalized entropies. For instance, we will show that if a tends to zero, d must tend to 1 (and vice versa) and so we recover the standard entropy and the usual mean value. Therefore, to obtain generalized entropies we must take $d \neq 1$ and $a \neq 0$.

Now, in order to obtain the generalized entropies we divide the problem into two steps. In the first step we use assumptions (I), (II) and (III.I) for the non-interacting systems to prove that $S(1/n, \ldots, 1/n) = \frac{1}{a}[-1 + n^b]$. The entropy at the equilibrium system is a non-decreasing function of its arguments

$$L(n) \equiv S(1/n, \ldots, 1/n, 0) \leqslant S(1/(n+1), \ldots, 1/(n+1)) = L(n+1). \tag{2.4}$$

Next, we consider m mutually independent schemes $A^{(j)}$ $(j = 1, 2, \ldots, m)$, each consisting of r equally likely events. If we consider them as a single scheme we have r^m equally likely events with entropy $L(r^m)$. If we now consider them as a product scheme the entropy is calculated (using (III.I)) as follows.

For $m = 2$, it is immediately verified that

$$L(r^2) = 2L(r) + a[L(r)]^2 = \frac{1}{a}[(1 + aL(r))^2 - 1]. \tag{2.5}$$

For $m = 3$, the product scheme can be calculated in the following way. We consider these three schemes as independent systems and each of them has r equally likely events. This way, we can first calculate the entropy by choosing any two of these systems and the result is given by (2.5). Then, we use this result with the third system and assumption (III.I) to obtain that

$$\begin{aligned} L(r^3) &= \frac{1}{a}[(1 + aL(r))^2 - 1] + L(r) + a \cdot L(r) \cdot \frac{1}{a}[(1 + aL(r))^2 - 1] \\ &= \frac{1}{a}[(1 + aL(r))^3 - 1]. \end{aligned}$$

For all positive integers r and m, we use the induction method, and we have

$$L(r^m) = \frac{1}{a}[(1 + aL(r))^m - 1]. \tag{2.6}$$

This is the generalization of standard entropy, i.e. in the limit $a \to 0$ we recover the well known result $L(r^m) = mL(r)$.

It is easy to guess that the following function satisfies (2.6)

$$L(r) = \frac{1}{a}[-1 + r^b] \tag{2.7}$$

where b is a constant.

It is noteworthy that, to recover the additivity property for entropy, we must take both, a and b from equation (2.7) tending to zero, simultaneously.

To prove (2.7) we assume r, s and n are arbitrary positive integers with m determined by

$$r^m \leqslant s^n \leqslant r^{m+1}. \tag{2.8}$$

By applying the logarithm in (2.8) we have

$$\frac{m}{n} \leqslant \frac{\ln s}{\ln r} \leqslant \frac{m+1}{n}. \tag{2.9}$$

On the other hand, we write $L(x)$ for each term x of (2.8), i.e. $L(r^m) \leqslant L(s^n) \leqslant L(r^{m+1})$. Then, using relation (2.6), we obtain

$$\frac{m}{n} \leqslant \frac{\ln(1 + aL(s))}{\ln(1 + aL(r))} \leqslant \frac{m+1}{n}. \tag{2.10}$$

Combining (2.9) and (2.10), we find that

$$\left| \frac{\ln(1 + aL(s))}{\ln(1 + aL(r))} - \frac{\ln s}{\ln r} \right| \leqslant \frac{1}{n}. \tag{2.11}$$

Since n is arbitrary, we can take n arbitrarily high value, and we obtain

$$\frac{\ln(1 + aL(s))}{\ln(s)} = \frac{\ln(1 + aL(r))}{\ln(r)} = b \tag{2.12}$$

where b is a constant independent of s and r. Therefore,

$$L(n) = \frac{1}{a}[-1 + n^b] \tag{2.13}$$

for all n. This is in accordance with equation (2.7). From equation (2.4), we immediately verify that $b \geqslant 0$.

In the next step, we consider two interacting systems A and B to prove that $S(p_1, \ldots, p_n) = \sum_i (p_i^d - 1)/a$. To do so, we suppose two interacting probability schemes A and B. Consider that the n probabilities of A giving by $p_i = g_i/g$, where the g's are positive integers with $\sum_i g_i = g$. Consider that a dependent scheme B has n groups of events so that the ith group has g_i events. If event A_i occurs, we assume that all events of the ith group are equally likely with probability $1/g_i$. Hence, the conditional entropy $S_i(B)$ of B, given A is in state A_i, is obtained by (2.13), replacing n by g_i

$$S_i(g_i) = \frac{1}{a}[-1 + g_i^b]. \tag{2.14}$$

By using (III.II) we have

$$S_A(B) = \frac{\sum_i p_i^d S_i(B)}{\sum_j p_j^d} = \frac{\sum_i p_i^d(-1 + g_i^b)}{a\sum_j p_j^d}. \tag{2.15}$$

However, the entropy of a composite system at the equilibrium state consists of g equally likely events

$$S(AB) = \frac{1}{a}[-1 + g^b]. \tag{2.16}$$

It follows from (III.II) that

$$S(A) = \frac{S(AB) - S_A(B)}{1 + aS_A(B)} = \frac{\sum_i (p_i^d - p_i^{b+d})}{a\sum_j p_j^{b+d}}. \tag{2.17}$$

As we can see, equation (2.17) contains three arbitrary constants a, b and d; to obtain the mean entropy we should take $b + d = 1$. Consequently,

$$S(A) = \frac{\sum_i p_i^d - 1}{a} \tag{2.18}$$

where we have used the condition $\sum_i p_i = 1$.

We should note that if we take the parameter d in (2.18) tending to 1 we must take a tending to zero and so we recover the standard entropy. As a consequence, to obtain the generalized entropies with a non-extensive property given by (III.I) we need to use the unusual average, i.e. it should have $d \neq 1$ and $a \neq 0$.

Of course, the entropy (2.18) can assume many different forms depending on the expressions of a and d. For example, for $d = q$ and $a = (1 - q)/k$ we obtain $b = 1 - q$ and $q \leqslant 1$, hence S becomes

$$S_q^T = k\frac{1 - \sum_i p_i^q}{q - 1}. \tag{2.19}$$

This is equal to equation (1.1). Moreover, S_q^T may be rewritten as follows

$$S_q^T = \frac{\sum_i p_i^q (p_i^{1-q} - 1)}{(q - 1)/k} = \sum_i p_i^q S_{qi}^T(p_i) \tag{2.20}$$

and

$$S_{qi}^T(p_i) = -k(1 - p_i^{1-q})/(q - 1). \tag{2.21}$$

So, we may interpret S_q^T as an average (or mean) that has been taken over a probability distribution $\{p_i^q\}$. In addition, it is important to emphasize that the form of this average is in agreement with the mean internal energy $U_q = \sum_i p_i^q \varepsilon_i$ obtained by Curado and Tsallis [4] which they have used to obtain the connection with the thermodynamics.

Now, for $d = q$ and $a = 2^{1-q} - 1$ we obtain the Daróczi entropy [5]

$$S_q^D = \frac{\sum_i p_i^q - 1}{2^{1-q} - 1}. \tag{2.22}$$

These entropies S_q^T and S_q^D are equivalent because they only differ by a multiplicative constant. Furthermore, we should note that in both cases, a and b tend to zero for $q \to 1$, simultaneously. This means that the extensivity limit is preserved for both entropies.

Certainly, we can make other kinds of entropies by taking different forms of $a(q)$ and $d(q)$. However, it is easy to show that they can be transformed from one to another and the differences among them are a multiplicative constant. Therefore, they constitute a family of equivalent generalized entropies.

3. General behaviour of S_q^T for two interacting systems

In this section we address the question of the general behaviour of the generalized Tsallis entropy for two interacting systems. In particular, we analyse the problem of decreasing Tsallis entropy related to the additional information on the system. To do so, let us use the following inequality, valid for any continuous strictly monotonic convex function $\phi(x)$ (see [6, theorem 86]).

$$\phi\left(\sum_i a_i x_i\right) \leqslant \sum_i [a_i \phi(x_i)] \tag{3.1}$$

where $a_i \geqslant 0$ and $x_i \geqslant 0$ such that $\sum_i x_i = 1$.

Now, for Tsallis entropy we consider $\phi(x_i)$ given by

$$\phi(x) = \frac{x - x^q}{1 - q} \qquad \text{for } q > 0. \tag{3.2}$$

From this and equation (3.1) we have that

$$\frac{g_j - g_j^q}{1 - q} = \frac{\sum_i p_i^q g_{ij} - (\sum_i p_i^q g_{ij})^q}{1 - q} \leqslant \sum_i \frac{p_i^q (g_{ij} - g_{ij}^q)}{1 - q} \tag{3.3}$$

where g_j is as the total probability $\sum_i p_i^q g_{ij}$ of finding the event B_j in system B. Now, adding over j and multiplying by $-k$ we immediately find that

$$S_A(B) \leqslant S(B). \tag{3.4}$$

This inequality can be easily extended to

$$S(AB) = S(A) + S_A(B) + (1 - q)S(A)S_A(B) \leqslant S(A) + S(B) + (1 - q)S(A)S(B). \tag{3.5}$$

Therefore, we have obtained the remarkable result that the additional information on the system decreases the entropy. The well known example of this result can be associated to Maxwell's demon, which is capable of decreasing entropy by using the information of the system without the performance of work.

4. Conclusion

Therefore, we have shown that in order to obtain the generalized entropies we have modified, in a smooth way, one of the three assumptions given in Landsberg [3]. In particular, we have replaced the additivity property $S(AB) = S(A) + S(B)$ by $S(AB) = S(A) + S(B) + aS(A)S(B)$ and the normalized mean conditional entropy $S_A(B) = \sum_i p_i S_i(B)$ by

$$S_A(B) = \frac{\sum_i p_i^d S_i(B)}{\sum_j p_j^d}$$

and we have also used the basic concepts of thermostatistics. By using the three assumptions (I), (II) and (III) we have shown that there is a family of equivalent generalized entropies including the cases of Tsallis and Daróczi entropies. In addition, we have obtained the mean entropy (2.20) which was taken over the probability distribution $\{p_i^q\}$. This is in accordance with the mean internal energy $U_q = \sum_i p_i^q \varepsilon_i$ obtained by Curado and Tsallis [4] which they have used to obtain the connection with the thermodynamics. Therefore, this approach is very close to the development of the standard statistics†.

We have also discussed the general behaviour of S_q^T for two interacting systems. It is shown by (3.5) that the entropy decreases with the additional information on the system. This means that Maxwell's demon also acts on S_q^T.

References

[1] Santos R J V 1997 *J. Math. Phys.* **38** 4104
[2] Tsallis C 1988 *J. Stat. Phys.* **52** 479
[3] Landsberg P T 1990 *Thermodynamic and Statistical Mechanics* (New York: Dover) section 9.1
[4] Curado E M F and Tsallis C 1991 *J. Phys. A: Math. Gen.* **24** L69 (corrigenda 1991 *J. Phys. A: Math. Gen.* **24** 3187 and 1992 **25** 1019)
[5] Daróczi Z 1970 *Inf. Control* **16** 36
[6] Hardy G H, Littlewood J E and Pólia G 1973 *Inequalities* (Cambridge: Cambridge University Press)
[7] Tsallis C 1994 *Phys. Lett.* A **195** 329

† An interesting analysis with Tsallis entropy which may be put in a wider context between generalized statistics and quantum groups is given in [7].