Maximum entropy approach to stretched exponential probability distributions

C Anteneodo†§ and A R Plastino‡∥

† Instituto de Biofísica, Universidade Federal do Rio de Janeiro, CCS G026, Cidade Universitária, Ilha do Fundão, CEP 21949-900, Rio de Janeiro, Brazil
‡ Centro Brasileiro de Pesquisas Físicas/CNPq, R. Xavier Sigaud 150, CEP 22290-180, Rio de Janeiro, Brazil

Received 14 August 1998

Abstract. We introduce a nonextensive entropy functional $S_\eta$ whose optimization under simple constraints (mean values of some standard quantities) yields stretched exponential probability distributions, which occur in many complex systems. The new entropy functional is characterized by a parameter $\eta$ (the stretching exponent) such that for $\eta = 1$ the standard logarithmic entropy is recovered. We study its mathematical properties, showing that the basic requirements for a well-behaved entropy functional are verified, i.e. $S_\eta$ possesses the usual properties of positivity, equiprobability, concavity and irreversibility and verifies Khinchin axioms except the one related to additivity since $S_\eta$ is nonextensive. The entropy $S_\eta$ is shown to be superadditive for $\eta < 1$ and subadditive for $\eta > 1$.

1. Introduction

The study of the mathematical properties and physical applications of new formulations of the maximum entropy (ME) principle based on generalized or alternative entropic measures constitutes a currently growing field of research in statistical physics [1–3]. This line of inquiry has been greatly stimulated by the work of Tsallis [4], who developed a complete and consistent thermostatistical formalism on the basis of a generalized nonextensive entropic functional.

The ME principle, introduced by Jaynes on the basis of Shannon’s information measure, is a powerful tool widely used in many areas of both experimental and theoretical science. Jaynes advanced this principle as a new foundation for Boltzmann–Gibbs statistical mechanics [5]. However, nowadays its range of applications embraces a variety of applied fields such as image reconstruction and other inverse problems with noisy and incomplete data [6], time series analysis [7] and the approximate solution of partial differential equations [8].

Despite the great success of the standard ME principle, it is a well known fact that there are many relevant probability distributions in nature which are not easily derivable from the Jaynes–Shannon prescription. Lévy distributions constitute an interesting example showing such difficulties. If one sticks to the standard logarithmic entropy, awkward constraints are needed in order to obtain Lévy-type distributions [9]. However, Jaynes ME principle suggests in a natural way the possibility of incorporating alternative entropy functionals to the variational principle. Actually, there exist many physical scenarios where the standard

§ E-mail address: celia@cat.cbpf.br
∥ Author for correspondence. Permanent address: Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, Paseo del Bosque s/n, 1900 La Plata, Argentina.
A statistical description based on the Boltzmann–Gibbs–Shannon entropy fails, such as self-gravitating systems [10], electron-plasma two-dimensional turbulence [11], and self-organized criticality [12], among many others (see [13]). An important feature shared by all these systems is their nonextensive behaviour, suggesting that a nonextensive (nonadditive) entropy functional might be appropriate for their thermostatistical description.

Indeed, it has been recently noted that many of the above-mentioned problems involve a family of probability distributions derivable from the extremalization of a generalized nonextensive entropy measure recently introduced by Tsallis [1]. This entropy, defined by

\[ S^T_q = \sum_{i=1}^w p_i - p_i^q \frac{q}{q-1} \]

is characterized by a real parameter \( q \) associated with the degree of nonextensivity. \( S^T_q \) coincides with the standard logarithmic entropy in the limit \( q \to 1 \). The optimization of Tsallis entropy under simple constraints yields power law probability distributions while, in the limit \( q \to 1 \), the simple exponential law is recovered. The new thermostatistical formalism derived from the \( S^T_q \) entropy has been shown to consistently generalize the relevant properties of the Boltzmann–Gibbs statistics [1, 14]. As a consequence, many interesting physical applications have already been worked out [15]. Besides the fundamental aspects, there are implications of practical relevance such as its application to the solution of partial differential equations [8] or to optimization problems [16].

To a great extent, the success of Tsallis proposal is due to the ubiquity of power law distributions in nature. However, other important families of distributions, the stretched exponential ones, are also frequent in complex systems. Stretched exponential probability distributions appear, for instance, in the description of turbulent flows [17]. Many other examples of these distributions in nature and economy are listed by Laherrere and Sornette [18]. These authors also call attention to the possibility that distributions usually classified as power laws may actually correspond to stretched exponential ones.

On the other hand, anomalous slow relaxations in disordered systems (glassy systems, quasicrystals, polymers, strongly interacting materials, etc) [19] often follow the stretched exponential form. The anomalous decay of the density of species in diffusion-controlled reactions may follow power laws as well as stretched exponential ones [20]. These latter examples do not involve probability distributions of the stretched exponential form directly. However, on the basis of extreme deviations in random multiplicative processes, stretched exponential distributions can be applied to rationalize stretched exponential relaxations [21].

All these considerations suggest the possibility that new useful entropy functionals may be lurking behind stretched exponential probabilities. Thus, our present goal is to explore the properties of a new nonextensive entropy functional whose optimization yields such probability distributions. A full understanding of the generic mechanisms underlying stretched exponential laws, which up to now remain mainly at the phenomenological level, is still lacking. Hence, the study of a variational approach to these ubiquitous functions may also throw new insights onto their physical origin.

This paper is organized as follows. In section 2 we present the new entropy \( S_\eta \). In sections 3 and 4 we show that \( S_\eta \) verifies the usual requirements for a mathematically well defined entropy functional. We illustrate these properties by means of two-state systems in section 5. Finally, section 6 contains some final remarks.
2. The new entropy functional

Let us begin by defining the following entropy functional $S_\eta$ associated with a given discrete probability distribution $\{p_i, i = 1, \ldots, w\}$ (the extension to the continuous case is straightforward)

$$S_\eta = \sum_{i=1}^{w} s_\eta(p_i) \quad (2)$$

where

$$s_\eta(p_i) \equiv \frac{1}{\eta} \left[ \Gamma \left( \frac{\eta + 1}{\eta}, -\ln p_i \right) - p_i \Gamma \left( \frac{\eta + 1}{\eta} \right) \right]. \quad (3)$$

Here, $\eta$ is a positive real number, $0(\mu, t) = \int_{t}^{\infty} \exp(-\mu y) \, dy = \int_{0}^{\exp(-t)} [-\ln x]^{\mu-1} \, dx \quad \mu > 0 \quad (4)$

is the complementary incomplete Gamma function, and $\Gamma(\mu) = \Gamma(\mu, 0)$ the Gamma function. By recourse to the definition (3), it is easy to verify that in the case $\eta = 1$ the standard entropy $S_1 = -\sum_{i=1}^{w} p_i \ln p_i$ is recovered. This particular choice of $s_\eta(p_i)$ will soon become clear.

Optimization of $S_\eta$ under the following constraints

$$\sum_{i=1}^{w} p_i = 1 \quad (5)$$

$$\langle O^{(r)} \rangle \equiv \sum_{i=1}^{w} p_i O^{(r)}_i = O^{(r)}_\eta \quad (r = 1, \ldots, R) \quad (6)$$

where $\{O^{(r)}\}$ are observables and $\{O^{(r)}_\eta\}$ are finite known quantities, yields

$$p_i = \exp \left( - \left[ \Gamma \left( \frac{\eta + 1}{\eta} \right) + \alpha + \sum_{r=1}^{R} \beta_r O^{(r)}_\eta \right]^{\eta} \right) \quad i = 1, \ldots, w \quad (7)$$

where $\alpha$ and $\{\beta_r\}$ are the Lagrange multipliers associated to the constraints (5) and (6), respectively. So, the optimization of $S_\eta$ constrained under fixed mean values of relevant quantities $O^{(r)}$ yields stretched exponentials of the form given by expression (7).

Actually, the function $s_\eta(p_i)$ was found by precisely requiring the probabilities $p_i$, arising from the stated variational problem, to be of the form (7). That is to say we have solved the inverse problem of obtaining the entropy functional from a given maximum entropy probability distribution. This procedure could be applied to arbitrary classes of probability distributions but here we are interested in the stretched exponential one.

3. Khinchin axioms

Khinchin proposed a set of four axioms [22], which are usually regarded as reasonable requirements for a well behaved information measure. Our entropy measure $S_\eta$ verifies the first three of them:

(i) $S_\eta = S_\eta(p_1, \ldots, p_w)$, i.e. the entropy is a function of the probabilities $p_i$ only.
(ii) $S_\eta(p_1, \ldots, p_w) \leq S_\eta(\frac{1}{w}, \ldots, \frac{1}{w}) \equiv S_\eta^{\text{equivr}}(w)$, i.e. $S_\eta$ adopts its extreme at equiprobability (this property will be proved in section 4).
(iii) $S_\eta(p_1, \ldots, p_w) = S_\eta(p_1, \ldots, p_w, 0)$ this property, known as expansibility, is clearly verified since $s_\eta(0) = 0$.
(iv) The fourth Khinchin axiom concerns the behaviour of the entropy of a composite system in connection to the entropies of the subsystems. We will comment on this axiom later.
4. General mathematical properties

Let us consider other interesting properties related to positivity, certainty, concavity, equiprobability, additivity and irreversibility.

4.1. Positivity

It is plain from equation (3) that \( s_\eta(0) = s_\eta(1) = 0 \) and also that

\[
\frac{d^2 s_\eta}{dp_i^2} = -\frac{1}{\eta} \frac{[\ln p_i]^{\frac{1}{\eta} - 1}}{p_i} < 0 \quad \text{for} \quad 0 < p_i < 1.
\]

Consequently, \( s_\eta(p_i) \) is a positive quantity for \( p_i \in (0, 1) \). This, in turn, implies the positivity condition:

\[
S_\eta \geq 0.
\]

4.2. Certainty

The equality symbol in equation (9) holds only at certainty, i.e.

\[
S_\eta(1, 0, \ldots, 0) = 0.
\]

Indeed, \( S_\eta \) vanishes if and only if we have certainty.

4.3. Concavity

Considering \( (p_1, \ldots, p_w) \) as independent variables, the second partial derivatives of \( S_\eta \) are

\[
\frac{\partial^2 S_\eta}{\partial p_j \partial p_k} = -\frac{1}{\eta} \frac{[\ln p_j]^{\frac{1}{\eta} - 1}}{p_j} \delta_{jk} < 0 \quad \text{for} \quad 0 < p_j < 1.
\]

Now, if we incorporate the constraint \( \sum_{i=1}^w p_i = 1 \), as it defines a convex set, expression (11) guarantees definite concavity over probability space.

4.4. Equiprobability

Taking into account the normalization condition, let \( p_w \) be the dependent probability. In that case the first derivatives of \( S_\eta \) are

\[
\frac{\partial S_\eta}{\partial p_j} = [-\ln p_j]\frac{1}{\eta} - [-\ln p_w]\frac{1}{\eta} \quad \forall j \neq w.
\]

Therefore, the first derivatives vanish for \( p_j = p_w, \forall j \). Since \( S_\eta \) has negative concavity, then it is maximal at equiprobability.

A well-behaved entropy should also be, at equiprobability, a monotonically increasing function of the number of states \( w \). We will show that \( S_\eta \) verifies this property. From the definition of \( S_\eta \),

\[
S_\eta^{\text{equipr.}}(w) = w \Gamma \left( \frac{\eta + 1}{\eta}, \ln w \right) - \Gamma \left( \frac{\eta + 1}{\eta} \right)
\]

then

\[
\frac{dS_\eta^{\text{equipr.}}}{dw} = \Gamma \left( \frac{\eta + 1}{\eta}, \ln w \right) - \frac{[\ln w]^{\frac{1}{\eta}}}{w}.
\]
By rewriting equation (14) as
\[
\frac{dS_{\text{eqi}}} {dw} = \int_{\ln w}^{\infty} y^{\frac{1}{\eta}} e^{-y} dy + \int_{\ln w}^{\infty} \frac{d(y^{\frac{1}{\eta}} e^{-y})} {dy} dy
\]
(15)
it is clear that
\[
\frac{dS_{\text{eqi}}} {dw} = \frac{1}{\eta} \int_{\ln w}^{\infty} y^{\frac{1}{\eta} - 1} e^{-y} dy = \frac{1}{\eta} \Gamma\left(\frac{1}{\eta}, \ln w\right)
\]
(16)
which is a positive quantity (recall that \(\eta > 0\)). Therefore, \(S_{\text{eqi}}(w)\) is an increasing function of \(w\).

4.5. Nonextensivity

We now analyse the relation of the entropy of a composite system with those of its subsystems.

Let us consider systems \(A\) and \(B\) with associated probabilities \(\{a_i, i = 1, \ldots, w_A\}\) and \(\{b_j, j = 1, \ldots, w_B\}\), respectively. If systems \(A\) and \(B\) are independent, i.e. system \(A \oplus B\) has associated probabilities \(\{a_i b_j; i = 1, \ldots, w_A; j = 1, \ldots, w_B\}\), then the entropy \(S_{\eta}(A \oplus B)\) of the composite system minus those of the subsystems, following the definition of \(S_{\eta}\), is
\[
\Delta S_{\eta}(A, B) = S_{\eta}(A \oplus B) - S_{\eta}(A) - S_{\eta}(B)
\]
(17)
with \(f_{\eta}(p) \equiv [-\ln p]^{1/\eta}\). After some manipulations, one gets
\[
\Delta S_{\eta}(A, B) = \sum_{i=1}^{w_A} \sum_{j=1}^{w_B} a_i b_j \int_{0}^{1} dp \, f_{\eta}(p) F_{\eta}(p, a_i, b_j)
\]
(18)
where
\[
F_{\eta}(p, a_i, b_j) \equiv \left(1 + \left[1 + \frac{\ln a_i}{\ln p} + \frac{\ln b_j}{\ln p}\right]^{1/\eta} - \left[1 + \frac{\ln a_i}{\ln p}\right]^{1/\eta} - \left[1 + \frac{\ln b_j}{\ln p}\right]^{1/\eta}\right).
\]
(19)
If \(\eta = 1\), then \(F_{\eta}(p, a_i, b_j) = 0\) and the additivity property of the standard entropy \(S_1(A \oplus B) = S_1(A) + S_1(B)\) is recovered. However, if \(\eta < 1\), \(F_{\eta}(p, a_i, b_j) = 0\) only if \(a_i = 1\) or \(b_j = 1\) (certainty for at least one of the subsystems), otherwise, \(F_{\eta}(p, a_i, b_j) > 0\). In fact, let us consider the function \(G(x, y) = 1 + [1 + x + y]^{1/\eta} - [1 + x]^{1/\eta} - [1 + y]^{1/\eta}\). Since \(G(0, y) = 0\), it is apparent that for \(x, y > 0\) and \(\eta < 1\), \(\partial G / \partial x > 0\) and, therefore, \(G(x, y) > 0\). An analogous reasoning is valid for \(\eta > 1\). In this latter case \(F_{\eta}(p, a_i, b_j) < 0\). Consequently, \(S_{\eta}(A \oplus B)\) is superadditive for \(\eta < 1\) and subadditive for \(\eta > 1\). The nonadditivity of \(S_{\eta}\) for \(\eta \neq 1\) reflects the nonextensivity of composite systems.

4.6. Irreversibility

One of the most important roles played by entropic functionals within theoretical physics is to characterize the ‘arrow of time’. When they verify an \(H\)-theorem, they provide a quantitative measure of macroscopic irreversibility. We will now show, for some simple systems, that the present measure \(S_{\eta}\) satisfies an \(H\)-theorem, i.e. its time derivative has a definite sign.
Let us calculate the time derivative of $S_\eta$

$$\frac{dS_\eta}{dt} = \sum_{i=1}^{w} [-\ln p_i]^{\frac{\eta}{\eta+1}} \frac{dp_i}{dt}$$

(20)

for a system whose probabilities $p_i$ evolve according to the master equation

$$\frac{dp_i}{dt} = \sum_{j=1}^{w} [P_{ij} p_j - P_{ij} p_i]$$

(21)

where $P_{ij}$ is the transition probability per unit time between microscopic configurations $i$ and $j$. Assuming detailed balance, i.e. $P_{ij} = P_{ji}$, we obtain from (20)

$$\frac{dS_\eta}{dt} = \frac{1}{2} \sum_{i=1}^{w} \sum_{j=1}^{w} P_{ij} (p_i - p_j) \left([-\ln p_j]^{\frac{\eta}{\eta+1}} - [-\ln p_j]^{\frac{\eta}{\eta+1}} \right).$$

(22)

In each term of the above expression, both factors involving $p_i$ and $p_j$ have the same sign for $\eta > 0$, then we obtain

$$\frac{dS_\eta}{dt} \geq 0.$$  

(23)

The equality holds for equiprobability, i.e. at equilibrium, while in any other cases the entropy $S_\eta$ increases with time. Therefore, $S_\eta$ exhibits irreversibility.

### 4.7. Jaynes thermodynamic relations

It is noteworthy that, within the present ME formalism, the usual thermodynamical relations involving the entropy, the relevant mean values, and the associated Lagrange multipliers, i.e.

$$\frac{\partial S_\eta}{\partial \langle O(r) \rangle} = \beta_r$$

(24)

are verified. Hence, our formalism exhibits the usual thermodynamical Legendre transform structure. Actually, this property is verified by a wide family of entropy functionals [2].

### 5. Two-state systems

In order to illustrate some of the above properties, we consider a two-state system (with associated probabilities $\{p, 1-p\}$). In this case, $S_\eta$ depends only on the variable $p$. In fact, from its definition, we have

$$S_\eta(p) = \Gamma\left(\frac{\eta+1}{\eta}, -\ln p\right) + \Gamma\left(\frac{\eta+1}{\eta}, -\ln[1-p]\right) - \Gamma\left(\frac{\eta+1}{\eta}\right).$$

(25)

The shape of $S_\eta(p)$ for different values of $\eta$ is shown in figure 1, which exhibits the positivity and concavity of $S_\eta$. In fact, from expression (25), the first derivative of $S_\eta(p)$ vanishes at $p = \frac{1}{2}$ and $d^2 S_\eta/dp^2 < 0 \forall p$. Since the second derivative is always negative, $S_\eta(p)$ is maximal at equiprobability. Moreover, as shown in the general case, taking into account the concavity of $S_\eta$ and that $S_\eta$ vanishes at the certainty, then $S_\eta$ is positive for all $p$.

The subadditivity and superadditivity of $S_\eta$ is illustrated in figure 2(a) for two identical and independent two-state systems $X$, through the plot of the relative difference $(\Delta S_\eta)_{\text{rel.}} = [S_\eta(X \oplus X) - 2S_\eta(X)]/S_\eta(X \oplus X)$ between the entropy of the composite system and those of the subsystems as a function of $p$. Figure 2(b) exhibits the behaviour of $(\Delta S_\eta^\text{eqpr.})_{\text{rel.}}$ for two-state systems, at equiprobability, as a function of $\eta$. This behaviour is qualitatively representative of that of any two systems with arbitrary number of states at equiprobability.
6. Final remarks

We have shown that the entropy functional $S_\eta$, inspired in stretched exponential probabilities, verifies the main properties usually regarded as essential requirements for useful entropy functionals. The entropy $S_\eta$ verifies the first three Khinchin axioms. It is superadditive for $\eta < 1$ and subadditive for $\eta > 1$. $S_\eta$ also satisfies the requirements of positivity, equiprobability, concavity and irreversibility. Jaynes thermodynamic relations are also verified.

It is worth remarking that these are not trivial properties that can be verified by any conceivable functional of a probability distribution. In fact, for instance, the well known Renyi entropies do not exhibit definite concavity for arbitrary values of their parameter. Although those properties were shown here for the discrete case, they are expected to be valid also for continuous variables.

Concerning irreversibility, it is worthwhile making some comments on the limitations of our treatment. The requirement of microscopic detailed balance is equivalent to assume...
that the equilibrium state is given by the uniform probability distribution. However, there are many important situations where the equilibrium distribution is not constant [23]. The proper way to encompass these more general settings is to formulate an $H$-theorem in terms of a relative entropy measure [23]. For example, the stationary solutions of the Fokker–Planck equation are, in general, not uniform [24]. Consequently, the Boltzmann–Gibbs entropy does not verify an $H$-theorem. However, the Kullback relative entropy between two time-dependent solutions of the Fokker–Planck equation does exhibit a monotonous time behaviour characterizing irreversibility [24]. Similar results have recently been obtained within Tsallis $q$-nonextensive formalism [25]. Some interesting related work has also been reported in [26]. $H$-theorems for the general Liouville equation, involving just one of its solutions, have been obtained there. The main ingredient of those derivations was the definition of a entropic functional on the basis of an adroitly chosen phase space measure [25, 26]. However, that procedure is tantamount to using the relative entropy between the considered solution of the Liouville equation and the stationary solution [25]. From the above comments we can conclude that in order to extend our $H$-theorem to the case of general master equations it is necessary to first generalize the concept of relative entropy. That line of development, though worth pursuing, is beyond the scope of this work.

The properties verified by $S_\eta$ suggest that it might be a useful measure for describing nonextensive physical phenomena as well as for practical applications. Our derivation of the stretched exponential distributions from a ME principle may lead to new physical insights, as happened in the case of Jaynes ME approach to the ensembles of standard statistical mechanics [5]. Once the proper ME variational scheme leading to a particular family of distributions is identified, the scope of possible applications may be increased in a considerable way. For instance, the ME prescription can be implemented to obtain useful ansatz for the approximate description of time dependent processes [8]. A case story illustrating the importance of an appropriate ME principle yielding a particular kind of probability distributions is provided by Tsallis nonextensive $q$-formalism [1]. Tsallis (power-law) distributions have been known, within a variegated set of (apparently) unrelated scenarios, for a long time. For instance, polytropic distributions in stellar dynamics have been known since the beginning of this century [10]. However, the underlying connections between some of the power distributions appearing in various physical situations only began to be understood after the discovery of Tsallis ME principle [4, 15]. In the case of the new nonextensive entropic measure that we are introducing here, it would be important to clarify the physical meaning of the parameter $\eta$, and to understand why nature chooses, in some situations, to maximize the functional $S_\eta$. The only way to attain such an understanding is by a detailed study of the dynamics of each particular system described by stretched exponential distributions. We hope that our present contribution may stimulate further work within this line of inquiry.

Acknowledgments

We are grateful to C Tsallis for useful discussions and for encouraging this work. We also want to thank R O Vallejos and G R Guerberoff for insightful remarks. We acknowledge the Brazilian agencies FAPERJ, PRONEX and CNPq for partial financial support.

References

ME approach to stretched exponential probability distributions

Abe S 1997 Phys. Lett. A 224 326
Mendes R S 1997 Physica A 242 299
An updated complete bibliography can be found in webpage http://tsallis.cat.cbpf.br/biblio.htm.
Katz A 1967 Statistical Mechanics (San Francisco, CA: Freeman)
Plastino A R and Plastino A 1995 Physica A 222 347
Plastino A R and Plastino A 1994 Physica A 202 438
Andrade R 1994 Physica A 203 486
Penna T J P 1995 Phys. Rev. E 51 R1
Laherriere J and Sornette D 1998 Preprint cond-mat/9801293
[12] Frisch U and Sornette D 1997 J. Physique I 7 1155
Khinchin A I 1957 Mathematical Foundations of Information Theory (New York: Dover)
Risken H 1989 The Fokker–Planck Equation (Berlin: Springer)