Physica A 441 (2016) 23–31

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Extensivity of Rényi entropy for the Laplace–de Finetti distribution

H. Bergeron [a], E.M.F. Curado [b,c,*], J.P. Gazeau [b,d], Ligia M.C.S. Rodrigues [b]

[a] Univ Paris-Sud, ISMO, UMR 8214, 91405 Orsay, France
[b] Centro Brasileiro de Pesquisas Fisicas, Brazil
[c] Instituto Nacional de Ciência e Tecnologia-Sistemas Complexos, Rua Xavier Sigaud 150, 22290-180 Rio de Janeiro, RJ, Brazil
[d] APC, UMR 7164, Univ Paris Diderot, Sorbonne Paris Cité, 75205 Paris, France

HIGHLIGHTS

- de Finetti distributions have extensive Rényi entropy.
- Lower and upper bounds for Boltzmann–Gibbs entropy for de Finetti distributions.
- Lower and upper bounds for Rényi entropy for de Finetti distributions.

ARTICLE INFO

Article history:
Received 24 March 2015
Received in revised form 30 July 2015
Available online 1 September 2015

Keywords:
Entropy
Rényi
Extensivity
Binomial distribution
Laplace–de Finetti representation

ABSTRACT

The Boltzmann–Gibbs entropy is known to be asymptotically extensive for the Laplace–de Finetti distribution. We prove here that the same result holds in the case of the Rényi entropy. We also show some interesting lower and upper bounds for the asymptotic limit of these entropies.

1. Introduction

Since the seminal works by Clausius and Lord Kelvin introducing the concept of entropy, S, which ended with the analytical formulation given by Clausius in 1865 in Eq. (59) of his fundamental work [1], this quantity is written as $dS = \delta Q/T$, where δQ is the heat exchanged in a thermodynamical transformation and T is the Kelvin temperature. The fact that this thermodynamical entropy is an extensive quantity can be immediately seen from the fundamental relation of thermodynamics, that can be written for a simple system as $U = TS - pV + \mu N$. As the energy U is extensive, the entropy, S, has to be extensive, since the temperature, T, is an intensive variable. From the statistical mechanics point of view we have to work with a microscopic definition of entropy. This quantity is defined on the density probability of the microscopic states of the system, whether classical or quantum. This route was initiated by L. Boltzmann, with his work of 1872 [2] and put in clearer form in his work of 1877 [3]. In both papers he essentially uses the microscopic quantity

* Corresponding author at: Centro Brasileiro de Pesquisas Fisicas, Brazil.
E-mail addresses: herve.bergeron@u-psud.fr (H. Bergeron), evaldo@cbpf.br (E.M.F. Curado), gazeau@apc.univ-paris7.fr (J.P. Gazeau), ligia@cbpf.br (L.M.C.S. Rodrigues).

http://dx.doi.org/10.1016/j.physa.2015.08.014
0378-4371/

$H = \int dx f(x,t) \log f(x,t)$, where $f(x,t)$ is a density probability on the microscopic states and x is the energy. The density $f(x,t)$ tends to a time-invariant distribution when time tends to infinite and the microscopic-based quantity H tends to a steady-state value, which can be associated with the thermodynamical entropy. This steady-state quantity H is extensive and associated with the equilibrium state. Since then, the nowadays so-called Boltzmann–Gibbs entropy (S_{BG}), related to H, or its quantum version, the von Neumann entropy, was accepted by the scientific community as the microscopic version of the Clausius thermodynamical entropy.

In 1948 C. Shannon formulated his "A Mathematical Theory of Communication" [4] where, based on a set of three axioms, he obtained a quantity that measures the information, $H = -K \sum_{i=1}^{n} p_i \log p_i$, depending on the probabilities $\{p_i\}$ of the messages. He also called this quantity, H, "entropy", in spite of the fact that this quantity was measuring a different object than the Boltzmann–Gibbs entropy was.

In 1957 E.T. Jaynes [5] established the connection between information theory and statistical mechanics, assuming that, in statistical mechanics, the $\{p_i\}$represents the probability of the microscopic states of a given physical system. If these probabilities of microscopic states are the equilibrium probabilities, the quantity H, which is extensive, gives the Clausius thermodynamical entropy. In information theory, in fact, after the work of Shannon, many quantities other than H that measure the information were proposed, but these other information measures were always considered as appropriate to the field of information theory and were not related to the thermodynamical entropy, reserved to the Boltzmann–Gibbs – or von Neumann – entropy.

The association of the Boltzmann–Gibbs entropy with the thermodynamical entropy has been questioned in the last twenty five years for systems with long-range interactions or long-time memory, presenting a huge contraction of the accessible phase space of the system [6]. In those cases there are examples suggesting that the entropic form which should be extensive is not anymore the Boltzmann–Gibbs entropy but, for example, the Tsallis entropy. These results suggest that, for these cases, the microscopic entropic form which should be related to the thermodynamical entropy is the Tsallis one and not BG, the extensivity of the entropic form when the size of the system goes to infinity being one of the main criteria to choose the "correct" microscopic entropic form.

However, in spite of the fact that some of these results concerning systems with long-range interactions are disputed, it is largely believed among the scientists that the only entropic form, depending on the probabilities of the microscopic states, which is extensive for any independent or weakly correlated systems is the Boltzmann–Gibbs–Shannon entropy. In fact, extensivity is to be expected from BG entropy's well known additivity property i.e., $S_{BG}(E_1 \cup E_2) = S_{BG}(E_1) + S_{BG}(E_2)$ where E_1 and E_2 are independent. Therefore, as it is well known that Rényi entropy is additive [7], its extensivity for any independent systems should be also expected. Simple physical examples are chain of independent spins and ideal gases.

In this work we revisit these basic concepts by comprehensively examining the extensivity properties of Boltzmann–Gibbs (BG) and Rényi entropies for the binomial distribution, which is appropriate for uncorrelated events, and for the Laplace–de Finetti representation [8], which deals with binary correlated systems. For an interesting discussion about the meaning of this representation, see Jaynes [9]. In recent years, application of the de Finetti representation theorem, in its more abstract versions, has been gaining more and more importance both in classical and quantum physics [10,11]; for more references, see Ref. [12].

For the binomial distribution we prove that both entropies are extensive, i.e. are proportional to the number n of events; we also prove that both entropies are upper-bounded by $n \log 2$, which is the value that both assume when win and loss are equiprobable. The same upper bound $n \log 2$ holds in the Laplace–de Finetti case, and we prove that this is the value asymptotically assumed at large n, which means that both BG and Rényi entropies are asymptotically extensive. We thus obtain new results about the extensivity of Rényi entropy and show unknown inequalities concerning both entropies, BG and Rényi.

In Section 2 we examine and prove the extensivity properties of both the Boltzmann–Gibbs and Rényi entropies for the binomial distribution. The same study is implemented in Section 3 for the Laplace–de Finetti distribution. Our results are discussed in 4 and Appendix is devoted to the study of asymptotic behavior of both distributions.

In writing the paper, we have opted for a pedagogical presentation of our results. Although the binomial case is one of the most familiar distributions, we recall elementary facts and proofs in order to make the paper self-contained.

2. Entropies for the binomial case

Let us consider a sequence of n binary events, "win" or "loss", $x_1, x_2, \ldots, x_n$. The n-point joint probability of this sequence, $p_n(x_1, x_2, \ldots, x_n)$, is called exchangeable if p_n remains symmetric in its arguments, for all n. Also, it is always possible to get p_n by summing over x_{n+1} through $p_n(x_1, x_2, \ldots, x_n) = \sum_{x_{n+1}} p_{n+1}(x_1, x_2, \ldots, x_n, x_{n+1})$. Therefore, the probability of a given sequence of n trials does not depend on the specific order of the binary events but only on the number of sequence elements that are in one of the states, "win' or "loss". If we have k "wins' and $n-k$ "losses", we have

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{1}$$

possible sequences and the probability of getting them is

$$\mathfrak{P}_k^{(n)} = \binom{n}{k} \varpi_k^{(n)}, \tag{2}$$