Available online at www.sciencedirect.com
SCIENCE DIRECT®

Physics Letters A 338 (2005) 217–224

PHYSICS LETTERS A

www.elsevier.com/locate/pla

A step beyond Tsallis and Rényi entropies

Marco Masi

Dipartimento di Fisica G. Galilei, Padova, Italy

Received 1 October 2004; received in revised form 11 January 2005; accepted 25 January 2005

Available online 10 March 2005

Communicated by A.P. Fordy

Abstract

Tsallis and Rényi entropy measures are two possible different generalizations of the Boltzmann–Gibbs entropy (or Shannon's information) but are not generalizations of each others. It is however the Sharma–Mittal measure, which was already defined in 1975 [J. Math. Sci. 10 (1975) 28] and which received attention only recently as an application in statistical mechanics [Physica A 285 (2000) 351, Eur. Phys. J. B 30 (2002) 543] that provides one possible unification. We will show how this generalization that unifies Rényi and Tsallis entropy in a coherent picture naturally comes into being if the q-formalism of generalized logarithm and exponential functions is used, how together with Sharma–Mittal's measure another possible extension emerges which however does not obey a pseudo-additive law and lacks of other properties relevant for a generalized thermostatistics, and how the relation between all these information measures is best understood when described in terms of a particular logarithmic Kolmogorov–Nagumo average.

PACS: 05.70; 65.50; 89.70; 05.70.L

Keywords: Generalized information entropy measures; Tsallis; Rényi; Sharma–Mittal

1. Introduction

To gain a unified understanding of the different entropy measures and how they relate to each others in the frame of a generalized picture, it is first necessary to recall what characterizes "classical" entropies and emphasize some aspects which are important for the present Letter.

1.1. *The Boltzmann–Gibbs entropy and Shannon's information measure*

As it is well known, given a probability distribution $P = \{p_i\}$ $(i = 1, \dots, N)$, with p_i representing the probability of the system to be in the ith microstate, the Boltzmann–Gibbs (BG) entropy reads

$$S_{\mathrm{BG}}(P) = -k \sum_{i=1}^{N} p_i \log p_i ,$$

E-mail addresses: marco.masi@spiro.fisica.unipd.it, marco_ masi2@tin.it (M. Masi).

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2005.01.094

where k is the Boltzmann constant and N the total number of microstates. If all states are equiprobable it leads to the famous Boltzmann principle $S = k \log W$ ($N = W$). BG entropy is equivalent to Shannon's expression if we set $k = 1$ (as we will do from now on) and use the immaterial base b for the logarithm function

$$S_{\mathrm{S}}(P) = -\sum_{i=1}^{N} p_i \log_b p_i.$$

It is common to use the natural base for the BG entropy, while base 2 has the advantage to deliver information quantities in bits.

What characterizes BG and Shannon's measure is additivity of information. Given two systems, described by two independent probability distributions A and B (i.e., $P(A \cap B) = P(A)P(B)$), using an additive information measure means that

$$S_{\mathrm{S}}(A \cap B) = S_{\mathrm{S}}(A) + S_{\mathrm{S}}(B|A),$$

with

$$S_{\mathrm{S}}(B|A) = \sum_i p_i(A) S_{\mathrm{S}}(B|A = A_i),$$

being the conditional entropy. In this case we are talking about *extensive systems*, i.e., systems where the entropy is given by the sum of all the entropies of their parts, as it is customary to do in standard statistical mechanics. The unique function which assures additivity is the logarithm. Also in the axiomatic derivation of Shannon's entropy performed by Khinchin [1], it is the additive property which leads to the appearance of the logarithm function. This is the real reason that stands behind the ubiquitous presence of the logarithm function in information theory, and we can confidently say that every modification to it reflects a deviation from the additive law.

We will from now on use the natural base. Shannon's entropy can be written in the form of a "linear" (the arithmetic) mean as

$$S_{\mathrm{S}}(P) = \langle I_i \rangle_{\mathrm{lin}} = \left\langle \log\left(\frac{1}{p_i}\right) \right\rangle_{\mathrm{lin}}, \tag{1.1}$$

where we will call the quantity

$$I_i = \log\left(\frac{1}{p_i}\right),$$

the *elementary information gain* associated to an event of probability p_i (in information theory it is sometimes called the *code length*). The quantity $\frac{1}{p_i}$ is also called the *surprise* (less probable events are considered more "surprising" than more probable ones), and we will see that it is this quantity which is really measured in one way or another, not $-\log p_i$.

1.2. Tsallis' entropy

Additivity is however not always preserved, especially in nonlinear complex systems, e.g., when we have to deal with long range forces, as it is in the case of the dynamic evolution of star clusters or in systems with long range microscopic memory, in fractal- or multifractal-like and self-organized critical systems, etc. We are dealing in this case with *non-extensive systems*; a case which received much attention in the last decade [2].

A generalization of the BG entropy to *non-extensive systems* is known as Tsallis entropy [3]. C. Tsallis noted that if non-extensivity enters into the play things are described better by power law distributions, p_i^q, so called *q-probabilities*, i.e., by scaled probabilities where q is a real parameter. This introduces the formal possibility not to set rare and common events on the same footing, as in BG or Shannon statistics, but it enhances or depresses them according to the parameter chosen (in complex systems rare events can have dramatic effects on the overall evolution).

With the introduction of the normalized *q-probabilities* it became customary to define so-called *escort-* or *zooming-distribution*

$$\pi_i(P, q) = \frac{p_i^q}{\sum_{i=1}^N p_i^q}, \quad q > 0,\ q \in \mathbb{R}.$$

In this frame Tsallis postulated his now famous generalization of Shannon's entropy to non-extensivity [3]:

$$S_{\mathrm{T}}(P, q) = \frac{\sum_{i=1}^N p_i^q - 1}{1 - q} = \frac{1}{q-1} \sum_{i=1}^{N} p_i \left(1 - p_i^{q-1}\right). \tag{1.2}$$

For $q \to 1$, Shannon's measure is recovered, i.e., $S_{\mathrm{T}}(P, 1) = S_{\mathrm{S}}(P)$.