Physica A 422 (2015) 167–174

Contents lists available at ScienceDirect

Physica A

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

Generalized relative entropies in the classical limit

A.M. Kowalski [a,b], M.T. Martin [a,c], A. Plastino [a,c,*]

[a] Instituto de Física (IFLP-CCT-Conicet), Fac. de Ciencias Exactas, Universidad Nacional de La Plata, C.C. 727, 1900 La Plata, Argentina
[b] Comision de Investigaciones Científicas (CIC), Argentina
[c] Argentina's National Research Council (CONICET), Argentina

HIGHLIGHTS

- Statistical quantifiers are compared in their ability to describe feature of the route towards the classical limit.
- The normalized Cressie–Read and relative Tsallis ones are shown to be equivalent.
- The Tsallis quantifier is seen to provide a better description than the Kullback–Leibler one.

ARTICLE INFO

Article history:
Received 1 October 2014
Received in revised form 11 December 2014
Available online 23 December 2014

Keywords:
Tsallis relative entropy
Cressie–Read quantifiers
Classical limit
Time-series

ABSTRACT

Our protagonists are (i) the Cressie–Read family of divergences (characterized by the parameter γ), (ii) Tsallis' generalized relative entropies (characterized by the q one), and, as a particular instance of both, (iii) the Kullback–Leibler (KL) relative entropy. In their normalized versions, we ascertain the equivalence between (i) and (ii). Additionally, we employ these three entropic quantifiers in order to provide a statistical investigation of the classical limit of a semiclassical model, whose properties are well known from a purely dynamic viewpoint. This places us in a good position to assess the appropriateness of our statistical quantifiers for describing involved systems. We compare the behaviour of (i), (ii), and (iii) as one proceeds towards the classical limit. We determine optimal ranges for γ and/or q. It is shown the Tsallis-quantifier is better than KL's for $1.5 < q < 2.5$.

1. Introduction

Entropic quantifiers (see as examples Refs. [1–4], and references therein) are useful in the study of time series' underlying dynamics. Systems that are characterized by either long-range interactions, long-term memories, or multi-fractality are best described by a generalized statistical mechanics' formalism [5] usually alluded to as deformed, Tsallis' q-statistics. The basic associated entity is the entropy [$q \in \mathcal{R}$ ($q \neq 1$)]

$$S_q = \frac{1}{(q-1)} \sum_{i=1}^{n} p_i \left[1 - {p_i}^{q-1}\right], \tag{1}$$

p_i being probabilities associated with the n different system-configurations. The entropic index (or deformation parameter) q describes the deviations of Tsallis entropy from the standard Boltzmann–Gibbs–Shannon-one. One has

$$S = -\sum_{i=1}^{n} p_i \ln p_i. \tag{2}$$

* Corresponding author at: Instituto de Física (IFLP-CCT-Conicet), Fac. de Ciencias Exactas, Universidad Nacional de La Plata, C.C. 727, 1900 La Plata, Argentina. Tel.: +54 11 4786 8114; fax: +54 11 4786 8114.
E-mail addresses: kowalski@fisica.unlp.edu.ar (A.M. Kowalski), mtmartin@fisica.unlp.edu.ar (M.T. Martin), plastino@fisica.unlp.edu.ar (A. Plastino).

http://dx.doi.org/10.1016/j.physa.2014.12.017
0378-4371/

Shannon's entropy works best for systems composed of either independent subsystems or interacting via short-range forces, whose subsystems can access all the available phase space [6]. For systems exhibiting long-range correlations, memory, or fractal properties, Tsallis' entropy becomes the most convenient quantifier [6–16].

In another vein, we have the Cressie–Read (CR) family of power divergences, defined through a class of additive convex functions. The CR power divergence measure encompasses a broad family of test statistics that leads to a large family of likelihood functions. They constitute a family of pseudo-distance measures from which to derive empirical probabilities associated with indirect noisy micro and macro data [17].

In order to assess how good these quantifiers are to statistically describe complex scenarios, we will apply the above mentioned quantifiers to a celebrated semiclassical system in its route towards the classical limit [18,19]. The system's dynamics exhibits regular zones, chaotic ones and other regions that, although not chaotic, display complex features. The system has been exhaustively investigated from a purely dynamic viewpoint [19] and also from a statistical one [20–22]. This last kind of study has a pre-requisite: how to extracted information from a time series (TS) [23]. The data at our disposal always possess a stochastic component due to noise [24,25], so that different extraction-procedures attain distinct degrees of quality. We will employ the Bandt and Pompe's approach [26], that determines the probability distribution associated to time series on the basis of the nature of the underlying attractor (see Appendix for the mathematical details).

Summing up, we will use the normalized versions of Tsallis relative entropy [9,27] and Cressie–Read family of divergences [17], to which we add Kullback–Leibler's relative entropy. It will be seen that the normalized CR coincides with the normalized Tsallis relative entropy for a special relationship between q and γ. With these entropies we will compare (i) the probability distribution functions (PDFs) associated to the system's dynamic equation's solutions in their route towards the classical limit [19] with (ii) the PDF associated to the classical solutions.

The relative entropies mentioned above are discussed in Section 2, which briefly recapitulates notions concerning the Tsallis relative entropy, the Kullback–Leibler relative entropy and the CR-divergence family of entropic functionals. As a test-scenario, the semiclassical system and its classical limit are described in Section 3, and the concomitant results are presented in Section 4. Finally, some conclusions are drawn in Section 5.

2. Kullback–Leibler relative entropy, Tsallis relative entropy and Cressie–Read family of divergences

The relative entropies (RE) quantify the difference between two probability distributions P and Q [28]. They provide an estimation of how much information P contains relative to Q and measure the expected number of extra bits required to code samples from P when using a code based on Q, rather than using a code based on P [28]. They can also be regarded as entropic *distances*, alternative means for comparing the distribution Q to P. The best representative is the Kullback–Leibler's (KL) one, based on the Shannon canonical measure (2). For two normalized, discrete probability distribution functions (PDF) $P = (p_1, \ldots, p_n)$ and $Q = (q_1, \ldots, q_n)$ $(n > 1)$, one has

$$D_{\mathrm{KL}}(P, Q) = \sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{q_i}\right), \tag{3}$$

with $D_{\mathrm{KLsn}}(P, Q) \geq 0$. $D_{\mathrm{KL}}(P, Q) = 0$ if and only if $P = Q$. One assumes that either $q_i \neq 0$ for all values of i, or that if one $q_i = 0$, then $p_i = 0$ as well [29]. In such an instance people take $0/0 = 1$ [29] (also, $0 \ln 0 = 0$, of course). It is convenient to work with a normalized KL-version, for the sake of a better comparison between different results. In this way the quantifier's values are restricted to the [0, 1] interval, via division by its maximum allowable value. If we divide D by $\ln n$, expression (3) becomes

$$D^{N}_{\mathrm{KL}}(P, Q) = \frac{1}{\ln n} \sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{q_i}\right), \tag{4}$$

with $0 \leq D^{N}_{\mathrm{KL}} \leq 1$. We will work with Eq. (4). KL can be seen as a particular case of the generalized Tsallis relative entropy [9,27]

$$D_q(P, Q) = \frac{1}{q-1} \sum_{i=1}^{n} p_i \left[\left(\frac{p_i}{q_i}\right)^{q-1} - 1\right], \tag{5}$$

when $q \to 1$ [9,27]. $D_q(P, Q) \geq 0$ if $q \geq 0$. For $q > 0$ one has $D_q(P, Q) = 0$ if and only if $P = Q$. For $q = 0$ one has $D_q(P, Q) = 0$ for all P and Q.

We pass now to define the Cressie–Read (CR) family of divergence measures [17]:

$$I(P, Q, \gamma) = \frac{1}{\gamma(\gamma+1)} \sum_{i=1}^{n} p_i \left[\left(\frac{p_i}{q_i}\right)^{\gamma} - 1\right], \tag{6}$$

where γ is a parameter that indexes members of the CR family. CR differs from $D_q(P, Q)$ because of the condition $I(P, Q, \gamma) \geq 0$, **for all** γ. In the two special cases where $\gamma = 0$ or -1, the notation $I(P, Q, 0)$ and $I(P, Q, -1)$ are to be interpreted as the limits, $\lim_{\gamma \to 0}$ or $\lim_{\gamma \to -1}$, respectively [17]. The $\gamma = 0$ case, corresponds to $D_{\mathrm{KL}}(p, q)$ [17], mimicking what happens with $D_q(P, Q)$ when $q \to 1$. On the other hand, $I(P, Q, -1) = D_{\mathrm{KL}}(Q, P)$ [17].