Physica A 507 (2018) 153–174

Contents lists available at ScienceDirect

Physica A

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

Comparing permutation entropy functions to detect structural changes in time series

J.S. Cánovas *, G. García-Clemente, M. Muñoz-Guillermo

Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, Antiguo Hospital de Marina, 30202 Cartagena, Murcia, Spain

HIGHLIGHTS

- A comparison between several entropy functions is done to find structural changes in time series.
- Among several entropy functions, Tsallis ones are the entropy functions that better show the structural changes.
- We apply our theoretical results to the analysis of an earthquake and some financial time series.

ARTICLE INFO

Article history:
Received 16 October 2017
Received in revised form 23 April 2018

ABSTRACT

Entropy can be taken as a measure of the complex dynamical behavior. In this paper, we consider different entropy functions and the permutation symbolic dynamics and we apply them to find structural changes in time series. We analyze what entropy functions are more suitable to show changes in simulated time series where the structural changes are know. Applications to seismic real data and economic data series are shown to illustrate how this type of tools can be used.

1. Introduction

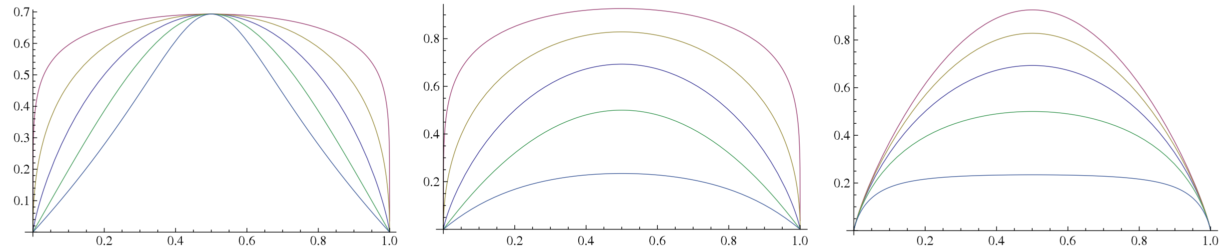
Different measures of complexity have been used to compare time series and distinguish between regular and chaotic behavior. The complexity of heart and brain data can distinguish healthy and sick subjects and sometimes even predict heart attack or epileptic seizure, [1,2]. Other natural processes, like earthquakes, are complex systems by nature. The number of variables to take into account is big enough to make difficult the attempts of modeling them as well as the possibility of predicting their behavior.

Entropy-based indicators are useful to study systems evolving with time with a degree of disorder or chaoticity. Time series appear as a source of information in which other approaches are not possible. Measuring the degree of disorder of such that systems is a non-trivial problem. In this frame, permutation entropy can be used as a measure of the chaoticity of the system, see [1,3–5]. In [6] permutations are used to encode the information of a data series and detect structural changes using the number of admissible permutations. On the other hand, permutations jointly with a limit of Renyi entropy functions as the parameter goes to infinity has been used in [7] to detect dynamical changes of time series. Finally, Shannon permutation entropy has been used to detect changes in time series (see [8,9]).

In this paper, following [6], permutations are used to codify the information contained in a data series in order to find structural changes but different entropy-measures (see [10]) are used to detect such structural changes. The main motivation

* Corresponding author.
E-mail addresses: Jose.Canovas@upct.es (J.S. Cánovas), gonzalo10993@gmail.com (G. García-Clemente), Maria.mg@upct.es (M. Muñoz-Guillermo).

https://doi.org/10.1016/j.physa.2018.04.101
0378-4371/



(a) Renyi entropies for two permutations with frequencies x and $1 - x$ for parameter values $r = 1/5, 1/2, 1$ (Shannon entropy), 2 and 5.

(b) Tsallis entropies for two permutations with frequencies x and $1 - x$ for parameter values $q = 1/5, 1/2, 1$ (Shannon entropy), 2 and 5.

(c) Gaussian entropies for two permutations with frequencies x and $1 - x$ for parameter values $q = 1/5, 1/2, 1$ (Shannon entropy), 2 and 5.

Fig. 1. Entropy functions with two frequencies x and $1 - x$. Shannon entropy divides the parameter region such that $q, r > 1$ and $q, r < 1$.

for this paper is to analyze what entropy measure is better to show structural changes in data series. Of course, this problem is quite difficult to analyze because families of entropy functions depend on real parameters. All of them include Shannon entropy as a limit case, which splits the parameter set in two disjoint subsets. Our study considers one parameter value in each of these disjoint subsets. After normalization, these chosen parameters give representative entropy functions with parameters on these subsets.

As usual, first we make experiments with simulated time series where structural changes are given and then, we apply the techniques to analyze real data time series analysis. In particular, we consider real seismic series, see [11–13], and we will apply the practical analysis of entropy-measures to that seismic real data. In addition, two different economic data have been analyzed as well.

The paper is organized as follows. Section 2 is devoted to introduce the basic notation and definitions. We highlight the way in which the time series is encoded and we describe the entropy-like measures that will be used to make the analysis. In Section 3 different time series are generated and we apply the study of the entropy, following the sketch of [6]. Several applications to real data series are given in Section 4. The situation of the village of Lorca (Spain) in the previous days to the earthquake in 11th of May of 2011 is studied. An earthquake of catastrophic consequences happened in that date. We make a thorough analysis, specially we are interested in facts that might indicate precursory behavior. Additionally, the existence of changes in economic data series is also addressed.

2. Permutations and entropy functions

In order to make this paper self-contained, basic notation and definitions will be introduced within this section. More details can be obtained in [6] and [10].

Definition 1. Let $(x_n)_{n=1}^T$, $T \in \mathbb{N}$ be a time series, where $x_n \in \mathbb{R}$ for each $n \in \{1, \dots, T\}$. An *sliding window* from $(x_n)_{n=1}^T$ is given by

$$x_k(l) = (x_l, x_{l+1}, \dots, x_{l+k-1}),$$

for $1 \leq l < T - k + 1$, where k is a fixed natural number called the embedding dimension.

Observe that the embedding dimension agrees with the number of terms of the sliding window. The adjective *sliding* is because the window changes its position. The window is *slid* until covering the entire series or a subwindow. This cover can be done with or without overlapping between the windows.

Let m be a fixed natural number and $\mathcal{S}_m$ be the group of permutations of length m. The cardinality of $\mathcal{S}_m$ is given by $|\mathcal{S}_m| = m!$, where $|A|$ denotes the cardinality of a set A.

Definition 2. Let $m \in \mathbb{N}$ be a fixed natural number and $\pi = (i_1, i_2, \dots, i_m) \in \mathcal{S}_m$ a permutation. The sliding window $x_m(l) = (x_l, x_{l+1}, \dots, x_{l+m-1})$ is said to be of π-type if π is the unique permutation such that the following conditions hold:

(1) $x_{l+i_1} \leq x_{l+i_2} \leq \cdots \leq x_{l+i_m}$.
(2) $i_{s-1} < i_s$ if $x_{l+i_{s-1}} = x_{l+i_s}$.

Definition 3 (*[1]*)**.** Let $(x_n)_{n=1}^T$ be a time series, $m \in \mathbb{N}$ and $\pi \in \mathcal{S}_m$, then the *relative frequency of* π, denoted by $p(\pi)$ is given by

$$p(\pi) = \frac{|\{j : x_m(j), \text{ is of } \pi - \text{type}, j = 1, 2, \dots, T - m + 1\}|}{T - m + 1}.$$