Physica A 486 (2017) 273–283

Contents lists available at ScienceDirect

Physica A

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

Permutation entropy analysis based on Gini–Simpson index for financial time series

Jun Jiang, Pengjian Shang *, Zuoquan Zhang, Xuemei Li

Department of Mathematics, Beijing Jiaotong University, No.3 of Shangyuan Residence Haidian District, Beijing 100044, PR China

HIGHLIGHTS

- We propose a new permutation entropy based on Gini–Simpson index (GPE).
- We verify the stability and precision of the GPE method by simulated time series (Logistic map).
- GPE analysis is applied to simulated signals, US and European and Chinese stock markets.
- We compare different stock markets by original dynamics and find the inner mechanism of those markets.

ARTICLE INFO

Article history:
Received 14 February 2017
Received in revised form 5 May 2017
Available online 1 June 2017

Keywords:
Permutation entropy
Gini–Simpson index
Logistic map
Financial time series

ABSTRACT

In this paper, a new coefficient is proposed with the objective of quantifying the level of complexity for financial time series. For researching complexity measures from the view of entropy, we propose a new permutation entropy based on Gini–Simpson index (GPE). Logistic map is applied to simulate time series to show the accuracy of the GPE method, and expound the extreme robustness of our GPE by the results of simulated time series. Meanwhile, we compare the effect of the different order of GPE. And then we employ it to US and European and Chinese stock markets in order to reveal the inner mechanism hidden in the original financial time series. After comparison of these results of stock indexes, it can be concluded that the relevance of different stock markets are obvious. To study the complexity features and properties of financial time series, it can provide valuable information for understanding the inner mechanism of financial markets.

1. Introduction

The research of economics has been in the spotlight while a large amount of studies have employed statistical mechanics to analysis the nature of economy. Econophysics [1] opens the door to investigate economic problems which is the term of interdisciplinary research. It is no doubt that information is the crux of market analysis or predicting the complexity of stock. Entropy is easily calculated for financial time series, which is a measure of degree of uncertainty to detect the system complexity and has widely applied in physical, medical, engineering, and also economic sciences. Nicholas Georgescu-Roegen who is a pioneer throughout his research of economic sphere for his path-breaking 1971 masterpiece *The Entropy Law and the Economic Process*, that is the first time when entropy appeared in the economy field. Christoph Bandt et al. introduced the interesting concept of permutation entropy (PE) [2], as a complexity measure for time series analysis, which is very similar to Lyapunov exponents for the well-know family of logistic maps [3].

* Corresponding author.
E-mail address: pjshang@bjtu.edu.cn (P. Shang).

http://dx.doi.org/10.1016/j.physa.2017.05.059
0378-4371/

As a relatively new mathematical statistics method, permutation entropy has become a hot topic among scholars and researches around the world. Nowadays, an increasingly number of scholars settle down to study the theoretical of permutation entropy what is widely used and works quite well in many fields, such as medicine [4], biology [5], image processing [6], economic [7], etc. Compared with Lyapunov exponent and fractal dimension, it holds the advantages of simple calculation and strong anti-noise ability. The essential feature of permutation entropy is to replace the data with a symbol sequence with finitely many symbols for the given time series. The higher the value of permutation entropy, the higher complexity of the data.

For PE, farther model assumptions are needless, videlicet, the symbol sequence must come naturally from the data. It is universally received that the time series show complexity behaviors in the real world. The good news is that by the generalized permutation entropy, we desire to investigate the complexity behavior between time series hidden in the permutation entropy. So, we propose a modified permutation entropy [8] based on Gini–Simpson index [9](GPE) which has the capacity of detecting the features of complexity for the financial time series. In this paper, the simulated signals generated by logistic maps, one of the typical mappings of complex nonlinear behavior, which are used to verify the result of GPE in order to show the good properties of our GPE.

Since PE was introduced, scholars of different fields study their academic field by using this method and have make some research results. N. Nicolaou et al. [10] investigated for the first time the use of PE as a feature for automated epileptic seizure detection. And compared with other algorithms, the results showed that the proposed algorithm can improve the accuracy of the classification. XL Sun et al. [11] , applying the PE complexity metric to abiotic stress response time series data in Arabidopsis thaliana, genes involved in stress response and signaling were found to be associated with the highest complexity not only under stress, but surprisingly, also under reference, non-stress conditions. And there are also many researchers propose modified permutation entropy, such as Y. Li et al. [12] proposed a new bearing vibration feature extraction method based on multiscale permutation entropy to characterize the complexity of the principal PF component in different scales. Here we also propose a new modified permutation entropy that is called GPE to study the complexity features and properties of financial time series [13–15].

The remainder of this paper is organized as follows. In the following section, we introduce the basic principle of PE and propose the GPE method and describe the detail process of GPE analysis. Section 3 presents the results with simulated modeling and Section 4 provides the results and analysis of applying GPE method to original financial time series, US and European and Chinese stock markets. Finally, we conclude the paper in Section 5.

2. Methodology

2.1. Permutation entropy (PE)

As we all know that a dynamical system can be properly represented and analyzed by using a symbolic sequence. Permutation entropy, that is a convenient way to map a continuous time series onto a symbolic sequence [16,17].

For a given scalar time series, we embed the scalar time series $\{x(i), i = 1, 2, \ldots, T\}$ to a n-dimensional space: $X_i = [x(i), x(i+\tau), \ldots, x(i+(n-1)\tau)]$, where n is called the embedding dimension and τ the delay time. Reordering the X_i to an increasing order:

$$x(i+(j_1-1)\tau) \le x(i+(j_2-1)\tau) \le \cdots \le x(i+(j_n-1)\tau).$$

When an equality occurs such as $x[i+(j_{i1}-1)\tau] = x[i+(j_{i2}-1)\tau]$, we order the quantities x according to the values of their corresponding j_i, namely if $j_1 < j_2$ we get it that $x[i+(j_1-1)\tau] \le x[i+(j_2-1)\tau]$. Therefore, we can obtain a set of symbol sequences by each row of it that the reconstructed matrix of any time series, where the symbol sequences just like $S(\pi) = (j_1, j_2, \ldots, j_n)$, where the $\pi = 1, 2, \ldots, K$ while K means the objectively quantity of $(j_1, j_2, \ldots, j_n)$ and $K \le n!$. In the sequel, any vector X_i is uniquely mapped onto $(1, 2, \ldots, n)$ or $(2, 1, 3, \ldots, n)$ or $\cdots$ or $(n, n-1, \ldots, 2, 1)$ in total $n!$ different situations. Let the probability of occurrence of the distinct symbols be P_π $(\pi = 1, 2, \ldots, K)$, namely, for each π we determine the relative frequency (# means number)

$$P_\pi = \frac{\#\{t \mid t \le T-n, (j_1, j_2, \ldots, j_n) \text{ has type } \pi\}}{T-n+1}. \quad (1)$$

The permutation entropy of order $n \ge 2$ is defined as the Shannon entropy [18] for the K distinct symbols

$$H(n) = -\sum P_\pi log P_\pi, \quad (2)$$

it is clear that $0 \le H(n) \le log n!$.

The magnitude of the $H(n)$ value represents the random degree of time series X_i. The smaller the value of $H(n)$ is, the more regular time series is, otherwise, the time series is close to random series. The changes of $H(n)$ reflect and magnify the minute details of the time series.

It is necessary to determine the range of order n, before using it. Obviously, if n is equal to 1 or 2, this method would be meaningless since there are only very few distinct states. To the contrary, if n is too large, as we have just talked about it that π runs over all $n!$ permutations, it will bring about a large amount of computation. So we let $n = 3, \ldots, 6$, for practical application.