Chaos, Solitons and Fractals 188 (2024) 115481

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

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

Asymptotic distribution of entropies and Fisher information measure of ordinal patterns with applications

Andrea Rey [a,b], Alejandro C. Frery [c,*], Juliana Gambini [a,b], Magdalena Lucini [d,e]

[a] Centro de Investigación y Desarrollo en Informática Aplicada, Universidad Nacional de Hurlingham, Tte. Manuel Origone 151, 1688, Villa Santos Tesei, Argentina
[b] Department of Mathematics, Universidad Tecnológica Nacional Facultad Regional Buenos Aires, Mozart 2300, 1407, Buenos Aires, Argentina
[c] School of Mathematics and Statistics, Victoria University of Wellington, 6140, Wellington, New Zealand
[d] Facultad de Ciencias Exactas, Naturales y Agrimensura, Universidad Nacional del Nordeste, Av. Libertad 5470, 3400, Corrientes, Argentina
[e] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Av. Libertad 5470, 3400, Corrientes, Argentina

ARTICLE INFO

Keywords:
Ordinal patterns
Shannon entropy
Rényi entropy
Tsallis entropy
Havrda–Charvát entropy
Fisher information measure
Asymptotic distribution

ABSTRACT

We present the asymptotic distribution of the Rényi and Tsallis/Havrda–Charvát entropies and the Fisher information measure of ordinal patterns embedding their serial correlation. We study the convergence behavior of the asymptotic variance for some types of dynamics and the permutation entropy to the limit distribution. These results lead to tests for comparing the underlying dynamics of two time series. We apply these tests to discriminate uniform white noise, logistic maps with Gaussian noise, fractional Brownian motion, and f^{-k} noise, with $k \in \{0.5, 1, 1.5, 2, 2.5\}$. We also applied these tests to cryptocurrency open prices data, with favorable results. We provide the R code that implements the functions.

1. Introduction

The idea of entropy was first developed by Clausius [1]. Then, Boltzmann [2] applied this concept in the field of thermodynamics and Planck [3] proposed the formula as it is known nowadays. The generalization of the Boltzmann entropy to thermodynamic systems with microstates having non-equal probabilities was given by Gibbs [4]. However, it was Shannon [5] who defined the notion of entropy in the field of Information Theory as a way to measure the uncertainty associated with a discrete random variable. Entropy can be thought of as the average amount of information conveyed in a language. The importance and applications of entropies have been widely studied in a vast number of works [6–15].

Bandt and Pompe [16] proposed a methodology based on the ordinal patterns of a time series, which has been successfully applied in signal analysis. This technique transforms each equal-sized overlapping window of the time series into an ordinal pattern, a symbol built from the sequence of indices obtained that sort the values in the window. Thus, the time series can be studied through the sequence of ordinal patterns by computing the histogram of proportions, followed by its entropy and statistical complexity.

This method has interdisciplinary applications as biomedical and economic signal analysis [17,18], body temperature classification [19], emotional states analysis of RRI time series from ECG [20], Alzheimer's disease research [21], characterization of the clinical electrophysiological evolution of patients [22], texture analysis of synthetic and biomedical images [23], and SAR image interpretation [24], among many others.

One of the greatest appeals of this approach is the ability to display the point with coordinates entropy and statistical complexity in a closed manifold. The point position reveals structural information about the underlying dynamics that produced the time series; see, for instance, Fig. 4 from Chagas et al. [25].

For the sake of notational simplicity, consider the real-valued time series of length $N = n + m - 1$; $m \geq 2$ is an integer called the embedding dimension. Denote this time series as $\boldsymbol{x} = (x_1, x_2, \ldots, x_{n+m-1})$. For $t = 1, 2, \ldots, n$, let $\boldsymbol{s}_t = (x_t, x_{t+1}, \ldots, x_{t+m-1})$ be an overlapping window of m consecutive values in $\boldsymbol{x}$. Assume these values are different. Then, map each subsequence $\boldsymbol{s}_t$ into a symbol π_t univocally determined by the indexes that sort $(x_t, x_{t+1}, \ldots, x_{t+m-1})$; this is the ordinal pattern of $\boldsymbol{s}_t$. This operation, called "Bandt–Pompe Symbolization" or "Ordinal Pattern Transformation", converts the time series $\boldsymbol{x}$ into the sequence of symbols $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$. Each symbol π_j can take one of $m!$ values: $\pi_j \in \Pi = \{\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(m!)}\}$. The probability that the pattern $\pi^{(i)}$ appears in the sequence $\boldsymbol{\pi}$ is denoted by q_i, for $i = 1, 2, \ldots, m!$.

The sequence of patterns is invariant to monotonically increasing functions of $\boldsymbol{x}$ and it is less sensitive to outliers than descriptors that use

* Corresponding author.
E-mail address: alejandro.frery@vuw.ac.nz (A.C. Frery).

https://doi.org/10.1016/j.chaos.2024.115481
Received 8 August 2024; Accepted 1 September 2024
Available online 20 September 2024
0960-0779/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

the original values [25]. These two features make the ordinal patterns efficient for signal analysis and interpretation [26].

In the case of a time series obtained from N independent trials with equally probable outcomes $b_1, b_2, \ldots, b_M$, as N tends to infinity, the Shannon entropy distribution converges to a χ^2 distribution with $M-1$ degrees of freedom [27]. This framework includes the case of white noise, i.e., $q_i = 1/m!$ for all $i = 1, 2, \ldots, m!$, when the entropy is maximum. The χ^2 distribution is also the asymptotic distribution for the plug-in estimators of the mutual information and the transfer entropy [28].

The marginal (as opposed to transitional) study of ordinal patterns is based on computing the histogram of proportions of symbols in $\boldsymbol{\pi}$. Denote $\widehat{\boldsymbol{q}} = (\widehat{q_1}, \ldots, \widehat{q_{m!}})$ the observed proportion of patterns:

$$\widehat{q_i} = \frac{\#\{\pi_j \in \boldsymbol{\pi} : \pi_j = \pi^{(i)}\}}{n}, \ 1 \le i \le m! \tag{1}$$

Ordinal patterns of series with structural differences have typically different proportions. Since the entropy and the Fisher information measure depend on these proportions, they will also differ.

The entropy is a measure of a system disorder and is a central concept in Information Theory. The Shannon entropy [5] has been widely used in multiple applications, but there are other forms of entropy as Tsallis [29] and Rényi [30]. Also, the Fisher information measure [31] can be computed from the histogram of patterns. This work provides the asymptotic distributions of these quantities, with which confidence intervals can be computed. The duality between confidence intervals and hypothesis tests can be used to determine whether two time series have similar dynamics.

The literature presents results about test statistics related to ordinal patterns and to entropy. Salicru et al. [32] proposed a test to study the diversity of populations using entropies. Esteban and Morales [33] presented tests for *(i)* a predicted value of the population entropy, *(ii)* for a common predicted value of several population entropies, *(iii)* the equality of several population entropies. On the other hand, Matilla-García [34] and Matilla-García and Marín [35] studied tests for serial independence. Matilla-García et al. [36] proposed a test for checking if two series are independent. These tests were revisited, and the results improved by Elsinger [37]. Weiß et al. [38] studied the distribution of symbols stemming from monotonic subsequences to identify white noise. Weiß and Schnurr [39] generalized the definition of ordinal patterns to test serial dependence in discrete-valued time series.

Chagas et al. [40], using an empirical approach, obtained confidence regions for the entropy and the statistical complexity of white noise. By making the simplification that the patterns are independent, Rey et al. [41] obtained the asymptotic distribution of the Tsallis and Rényi entropies and of the Fisher information measure when the proportions can be modeled by a Multinomial distribution. However, due to the overlapping incurred, the sequence of ordinal patterns has serial correlation [37] and, therefore, the hypothesis of a Multinomial distribution for the proportions can be improved. Rey et al. [42] obtained the asymptotic distribution of the Shannon permutation entropy considering correlated patterns.

In this article, we extend those works and obtain asymptotic distributions of three descriptors from correlated patterns: Tsallis and Rényi entropies, and Fisher information measure. Our main contributions can be summarized as follows:

- Provide the asymptotic distribution for a group of entropies obtained from the ordinal patterns.
- Obtain the asymptotic variance required for a hypothesis test that can be applied to distinguish time series with different dynamics.
- Furnish the users with an R package that computes the asymptotic variance and implements the proposed hypothesis test.

These results can be used in works that compare the underlying dynamics of time series, provided they have a stochastic component; e.g., Spichak and Aragoneses [43].

Table 1
Normalized entropies.

Entropy	Normalized version
Shannon	$H^{(S)}[\boldsymbol{p}] = \dfrac{S^{(S)}[\boldsymbol{p}]}{\log k}$
Tsallis	$H^{(T_\beta)}[\boldsymbol{p}] = \dfrac{\beta-1}{1-k^{1-\beta}} S^{(T_\beta)}[\boldsymbol{p}]$
Rényi	$H^{(R_\beta)}[\boldsymbol{p}] = \dfrac{S^{(R_\beta)}[\boldsymbol{p}]}{\log k}$
Fisher	$H^{(F)}[\boldsymbol{p}] = 4\sum_{i=1}^{k-1}(\sqrt{p_{i+1}} - \sqrt{p_i})^2$.

This paper unfolds as follows. Section 2 recalls the definitions of the descriptors studied in this work. Section 3 discusses the main properties of ordinal patterns that are needed to proceed to Section 4, where we obtain the asymptotic distribution of the Tsallis and Rényi entropies, and of the Fisher information measure. Using those distributions, in Section 5 we formulate test statistics for checking if features from two time series are indistinguishable. In Section 6, when applied to finite sample size, we assess the theoretical results by numerical validations, comparison with other techniques, and asymptotic behavior. Moreover, the proposed test is evaluated using simulated and real-world time series. Finally, Section 7 concludes the paper by discussing its limitations and outlining future research avenues.

2. Entropies

Let $\boldsymbol{p} = (p_1, p_2, \ldots, p_k)$ be a probability vector of size k, the entropies based on the $\boldsymbol{p}$ vector are defined as follows:

- Shannon entropy:

$$S^{(S)}[\boldsymbol{p}] = -\sum_{i=1}^{k} p_i \log p_i, \tag{2}$$

  where, by convention, $0 \log 0 = 0$.
- Tsallis entropy with index $\beta \in \mathbb{R} \setminus \{1\}$:

$$S^{(T_\beta)}[\boldsymbol{p}] = \sum_{i=1}^{k} \frac{p_i - p_i^\beta}{\beta - 1}. \tag{3}$$

  This form of entropy was originally proposed by Havrda and Charvát [44].
- Rényi entropy of order $\beta \in \mathbb{R}^+ \setminus \{1\}$:

$$S^{(R_\beta)}[\boldsymbol{p}] = \frac{1}{1-\beta} \log \sum_{i=1}^{k} p_i^\beta. \tag{4}$$

- Fisher information measure (for the discrete case):

$$H^{(F)}[\boldsymbol{p}] = 4\sum_{i=1}^{k-1}\left(\sqrt{p_{i+1}} - \sqrt{p_i}\right)^2. \tag{5}$$

  This form is obtained by approximating the derivative of the probability density function $f(x)$ with domain Ω involved in the Fisher information measure, which is given by $\int_\Omega f(x)\left(d[\ln f(x)]/dx\right)^2 dx$ [45].

All these descriptors, except the last one, achieve their maximum value when $\boldsymbol{p}$ is the equiprobable vector. If the entropy is divided by its maximum value, we obtain a normalized entropy that ranges from 0 to 1. The normalized versions of these descriptors are shown in Table 1. Notice that Eq. (5) is already normalized, but we added it to the table for completeness.

3. Ordinal patterns

In this section, we briefly described the methodology based on ordinal patterns proposed by Bandt and Pompe [16]. Let $\boldsymbol{x} = (x_1, x_2, \dots, x_{n+m-1})$ be a real-valued time series, where n and $m \geq 2$ are positive integers. For $t = 1, 2, \dots, n$, let $\boldsymbol{s}_t = (x_t, x_{t+1}, \dots, x_{t+m-1})$ be an overlapping window of m consecutive different values in $\boldsymbol{x}$. We consider $\Pi = \{\pi^{(1)}, \pi^{(2)}, \dots, \pi^{(m!)}\}$, the set of labeled permutations, or symbols, of the elements $0, 1, \dots, m-1$, which represent the possible ordinal patterns defined as follows. Given $t \in \{1, 2, \dots, n\}$, the subsequence $\boldsymbol{s}_t$ is said to be $\pi^{(i)}$-type, where $\pi^{(i)} = (i_1, i_2, \dots, i_m)$, if $x_{t+i_1} \leq x_{t+i_2} \leq \cdots \leq x_{t+i_m}$. In the presence of tied data, the order is defined by the sequential appearance of the repeated values, e.g., for $m = 3$ the subsequence $\boldsymbol{s}_t = (8, 5, 8)$ is of type $(1, 0, 2)$. Thus, from the time series $\boldsymbol{x}$, we obtain a sequence of symbols, its ordinal patterns, $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n)$, where $\pi_j \in \Pi$ for $j = 1, 2, \dots, n$.

The ordinal pattern probability vector $\boldsymbol{q} = (q_1, q_2, \dots, q_{m!})$ depends on the underlying dynamics of the time series. In this case, q_i denotes the relative frequency of π_i in the sequence of ordinal patterns, for $i = 1, 2, \dots, m!$. There are two limit cases: *(i)* $q_i = 1$ for a single i and $q_j = 0$ for all $j \neq i$, which is obtained by a monotonic time series, and *(ii)* $q_i = 1/m!$ for all $i = 1, 2, \dots, m!$, which represents a completely random system.

Ordinal patterns π_t and $\pi_{t+\ell}$ are correlated for $\ell = 1, 2, \dots, m-1$ and $t = 1, 2, \dots, n-\ell$ because π_t was computed from the subsequence $\boldsymbol{s}_t$ while $\pi_{t+\ell}$ was computed from $\boldsymbol{s}_{t+\ell}$, and these subsequences overlap. Such overlap captures information about the underlying unobserved time series dynamics. Then, studying the probability of transitions between consecutive patterns is interesting.

For $\ell = 1, 2, \dots, m-1$, the ordinal pattern transition matrix of order ℓ can be defined as $\mathbf{Q}^{(\ell)} \in \mathbb{R}^{m! \times m!}$, whose elements are

$$q_{ij}^{(\ell)} = \Pr\left(\pi_t = \pi^{(i)} \wedge \pi_{t+\ell} = \pi^{(j)}\right) = q_i \Pr(\pi_{t+\ell} = \pi^{(j)} \mid \pi_t = \pi^{(i)}), \tag{6}$$

for $i, j = 1, 2, \dots, m!$.

Let $\boldsymbol{N} = (N_1, N_2, \dots, N_{m!})$ be the vector whose ith component N_i counts the number of times the ordinal pattern π_i appears in $\boldsymbol{\pi}$, for $i = 1, 2, \dots, m!$. In other words, $N_i = \#\{\pi_t \in \boldsymbol{\pi} : \pi_t = \pi^{(i)}\}$. Notice that N_i/n is an estimator of q_i. Thus, we can define the following sequence of random vectors:

$$\widehat{\boldsymbol{q}}_n = \left(\frac{N_1}{n}, \frac{N_2}{n}, \dots, \frac{N_{m!}}{n}\right) = \left(\widehat{q}_{1,n}, \widehat{q}_{2,n}, \dots, \widehat{q}_{m!,n}\right). \tag{7}$$

In what follows, $\mathbf{D}_{\boldsymbol{q}} = \mathrm{Diag}(q_1, q_2, \dots, q_{m!})$ denotes the diagonal matrix obtained by $\boldsymbol{q}$. By [46, Eq. (30)], it holds that

$$\sqrt{n}(\widehat{\boldsymbol{q}}_n - \boldsymbol{q}) \xrightarrow[n\to\infty]{D} \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}), \tag{8}$$

with

$$\boldsymbol{\Sigma} = \mathbf{D}_{\boldsymbol{q}} - (2m-1)\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}} + \sum_{\ell=1}^{m-1}\left(\mathbf{Q}^{(\ell)} + \mathbf{Q}^{(\ell)^{\mathrm{T}}}\right), \tag{9}$$

where the superscript T indicates the transpose operator. Direct computations show that, for $i, j = 1, \dots, m!$

$$\Sigma_{ij} = \begin{cases} q_i - (2m-1)q_i^2 + 2\sum_{\ell=1}^{m-1} q_{ii}^{(\ell)} & \text{if } i = j, \\ -(2m-1)q_i q_j + \sum_{\ell=1}^{m-1}\left(q_{ij}^{(\ell)} + q_{ji}^{(\ell)}\right) & \text{if } i \neq j. \end{cases} \tag{10}$$

The covariance matrix depends on the transition probabilities defined in (6). The theoretical matrices $\boldsymbol{Q}^{(\ell)}$ were calculated for a random walk [46], for white noise [46], and for the process of 3 drawn of a random variable with equal probable possible values ± 1 [37], using $m = 3$. Given a time series, these matrices can be estimated by the relative frequencies of the observed pattern transitions.

This context is suitable to apply the multivariate version of the Delta Method ([47, Theorem 8.22]). Let $h_1, h_2, \dots, h_{m!}$ be continuously differentiable real functions defined in a neighborhood of the parameter point $\boldsymbol{q}$. Consider $\boldsymbol{J}$ the matrix of partial derivatives; i.e. $J_{ij} = \partial h_i/\partial q_j$ for $i, j = 1, 2, \dots, m!$. Assuming that $\boldsymbol{J}$ is non-singular in this neighborhood, the Delta Method theorem states that

$$\sqrt{n}\left[h_1(\widehat{\boldsymbol{q}}_n) - h_1(\boldsymbol{q}), h_2(\widehat{\boldsymbol{q}}_n) - h_2(\boldsymbol{q}), \dots, h_{m!}(\widehat{\boldsymbol{q}}_n) - h_{m!}(\boldsymbol{q})\right] \xrightarrow[n\to\infty]{D} \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{J}\boldsymbol{\Sigma}\boldsymbol{J}^{\mathrm{T}}\right). \tag{11}$$

Let $\alpha_1, \alpha_2, \dots, \alpha_{m!}$ be real numbers. The following convergence for the linear combination of the components of Eq. (11) [48, Theorem 5.2] holds:

$$\sqrt{n}\left[\sum_{i=1}^{m!}\alpha_i h_i(\widehat{\boldsymbol{q}}_n) - \sum_{i=1}^{m!}\alpha_i h_i(\boldsymbol{q})\right] \xrightarrow[n\to\infty]{D} \mathcal{N}\left(0, \sum_{i=1}^{m!}\alpha_i^2\left(\boldsymbol{J}\boldsymbol{\Sigma}\boldsymbol{J}^{\mathrm{T}}\right)_{ii} + 2\sum_{i=1}^{m!-1}\sum_{j=i+1}^{m!}\alpha_i\alpha_j\left(\boldsymbol{J}\boldsymbol{\Sigma}\boldsymbol{J}^{\mathrm{T}}\right)_{ij}\right). \tag{12}$$

4. Asymptotic results for ordinal pattern entropies

In this section, we will extend the results obtained by Rey et al. [42] for the Permutation Entropy, which coincides with the Shannon entropy using the ordinal pattern probability vector $\boldsymbol{q}$. In that work, we detailed the computations for $m = 3$. A high computational effort is required for larger values of m.

In order to apply the Delta Method for the entropy, we need to define the functions h_i for $i = 1, 2, \dots, m!$ in an appropriate way. Let $\mathcal{I}$ be the interval $(0, 1)$. Then, we consider the following real functions with domain $(\mathcal{I})^{m!}$:

$$h_i^{(S)}(q_1, q_2, \dots, q_{m!}) = q_i \log q_i, \ i \in \{1, 2, \dots, m!\}, \tag{13}$$

$$h_i^{(T_\beta)}(q_1, q_2, \dots, q_{m!}) = q_i - q_i^\beta, \ i \in \{1, 2, \dots, m!\}, \tag{14}$$

$$h_i^{(R_\beta)}(q_1, q_2, \dots, q_{m!}) = q_i^\beta, \ i \in \{1, 2, \dots, m!\}, \tag{15}$$

$$h_i^{(F)}(q_1, q_2, \dots, q_{m!}) = \left(\sqrt{q_{i+1}} - \sqrt{q_i}\right)^2, \ i \in \{1, 2, \dots, m! - 1\}. \tag{16}$$

Notice that the case of Fisher information measure is quite different since the sum has $m!-1$ terms. Hence, $h_i^{(F)}$ is defined for $i = 1, 2, \dots, m!-1$. All these functions are continuous and differentiable in their domain. It is worth noticing that the proposed methodology is non-applicable to the particular case where $q_i = 1/m!$ for all $i = 1, 2, \dots, m!$, since the functions (13), (14), (15), and (16) are constant and, thus, the corresponding matrix of partial derivatives is singular.

Let $\boldsymbol{J}^{(\mathcal{H})}$ be the matrix of partial derivatives of the functions $h_i^{(\mathcal{H})}$, for $\mathcal{H} \in \{S, T_\beta, R_\beta, F\}$. Thus, for $i, j = 1, 2, \dots, m!$, it holds that:

$$J_{ij}^{(S)} = \frac{\partial h_i^{(S)}}{\partial q_j} = \begin{cases} \log q_i + 1 & \text{if } j = i, \\ 0 & \text{otherwise;} \end{cases} \tag{17}$$

$$J_{ij}^{(T_\beta)} = \frac{\partial h_i^{(T_\beta)}}{\partial q_j} = \begin{cases} 1 - \beta q_i^{\beta-1} & \text{if } j = i, \\ 0 & \text{otherwise;} \end{cases} \tag{18}$$

$$J_{ij}^{(R_\beta)} = \frac{\partial h_i^{(R_\beta)}}{\partial q_j} = \begin{cases} \beta q_i^{\beta-1} & \text{if } j = i, \\ 0 & \text{otherwise;} \end{cases} \tag{19}$$

We have to treat the Fisher information measure separately. For $i = 1, 2, \dots, m! - 1$ and $j = 1, 2, \dots, m!$, it holds that:

$$J_{ij}^{(F)} = \frac{\partial h_i^{(F)}}{\partial q_j} = \begin{cases} \frac{\sqrt{q_{i+1}} - \sqrt{q_i}}{(-1)^{i+j-1}\sqrt{q_j}} & \text{if } j = i \text{ or } j = i+1, \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Since all these matrices of partial derivatives are non-singular in their domain, by (11), we conclude that for $\mathcal{H} \in \{S, T_\beta, R_\beta\}$,

$$\sqrt{n}\left[h_1^{(\mathcal{H})}(\widehat{\boldsymbol{q}}_n) - h_1^{(\mathcal{H})}(\boldsymbol{q}), h_2^{(\mathcal{H})}(\widehat{\boldsymbol{q}}_n) - h_2^{(\mathcal{H})}(\boldsymbol{q}), \dots, h_{m!}^{(\mathcal{H})}(\widehat{\boldsymbol{q}}_n) - h_{m!}^{(\mathcal{H})}(\boldsymbol{q})\right] \xrightarrow[n\to\infty]{D} \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{q}}^{(\mathcal{H})}\right). \tag{21}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(\mathcal{H})} = \boldsymbol{J}^{(\mathcal{H})}\boldsymbol{\Sigma}(\boldsymbol{J}^{(\mathcal{H})})^{\mathrm{T}}$. And,

$$\sqrt{n}\left[h_1^{(F)}(\hat{\boldsymbol{q}}_n)-h_1^{(F)}(\boldsymbol{q}), h_2^{(F)}(\hat{\boldsymbol{q}}_n)-h_2^{(F)}(\boldsymbol{q}), \ldots, h_{m!-1}^{(F)}(\hat{\boldsymbol{q}}_n)-h_{m!-1}^{(F)}(\boldsymbol{q})\right] \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{q}}^{(F)}\right), \tag{22}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(F)} = \boldsymbol{J}^{(F)}\boldsymbol{\Sigma}(\boldsymbol{J}^{(F)})^{\mathrm{T}}$. This matrix is of order $m! - 1$.

Since $\boldsymbol{J}^{(\mathcal{H})}$ is diagonal if $\mathcal{H} \in \{S, T_\beta, R_\beta\}$, for $1 \le i, j \le m!$, then it holds that

$$\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(\mathcal{H})}\right)_{ij} = \begin{cases} \left(J_{ii}^{(\mathcal{H})}\right)^2 \Sigma_{ii} & \text{if } i = j, \\ J_{ii}^{(\mathcal{H})} J_{jj}^{(\mathcal{H})} \Sigma_{ij} & \text{if } i \neq j. \end{cases} \tag{23}$$

In addition (see [41] for details), for $1 \le i, j \le m! - 1$, it is verified that

$$\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(F)}\right)_{ij} = \left(J_{ii}^{(F)}\Sigma_{ij} + J_{i,i+1}^{(F)}\Sigma_{i+1,j}\right)J_{jj}^{(F)} + \left(J_{ii}^{(F)}\Sigma_{i,j+1} + J_{i,i+1}^{(F)}\Sigma_{i+1,j+1}\right)J_{j,j+1}^{(F)}. \tag{24}$$

Shannon entropy. Since

$$H^{(S)}[\boldsymbol{q}] = \sum_{i=1}^{m!} \alpha_i h_i^{(S)}(\boldsymbol{q}), \tag{25}$$

with $\alpha_i = -1/\log(m!)$ for all $i = 1, 2, \ldots, m!$, by applying (12) to (21), we can state that

$$\sqrt{n}\left(H^{(S)}[\hat{\boldsymbol{q}}_n] - H^{(S)}[\boldsymbol{q}]\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, (\sigma_{\boldsymbol{q}}^{(S)})^2\right), \tag{26}$$

for

$$\left(\sigma_{\boldsymbol{q}}^{(S)}\right)^2 = \frac{1}{\log^2(m!)}\left[\sum_{i=1}^{m!}\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(S)}\right)_{ii} + 2\sum_{i=1}^{m!-1}\sum_{j=i+1}^{m!}\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(S)}\right)_{ij}\right] > 0. \tag{27}$$

This implies that for n sufficiently large, $H^{(S)}[\hat{\boldsymbol{q}}_n] \sim \mathcal{N}\left(H^{(S)}[\boldsymbol{q}], (\sigma_{\boldsymbol{q}}^{(S)})^2/n\right)$.

Tsallis entropy. Since

$$H^{(T_\beta)}[\boldsymbol{q}] = \sum_{i=1}^{m!} \alpha_i h_i^{(T_\beta)}(\boldsymbol{q}), \tag{28}$$

with $\alpha_i = 1/(1 - m!^{1-\beta})$ for all $i = 1, 2, \ldots, m!$, by applying (12) to (21), we can state that

$$\sqrt{n}\left(H^{(T_\beta)}[\hat{\boldsymbol{q}}_n] - H^{(T_\beta)}[\boldsymbol{q}]\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, (\sigma_{\boldsymbol{q}}^{(T_\beta)})^2\right), \tag{29}$$

for

$$\left(\sigma_{\boldsymbol{q}}^{(T_\beta)}\right)^2 = \frac{1}{(1 - m!^{1-\beta})^2}\left[\sum_{i=1}^{m!}\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(T_\beta)}\right)_{ii} + 2\sum_{i=1}^{m!-1}\sum_{j=i+1}^{m!}\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(T_\beta)}\right)_{ij}\right] > 0. \tag{30}$$

In other words, for n sufficiently large, $H^{(T_\beta)}[\hat{\boldsymbol{q}}_n] \sim \mathcal{N}\left(H^{(T_\beta)}[\boldsymbol{q}], (\sigma_{\boldsymbol{q}}^{(T_\beta)})^2/n\right)$.

Rényi entropy. This case is different. Let $f[\boldsymbol{q}] = \sum_{i=1}^{m!} h_i^{(R_\beta)}(\boldsymbol{q})$. By applying (12) to (21), it holds that

$$\sqrt{n}\left(f[\hat{\boldsymbol{q}}_n] - f[\boldsymbol{q}]\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, (\sigma_{\boldsymbol{q}}^{(R_\beta)})^2\right). \tag{31}$$

for

$$\left(\sigma_{\boldsymbol{q}}^{(R_\beta)}\right)^2 = \sum_{i=1}^{m!}\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(R_\beta)}\right)_{ii} + 2\sum_{i=1}^{m!-1}\sum_{j=i+1}^{m!}\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(R_\beta)}\right)_{ij} > 0. \tag{32}$$

This implies that if n is sufficiently large, then $f[\hat{\boldsymbol{q}}_n] \sim \mathcal{N}\left(f[\boldsymbol{q}], (\sigma_{\boldsymbol{q}}^{(R_\beta)})^2/n\right)$. Since

$$H^{(R_\beta)}[\boldsymbol{q}] = \frac{1}{(1-\beta)\log(m!)}\log\left(f[\boldsymbol{q}]\right), \tag{33}$$

for n is sufficiently large, $H^{(R_\beta)}[\hat{\boldsymbol{q}}_n] \sim \mathcal{P}_\beta$, where

$$\mathcal{P}_\beta(x) = \Pr(H^{(R_\beta)}[\hat{\boldsymbol{q}}_n] \le \lambda) = \boldsymbol{\Phi}\left(\frac{\exp[(1-\beta)\log(m!)\lambda] - f[\boldsymbol{q}]}{\sigma_{\boldsymbol{q}}^{(R_\beta)}/\sqrt{n}}\right), \tag{34}$$

and $\boldsymbol{\Phi}(x)$ denotes the cumulative distribution function of the standard Normal distribution. We can omit the modulus in the logarithm argument in (33) because $f[\boldsymbol{q}]$ is always positive.

Fisher information measure. Since

$$H^{(F)}[\boldsymbol{q}] = \sum_{i=1}^{m!-1} \alpha_i h_i^{(F)}(\boldsymbol{q}), \tag{35}$$

with $\alpha_i = 4$ for all $i = 1, 2, \ldots, m! - 1$, by applying (12) to (22), we can state that

$$\sqrt{n}\left(H^{(F)}[\hat{\boldsymbol{q}}_n] - H^{(F)}[\boldsymbol{q}]\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, (\sigma_{\boldsymbol{q}}^{(F)})^2\right). \tag{36}$$

for

$$\left(\sigma_{\boldsymbol{q}}^{(F)}\right)^2 = 16\left[\sum_{i=1}^{m!-1}\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(F)}\right)_{ii} + 2\sum_{i=1}^{m!-2}\sum_{j=i+1}^{m!-1}\left(\boldsymbol{\Sigma}_{\boldsymbol{q}}^{(F)}\right)_{ij}\right] > 0. \tag{37}$$

This is equivalent to say that, for n sufficiently large, $H^{(F)}[\hat{\boldsymbol{q}}_n] \sim \mathcal{N}\left(H^{(F)}[\boldsymbol{q}], (\sigma_{\boldsymbol{q}}^{(F)})^2/n\right)$.

In practical applications, there may be missing observed patterns. This may be due to forbidden patterns, as is the case of some chaotic maps, or to the fact that we are dealing with finite-size time series. We will refer to this situation as "missing patterns" because, in principle, the user does not know the nature of the system that produced the time series.

The present methodology has to be slightly modified for time series with missing patterns. For simplicity, let us suppose a unique $r \in \{1, 2, \ldots, m!\}$ exists such that the pattern π_r has zero probability of occurrence, i.e., $q_r = 0$. This implies that the functions $h_r^{(\mathcal{H})}$, for $\mathcal{H} \in \{S, T_\beta, R_\beta, F\}$ vanish and, thus, the matrices $\boldsymbol{J}^{(\mathcal{H})}$ are singular and violate the hypothesis of the Delta Method. However, this method can be applied to $\widetilde{\boldsymbol{J}}^{(\mathcal{H})}$, which is the matrix obtained from $\boldsymbol{J}^{(\mathcal{H})}$ by eliminating the rth row and the rth column. Since the pattern π_r never occurs, there is no transition from or towards this pattern, which means that the rth row and the rth column of $\mathbf{Q}^{(\ell)}$ are null for $\ell = 1, 2, \ldots, m-1$. Let $\widetilde{\mathbf{Q}}^{(\ell)}$ be the matrix obtained from $\mathbf{Q}^{(\ell)}$ by eliminating the rth row and the rth column, and $\widetilde{\boldsymbol{q}}$ be the vector obtained from $\boldsymbol{q}$ by eliminating the rth component. Finally, we consider $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{q}}^{(\mathcal{H})} = \widetilde{\boldsymbol{J}}^{(\mathcal{H})}\widetilde{\boldsymbol{\Sigma}}(\widetilde{\boldsymbol{J}}^{(\mathcal{H})})^{\mathrm{T}}$, where $\widetilde{\boldsymbol{\Sigma}}$ is obtained by Eq. (9) replacing $\boldsymbol{q}$ and $\mathbf{Q}^{(\ell)}$ by $\widetilde{\boldsymbol{q}}$ and $\widetilde{\mathbf{Q}}^{(\ell)}$, respectively. This reasoning can be extended in the case of several missing patterns, and translates immediately to the sample quantities.

5. Hypothesis test

In this section, we test whether two time series share the same underlying behavior. Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_{n_{\boldsymbol{x}}+m-1})$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_{n_{\boldsymbol{y}}+m-1})$ be two time series, with $n_{\boldsymbol{x}}$ and $n_{\boldsymbol{y}}$ sufficiently large. We find the sequences of ordinal patterns $\boldsymbol{\pi}(\boldsymbol{x})$ and $\boldsymbol{\pi}(\boldsymbol{y})$. Then, using the formula (7), we compute the vectors $\hat{\boldsymbol{q}}_{n_{\boldsymbol{x}}}(\boldsymbol{x})$ and $\hat{\boldsymbol{q}}_{n_{\boldsymbol{y}}}(\boldsymbol{y})$. If $\boldsymbol{x}$ and $\boldsymbol{y}$ have dynamics with the same distribution of ordinal patterns $\boldsymbol{q}$, we expect to observe a statistically insignificant difference between the values $H^{(\mathcal{H})}[\hat{\boldsymbol{q}}_{n_{\boldsymbol{x}}}(\boldsymbol{x})]$ and $H^{(\mathcal{H})}[\hat{\boldsymbol{q}}_{n_{\boldsymbol{y}}}(\boldsymbol{y})]$, for $\mathcal{H} \in \{S, T_\beta, R_\beta, F\}$.

We state the following hypothesis test:

$$\begin{cases} \mathscr{H}_0: & H^{(\mathcal{H})}[\boldsymbol{q}(\boldsymbol{x})] = H^{(\mathcal{H})}[\boldsymbol{q}(\boldsymbol{y})], \\ \mathscr{H}_1: & H^{(\mathcal{H})}[\boldsymbol{q}(\boldsymbol{x})] \neq H^{(\mathcal{H})}[\boldsymbol{q}(\boldsymbol{y})]. \end{cases} \tag{38}$$

For $\mathcal{H} \in \{S, T_\beta, F\}$, the proposed statistic is

$$W^{(\mathcal{H})} = \frac{H^{(\mathcal{H})}[\hat{\boldsymbol{q}}_{n_{\boldsymbol{x}}}(\boldsymbol{x})] - H^{(\mathcal{H})}[\hat{\boldsymbol{q}}_{n_{\boldsymbol{y}}}(\boldsymbol{y})]}{\sqrt{\frac{\left(\sigma_{\hat{\boldsymbol{q}}_{n_{\boldsymbol{x}}}(\boldsymbol{x})}^{(\mathcal{H})}\right)^2}{n_{\boldsymbol{x}}} + \frac{\left(\sigma_{\hat{\boldsymbol{q}}_{n_{\boldsymbol{y}}}(\boldsymbol{y})}^{(\mathcal{H})}\right)^2}{n_{\boldsymbol{y}}}}} \sim \mathcal{N}(0, 1). \tag{39}$$

Again, the Rényi entropy must be treated in a different way because its asymptotic distribution does not verify normality. From expression (33), it is straightforward by injectivity that $\mathscr{H}_0: H^{(R_\beta)}[\boldsymbol{q}(\boldsymbol{x})] =$

$H^{(R_\beta)}[\boldsymbol{q}(\boldsymbol{y})]$ is verified if and only if $f[\boldsymbol{q}(\boldsymbol{x})] = f[\boldsymbol{q}(\boldsymbol{y})]$. Using the asymptotic distribution given by (31), the proposed statistic is:

$$W^{(R_\beta)} = \frac{f[\hat{\boldsymbol{q}}_{n_x}(\boldsymbol{x})] - f[\hat{\boldsymbol{q}}_{n_y}(\boldsymbol{y})]}{\sqrt{\frac{\left(\sigma^{(R_\beta)}_{\hat{q}_{n_x}(x)}\right)^2}{n_x} + \frac{\left(\sigma^{(R_\beta)}_{\hat{q}_{n_y}(y)}\right)^2}{n_y}}} \sim \mathcal{N}(0,1). \tag{40}$$

The statistics introduced in (39) and (40) are analogous to the one derived from the test that states the equal diversity of two independent populations [32, Theorem 2.2].

6. Finite sample size assessment

In this section, we present experiments to evaluate the performance of the theoretical results introduced in this work. We first describe the time series processes and dynamics that will be used for this aim.

The random walk (RW) is a process generated by the recursive formula

$$x_t = x_{t-1} + \xi_t, \tag{41}$$

where $\xi_t \sim \mathcal{N}(0,1)$. In [46, Eqs. (41) and (43), respectively], the authors prove that the theoretical vector of ordinal pattern probabilities is

$$\boldsymbol{q}(\mathrm{RW}) = \left(\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{4}\right), \tag{42}$$

and that the theoretical covariance matrix is

$$\boldsymbol{\Sigma}(\mathrm{RW}) = \frac{1}{192}\begin{pmatrix} 60 & -6 & -6 & -6 & -6 & -36 \\ -6 & 15 & 7 & -9 & -1 & -6 \\ -6 & 7 & 15 & -1 & -9 & -6 \\ -6 & -9 & -1 & 15 & 7 & -6 \\ -6 & -1 & -9 & 7 & 15 & -6 \\ -36 & -6 & -6 & -6 & -6 & 60 \end{pmatrix}. \tag{43}$$

Notice that we had to adapt these results to the way the ordinal patterns are labeled.

Another process regards a time series whose values are independent and can be 1 or −1 with equal probability. We denote this process as "TS±1". In Ref. [37, page 11], the author computes the theoretical ordinal pattern distribution which is given by the vector

$$\boldsymbol{q}(\mathrm{TS}\pm1) = \left(\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, 0\right), \tag{44}$$

as well as the theoretical covariance matrix given by

$$\boldsymbol{\Sigma}(\mathrm{TS}\pm1) = \frac{1}{64}\begin{pmatrix} 16 & -8 & -8 & 0 & 0 & 0 \\ -8 & 7 & 3 & -3 & 1 & 0 \\ -8 & 3 & 7 & 1 & -3 & 0 \\ 0 & -3 & 1 & 3 & -1 & 0 \\ 0 & 1 & -3 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{45}$$

Notice that in this case, the pattern π_6 is missing and, hence, we have to consider the modification introduced at the end of Section 4 that leads us to work with

$$\widetilde{\boldsymbol{q}}(\mathrm{TS}\pm1) = \left(\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\right) \tag{46}$$

and

$$\widetilde{\boldsymbol{\Sigma}}(\mathrm{TS}\pm1) = \frac{1}{64}\begin{pmatrix} 16 & -8 & -8 & 0 & 0 \\ -8 & 7 & 3 & -3 & 1 \\ -8 & 3 & 7 & 1 & -3 \\ 0 & -3 & 1 & 3 & -1 \\ 0 & 1 & -3 & -1 & 3 \end{pmatrix}. \tag{47}$$

We also consider time series simulated under the following dynamic behaviors.

1. **Uniform White Noise (UWN).** The time series observations are uniformly distributed in the interval $[0,1]$. Although the present proposal is not applicable to "perfect" WN since its entropy converges to a χ^2 distribution, we applied our methodology because the exact equiprobable distribution of patterns is seldom observed in practice, even when the signal is WN.
2. **Chaotic Logistic Map (CLM).** The logistic map is generated by the recursion $x_{t+1} = \lambda x_t(1-x_t)$, where $x_t \in (0,1)$ and $t = 0, 1, \ldots$. For $\lambda \in (3.57, 4)$, this map is chaotic, and one of the ordinal patterns is missing, since the time series does not contain a window $\boldsymbol{s}_t$ of the form $x_{t+2} \leq x_{t+1} \leq x_t$ [49]. As we mention in Section 4, our methodology needs a modification in this case. In this work, we use $\lambda = 3.7$ and $x_0 \in [0.1, 0.9]$.
3. **Factional Brownian Motion (FBM).** A real stochastic process $B^{\mathscr{H}}(t)$, with Hurst exponent $0 < \mathscr{H} < 1$ and $t \geq 0$, is a fractional Brownian motion if
   (a) $B^{\mathscr{H}}(0) = 0$,
   (b) $\mathrm{E}[B^{\mathscr{H}}(t)] = 0$ for all $t \geq 0$, and
   (c) $\mathrm{E}[B^{\mathscr{H}}(t)B^{\mathscr{H}}(s)] = \left(t^{2\mathscr{H}} - s^{2\mathscr{H}} - |t-s|^{2\mathscr{H}}\right)/2$, for all $t, s \geq 0$.

   In this work, we use $\mathscr{H} = 0.7$.
4. f^{-k} **noises** (k**N**). The time series is a process whose power spectral density is $S(f) \propto 1/f^k$. In this work, we use $k \in \{0.5, 1, 1.5, 2, 2.5\}$.

6.1. Numerical validation

This section provides a numerical validation of the mathematical derivation developed in Section 4. For this reason, we conducted a Monte Carlo numerical experiment for the two processes, RW and TS±1. Throughout this section, we consider an embedding dimension equal to 3.

We generated $R = 10^5$ time series of length $N = 10^5$ following each one of the two processes. Then, for each time series, we computed the entropies $H^{(\mathcal{H})}$ for $\mathcal{H} \in \{S, T_{1/2}, R_{1/3}, F\}$. The theoretical asymptotic means and variances related to each process are presented in Table 2. Figs. 1 and 2 show the corresponding empirical densities and histograms using the Freedman–Diaconis rule [50].In all cases, we can appreciate the proper fitting of the asymptotic distribution model proposed in this work.

Table 2
Theoretical asymptotic mean and variances of two time series process using $m = 3$, where $\boldsymbol{q} = \boldsymbol{q}(\mathrm{RW})$ and $\boldsymbol{q} = \boldsymbol{q}(\mathrm{TS}\pm1)$ for the first and second processes, respectively.

Process	Entropy	Asymptotic mean	Asymptotic variance
RW	Shannon	$H^{(S)}[\boldsymbol{q}] = 0.9671320$	$(\sigma^{(S)}_{\boldsymbol{q}})^2 = 0.0374138$
	Tsallis	$H^{(T_{1/2})}[\boldsymbol{q}] = 0.9756630$	$(\sigma^{(T_{1/2})}_{\boldsymbol{q}})^2 = 0.0204154$
	Rényi	$f[\boldsymbol{q}] = 3.2599210$	$(\sigma^{(R_{1/3})}_{\boldsymbol{q}})^2 = 0.0608574$
	Fisher	$H^{(F)}[\boldsymbol{q}] = 0.1715729$	$(\sigma^{(F)}_{\boldsymbol{q}})^2 = 1.2287640$
TS±1	Shannon	$H^{(S)}[\boldsymbol{q}] = 0.7737056$	$(\sigma^{(S)}_{\boldsymbol{q}})^2 = 0.1496551$
	Tsallis	$H^{(T_{1/2})}[\boldsymbol{q}] = 0.7735966$	$(\sigma^{(T_{1/2})}_{\boldsymbol{q}})^2 = 0.0594949$
	Rényi	$f[\boldsymbol{q}] = 2.7937010$	$(\sigma^{(R_{1/3})}_{\boldsymbol{q}})^2 = 0.1616843$
	Fisher	$H^{(F)}[\boldsymbol{q}] = 1.0000000$	$(\sigma^{(F)}_{\boldsymbol{q}})^2 = 5.0000000$

6.2. Comparison with other methods

The Shannon entropy's asymptotic distribution has already been studied in several works, especially concerning finite-sample sizes. For instance, Ricci [51] proves that the Shannon entropy of a process obtained by a regular Markov chain is asymptotically normally distributed. To compare our proposal with this result, we need the theoretical ordinal patterns transition matrices, for which we need $\Pr(\pi_t = \pi^{(i)} \wedge \pi_{t+1} = \pi^{(j)}) = \Pr(\pi_{t+1} = \pi^{(j)} \mid \pi_t = \pi^{(i)}) \cdot \Pr(\pi_t = \pi^{(i)})$, where the conditional probability is given by the transition matrix that defines the Markov chain but $\Pr(\pi_t = \pi^{(i)})$ depends on the instant t. This interesting case is out of the scope of the present proposal.

Table 3
Comparison with means and variances of the Shannon entropy for the two processes obtained by Basharin's method.

Process	Measure	Empirical	Basharin	Our proposal
RW	Mean	0.9669819	0.9669925	0.9671320
	Variance	0.0000039	0.0000037	0.0000037
TS ± 1	Mean	0.7735922	0.7735661	0.7737056
	Variance	0.0000150	0.0000150	0.0000150

On the other hand, Basharin [52] found an asymptotic distribution for the Shannon entropy given by Eq. (2) using the logarithm in base 2. If we adapt this result to our context, for n sufficiently large it holds that $H^{(S)}[\widehat{\boldsymbol{q}}_n] \sim \mathcal{N}\left(\mu_{\text{Basharin}}, \sigma^2_{\text{Basharin}}\right)$, where

$$\mu_{\text{Basharin}} = H^{(S)}[\boldsymbol{q}] - \frac{m! - 1}{2n\ln(m!)}, \text{ and} \tag{48}$$

$$\sigma^2_{\text{Basharin}} = \frac{1}{n\ln^2(m!)}\left[\sum_{i=1}^{m!} q_i \ln^2(q_i) - \left(S^{(S)}[\boldsymbol{q}]\right)^2\right]. \tag{49}$$

We computed the means and variances of the Shannon entropy for the two processes described in this section using the method proposed by Basharin. Then, we compare them with our proposal and with the empirical values. Table 3 shows these results. The fitting of the two asymptotic models is similar, as can be seen in Fig. 3.

6.3. Convergence analysis

In this section, we analyze the empirical convergence behavior of the asymptotic variance in terms of the ordinal patterns sequence length. We checked that its value stabilizes and that ensembles of entropies can be described as normal samples.

Firstly, we considered a time series $\boldsymbol{x}$ simulated under the dynamics CLM, FBM, and f^{-k} noises for $k \in \{0.5, 1, 1.5, 2, 2.5\}$. Then, we computed the asymptotic variance estimations $\sigma^{(H)}_{\widehat{q}_n(\boldsymbol{x})}/n$ for $n \in \{100, 500, 1000, 2000, \ldots, 25000\}$, $\mathcal{H} \in \{S, T_{1/2}, R_{1/3}, F\}$, and $m = 3$. Fig. 4 shows the obtained results, where we can observe a stabilization. In general, the Fisher information measure variances are considerably larger than the other information measures. The fractional Brownian motion presents the largest variances in all cases apart from the Tsallis entropy, for which the chaotic logistic map achieves the most significant values. The variance of f^{-k} noises for $k = 0.5, 1$ approaches zero as n tends to infinity in all cases. Except for the Fisher information measure, this property is also verified by $f^{-1.5}$ noises. The variances of the f^{-k} noises for $k = 2, 2.5$ seem to approach a value near zero for Shannon, Tsallis, and Rényi entropies. We notice that the variances increase as the value of k also increases.

To assess the convergence rate, we found the value of n for which the difference between the asymptotic variance estimations in two consecutive steps is less than 10^{-3}. The corresponding values for each entropy measure are presented in Table 4. We see that the Fisher information measure required a larger n to stabilize, followed by the Rényi entropy. Meanwhile, the Tsallis entropy generally achieved an asymptotic behavior for smaller values of n, followed by the Shannon entropy.

In the second stage, we conducted a Monte Carlo study to assess the speed of convergence of the permutation entropy to the limit distribution. Algorithm 1 outlines the structure of the code. Given a significance level η, the types of noise, the embedding dimensions, the time series lengths, we produce N independent time series to form an ensemble of Shannon permutation entropy values. This ensemble is submitted to the Shapiro–Wilk test [53] to verify the hypothesis of normality, and we obtain a p-value. This procedure is repeated a number of replications R, and we verify the percentage of p-values lower than the test significance η.

Table 4
Values of n for which the difference between two consecutive estimations of the asymptotic variance of each entropy measure is less than 10^{-3}. Hyphens indicate that this condition is not satisfied by $n \leq 25000$.

Dynamic	Shannon	Tsallis	Rényi	Fisher
CLM	500	7000	6000	22 000
FBM	13 000	4000	3000	–
0.5N	2000	500	4000	5000
1.0N	2000	1000	2000	13 000
1.5N	1000	500	2000	18 000
2.0N	4000	2000	6000	15 000
2.5N	6000	500	20 000	–

Algorithm 1: Convergence of the permutation entropy to the limit distribution

Input: Seed
Input: Types of f^{-k} noise (values of k); embedding dimensions m; time series lengths $n + m - 1$; number of time series to form the ensemble of permutation entropy values N; number of replications R; test significance η
Result: Number of replications for which the normal hypotheses for the ensemble are rejected

```
for each k do
    for each m do
        for each n + m − 1 do
            for 1 ≤ i ≤ R do
                for 1 ≤ j ≤ N do
                    Generate x, a time series of length n + m − 1 of noise of type k;
                    Compute π, the sequence of ordinal patterns of x;
                    Compute p̂, the proportions of patterns in π;
                    Compute H = H^(S)[p̂], the permutation entropy of the proportions;
                end
                H = (H_1, H_2, ..., H_N) is a vector of N independent values of the permutation entropy from time series of type k, embedding dimension m, and length n + m − 1;
                Compute the Shapiro–Wilk test of H and obtain the p-value;
            end
            pv = (pv_1, pv_2, ..., pv_R) is a vector of size R of independent p-values of the hypothesis of normality for the Shannon permutation entropy;
            Compute the proportion of entries of pv smaller than η;
        end
    end
end
```

We used $R = 100$ replications of $N = 1000$ independent time series of lengths $\{500, 1000, 5000, 10000, 50000\}$. The results are displayed in Fig. 5 where each row is a type of f^{-k} noise ($k \in \{0.5, 1, 1.5, 2, 2.5\}$) and each column is an embedding dimension ($m \in \{3, 4, 5, 6\}$). Notice that if the time series has 500 observations, the choice $m = 6$ is inadequate because $m! = 720$ is larger than the number of observations. Although this case has no statistical or practical significance, we include it for completeness in the plots.

Fig. 5 shows that the convergence to the asymptotic result is slow for f^{-k} noise with $k = 0.5, 1$ and small embedding dimensions ($m = 3, 4$). There is no evidence against the Normal distribution as a model

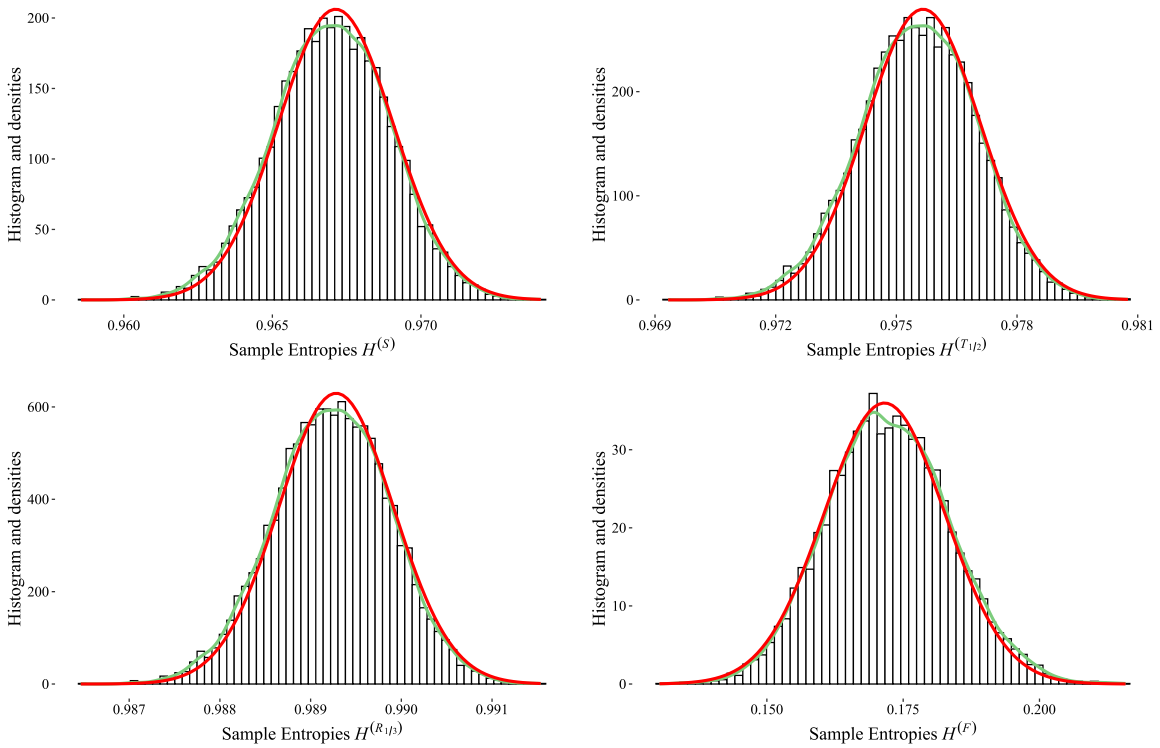

Fig. 1. Histograms of 10^5 samples of different entropies using $m = 3$, for the random walk, along with the empirical densities (green lines) and asymptotic models (red lines). Shannon (top left), Tsallis with $\beta = 1/2$ (top right), Rényi with $\beta = 1/3$ (bottom left), and Fisher information measure (bottom right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for the permutation entropy with time series with more than 1000 observations.

6.4. Hypothesis test assessment

We consider the eight groups of time series represented by the mentioned dynamic behaviors: UWN, CLM, FBM, and kN for $k \in \{0.5, 1.0, 1.5, 2.0, 2.5\}$. For each group, we generated 100 time series of length 300, 3000, and 30000. Then, we applied the proposed hypothesis tests to contrast these types of dynamics for an embedding dimension equal to 3, using $\beta = 1/2$ for Tsallis entropy and $\beta = 1/3$ for Rényi entropy. More precisely, let G and G' be two of these groups with equal-length time series. We contrasted each time series in G versus each time series in G', giving a total of 10000 hypothesis tests, and registered the proportion of p-values less than $\eta = 0.05$. This percentage of rejection is saved in the cell corresponding to both groups. Figs. 6, 7, 8, and 9 show the percentage of rejection in a pairwise comparison. The perfect case is represented by a table whose diagonal is green (no rejection within the same group) and red cells in the upper triangle (total rejection between different groups). We also estimated the power of the test in each case, i.e., the probability of rejecting the null hypothesis when it is false. Table 5 contains these results, where the power of the tests when comparing long-term time series in decreasing order are: 0.9517 for Fisher information measure, 0.9512 for Shannon entropy, 0.9159 for Rényi entropy with $\beta = 1/3$, and 0.9081 for Tsallis entropy with $\beta = 1/2$.

The obtained results show that Tsallis entropy is not effective at distinguishing different dynamics in short-length time series. On the other hand, this entropy performed ideally in identifying middle- and long-length time series within the same group. Even though all entropies showed an adequate behavior when comparing different groups of long-length time series, Fisher information measure, followed by Shannon entropy, achieved the best results.

Table 5
Estimation of the power of the proposed test using an embedding dimension equal to 3 and different time series' lengths.

Time series' length	Shannon	Tsallis	Rényi	Fisher
300	0.5989	0.2697	0.5738	0.5831
3000	0.8562	0.7042	0.8458	0.8452
30000	0.9512	0.9081	0.9159	0.9517

FBM is deemed compatible with f^{-k} noise for $k = 0.5$ by all entropy measures and for $k = 2.5$ by Shannon, Tsallis, and Rényi entropies. The test seems to confuse these f^{-k} noises with either the fractional Brownian motion or the fractional Gaussian noise (FGN) defined by the FBM increments. Recall that power spectral density for FBM and FGN are $S(f) \propto 1/f^{2\mathscr{H}+1}$ and $S(f) \propto 1/f^{2\mathscr{H}-1}$, respectively [54]. Since we use $\mathscr{H} = 0.7$, these spectra are of the form $1/f^{2.4}$ and $1/f^{0.4}$.

6.5. Applications to real data

In this section, we apply the proposed test to real data from the dataset Cryptocurrency Prices Data [55]. This dataset contains the historical daily open, high, low, and close prices for 56 cryptocurrencies from May 2013 to October 2022. We selected the following cryptocurrencies: Bitcoin (https://bitcoin.org/), Ethereum Classic (https://ethereumclassic.org/), Litecoin (http://www.litecoin.org/), and Tether (https://tether.to/). The time series of the open prices of these cryptocurrencies are shown in the top row of Fig. 10, and have lengths equal to 3248 (Bitcoin), 2072 (Ethereum Classic), 3248 (Litecoin), and 2582 (Tether). We can observe a similar shape in terms of pronounced peaks between Bitcoin and Litecoin. Meanwhile, Tether shows the most different behavior. These remarks are supported by the ordinal pattern frequency plots exhibited in the bottom row of Fig. 10, where we used the following notation: $\pi^{(1)} = (0, 1, 2)$, $\pi^{(2)} = (0, 2, 2)$, $\pi^{(3)} = (1, 0, 2)$, $\pi^{(4)} = (1, 2, 0)$, $\pi^{(5)} = (2, 0, 1)$, and $\pi^{(6)} = (2, 1, 0)$. In Bitcoin and litecoin,

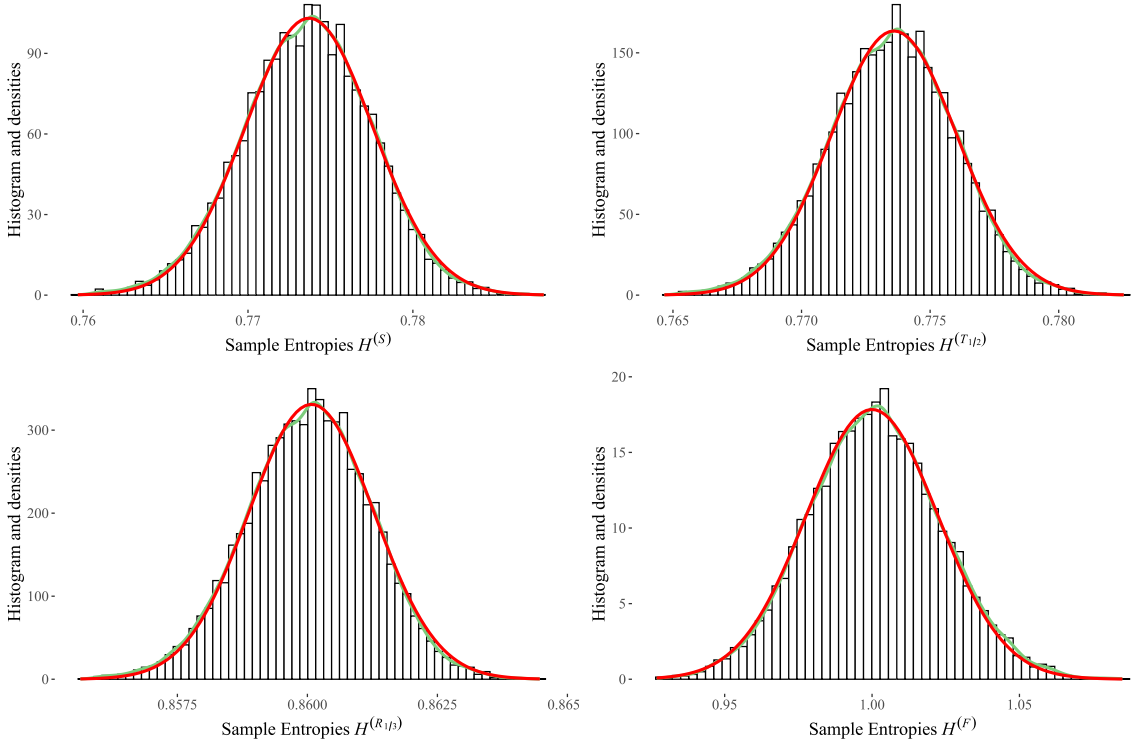

Fig. 2. Histograms of 10^5 samples of different entropies using $m = 3$, for the time series with ± 1 values, along with the empirical densities (green lines) and asymptotic models (red lines). Shannon (top left), Tsallis with $\beta = 1/2$ (top right), Rényi with $\beta = 1/3$ (bottom left), and Fisher information measure (bottom right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

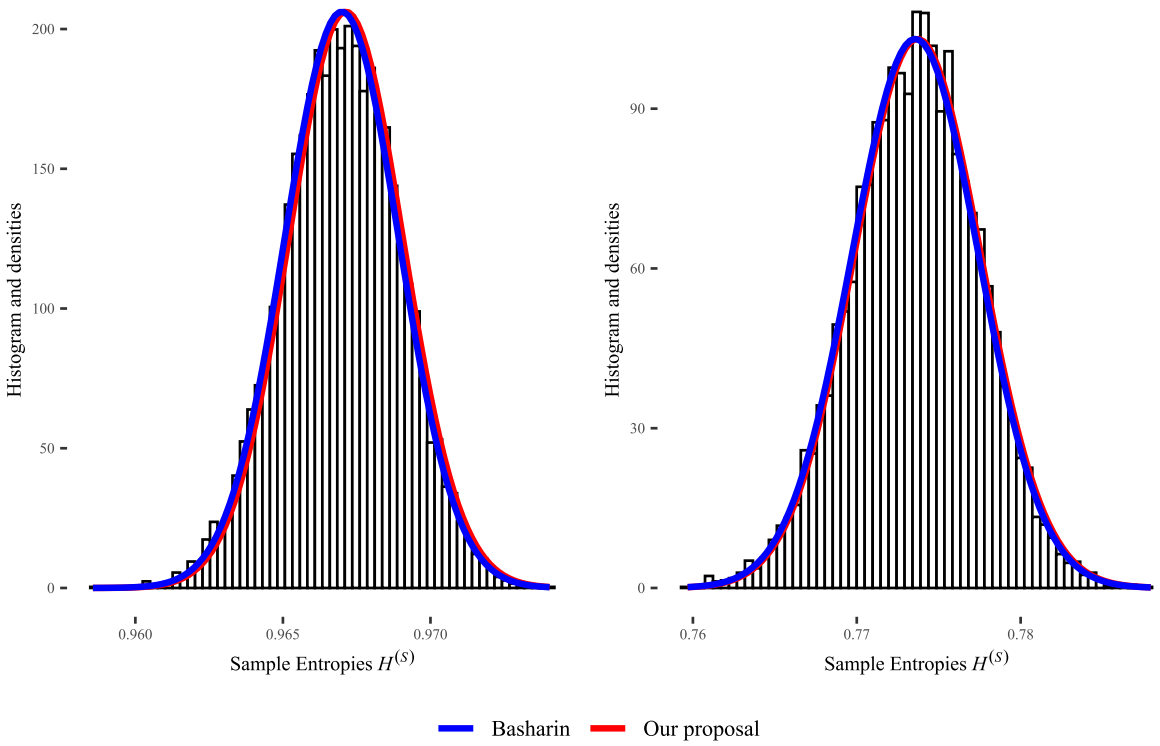

Fig. 3. Comparison with Basharin's method to obtain the asymptotic distribution of the Shannon entropy for RW (left) and for TS± 1 (right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

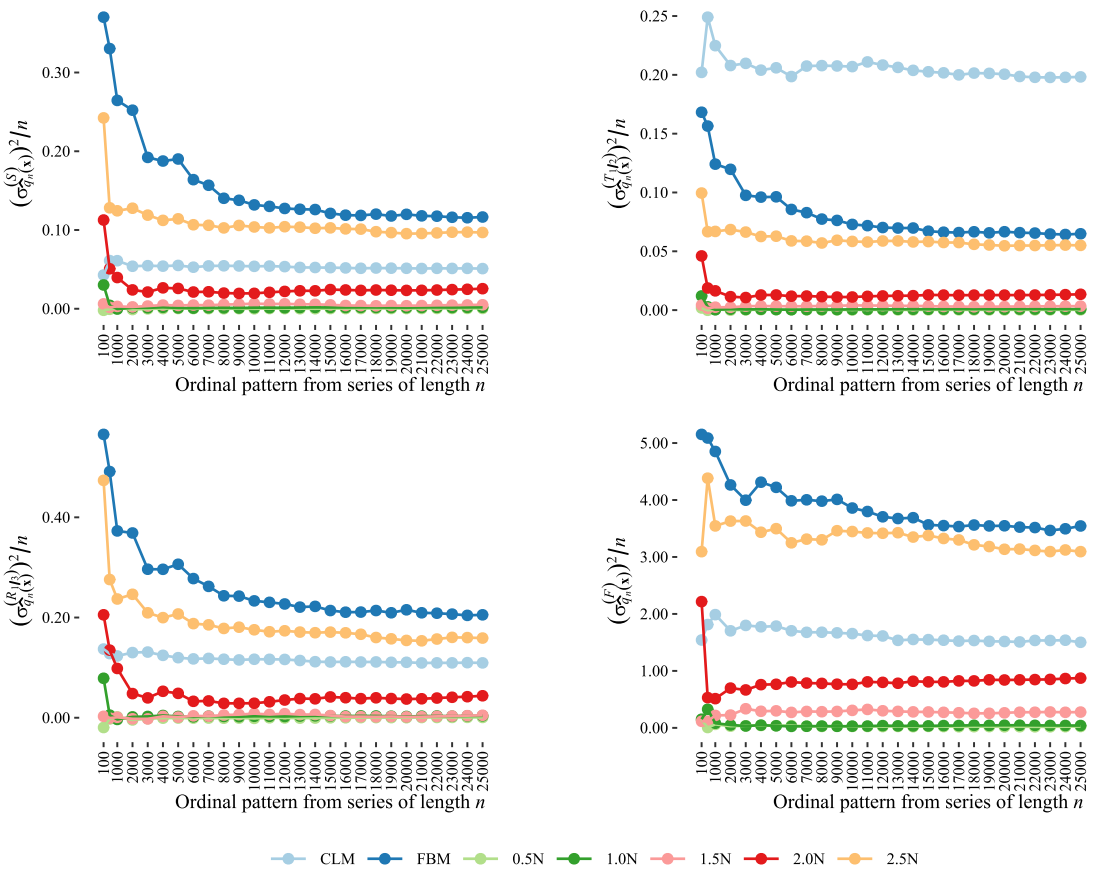

Fig. 4. Evolution of the asymptotic variance estimations $\sigma^{(\mathcal{H})}_{\hat{q}_n(\mathbf{x})}/n$ for $\mathcal{H} = S$ (top left), $\mathcal{H} = T_{1/2}$ (top right), $\mathcal{H} = R_{1/3}$ (bottom left), and $\mathcal{H} = F$ (bottom right), considering different types of dynamic behaviors.

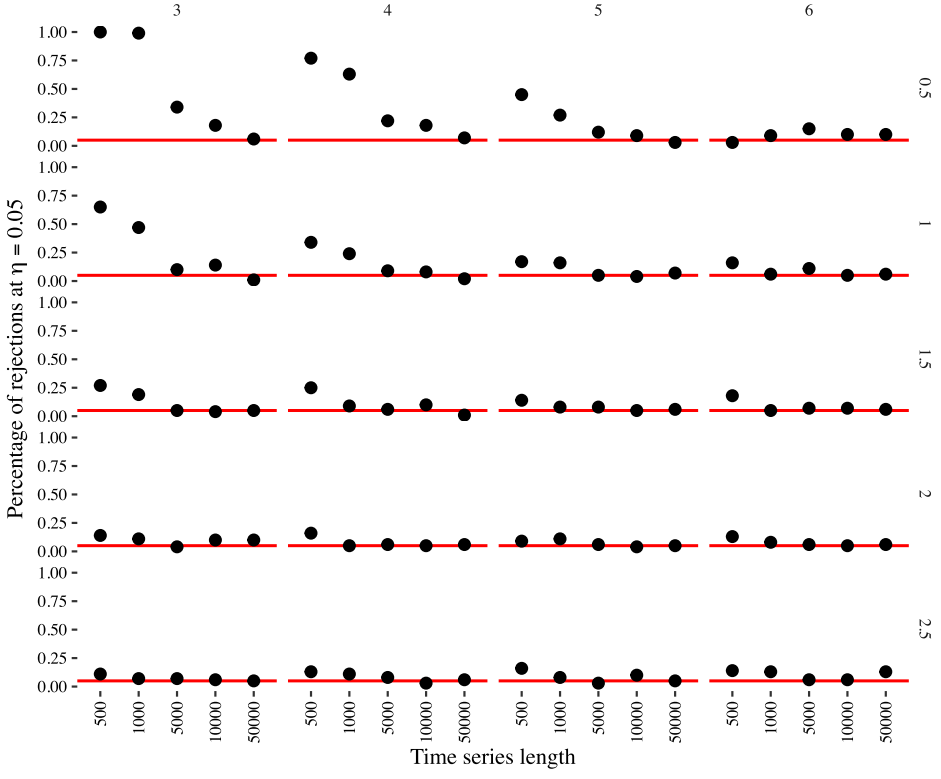

Fig. 5. Percentage of ensembles with p-values smaller than 0.05 for the normality hypothesis in one hundred replications for each combination of factors: type of noise (rows) and embedding dimension (columns).

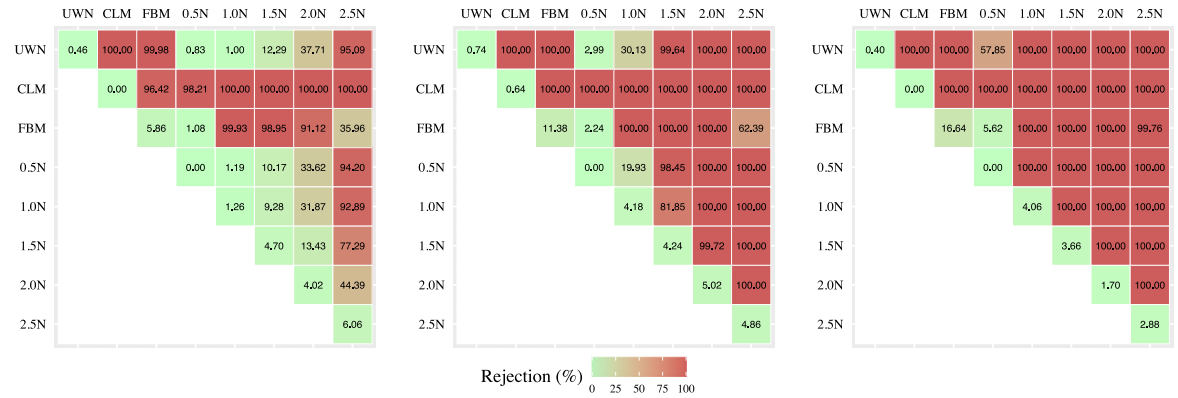

Fig. 6. Percentage of rejections when the hypothesis test (38) is applied using the Shannon entropy ($\mathcal{H} = S$) and embedding dimension equal to 3. Groups with time series of length 300, 3000 and 30000 are represented in left, middle, and right tables, respectively.

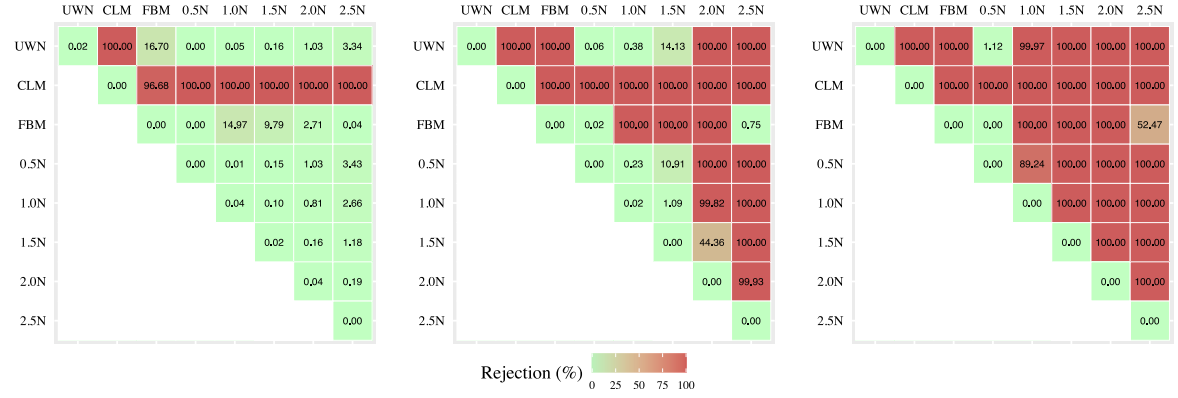

Fig. 7. Percentage of rejections when the hypothesis test (38) is applied using the Tsallis entropy with $\beta = 1/2$ ($\mathcal{H} = T_{1/2}$) and embedding dimension equal to 3. Groups with time series of length 300, 3000 and 30000 are represented in left, middle, and right tables, respectively.

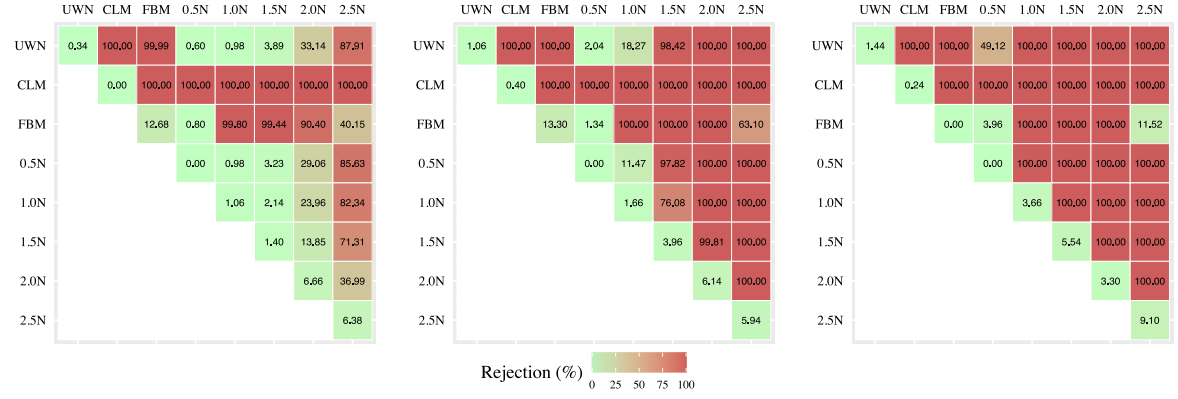

Fig. 8. Percentage of rejections when the hypothesis test (38) is applied using the Rényi entropy with $\beta = 1/3$ ($\mathcal{H} = R_{1/3}$) and embedding dimension equal to 3. Groups with time series of length 300, 3000, and 30000 are represented in left, middle, and right tables, respectively.

the patters $\pi^{(1)}$ and $\pi^{(6)}$ are the most frequent, followed by the pair $\pi^{(2)}$ and $\pi^{(4)}$, and finally by the pair $\pi^{(3)}$ and $\pi^{(4)}$. In Ethereum, the patterns $\pi^{(1)}$ and $\pi^{(6)}$ are again the most frequent, but the rest of the ordinal patterns show a similar frequency. In Tether, $\pi^{(1)}$ is the most frequent, and the rest ordinal patterns seem equiprobable.

We then applied the entropy test for the Shannon entropy and $m = 3$ to give the previous remarks statistical significance. Figs. 11, 12, 13, and 14 show the p-values and decisions at the 5 % confidence level. The decisions are also coded in colors: red shows rejection, while green indicates insufficient evidence to reject the null hypothesis that the series came from the same process. We can observe from Fig. 11 that for the Shannon entropy, the null hypothesis is always rejected (as expected) except for the comparisons of Bitcoin vs. Litecoin (as expected), and Litecoin vs. Ethereum Classic (with a p-value not much larger than 0.05). For the Tsallis entropy with $\beta = 1/2$, the decisions coincide with using the Shannon entropy as shown in Fig. 12. We replicated the procedure for the Rényi entropy with $\beta = 1/3$, and the "ideal" results are presented in Fig. 13. Fig. 14 exhibits the results for the Fisher information measure, which could not distinguish the subtle difference between Ethereum Classic and Bitcoin or Litecoin.

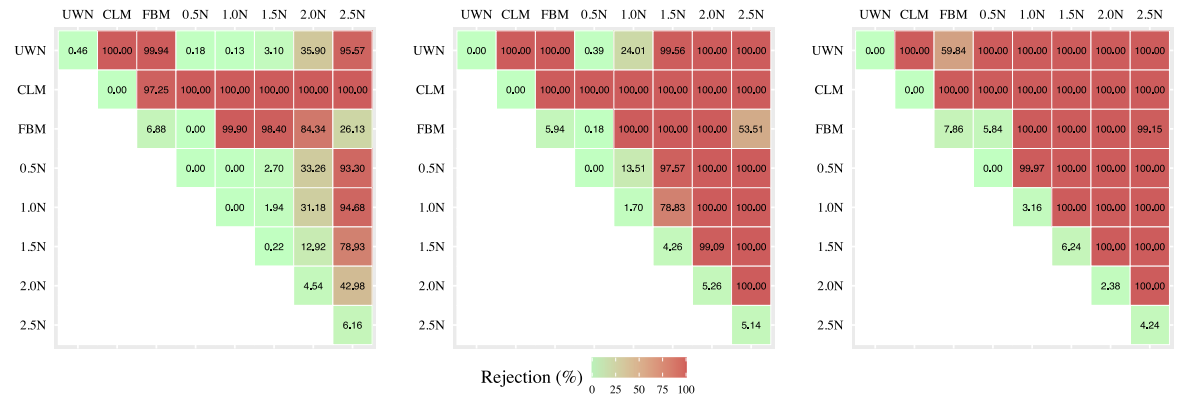

Fig. 9. Percentage of rejections when the hypothesis test (38) is applied using the Fisher information measure ($\mathcal{H} = F$) and embedding dimension equal to 3. Groups with time series of length 300, 3000 and 30000 are represented in left, middle, and right tables, respectively.

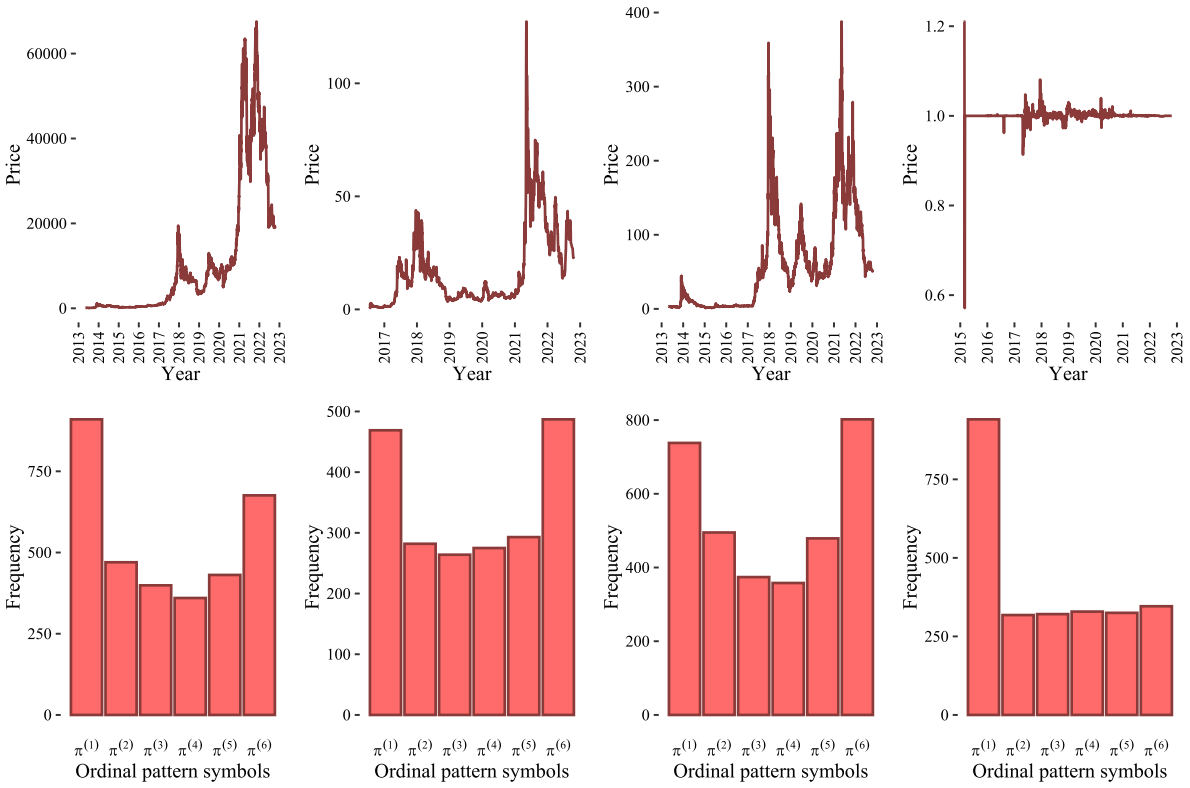

Fig. 10. Time series (top row) and the frequency of ordinal patterns (bottom row) for the open prices of the cryptocurrencies Bitcoin, Ethereum Classic, Litecoin, and Tether (from left to right).

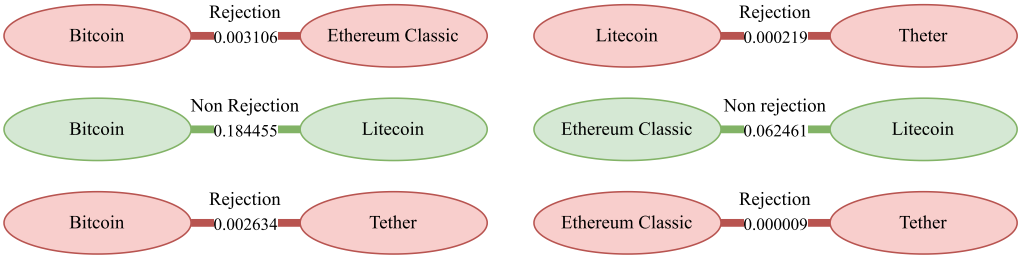

Fig. 11. Decisions at the 95 % confidence level and p-values of applying the Shannon entropy hypothesis test with $m = 3$ to the times series given by the opening prices of cryptocurrencies.

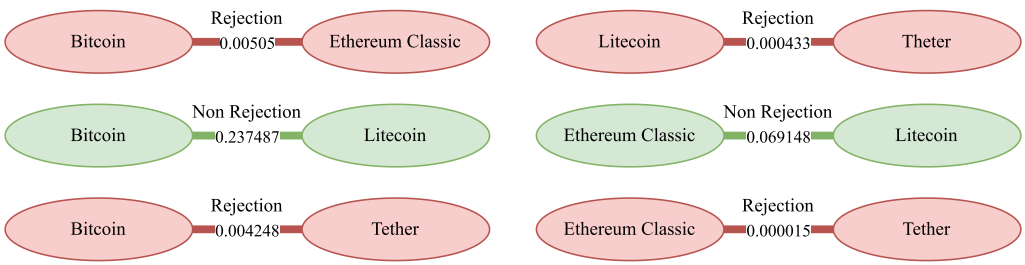

Fig. 12. Decisions at the 95 % confidence level and p-values of applying the Tsallis entropy hypothesis test with $\beta = 1/2$ and $m = 3$ to the times series given by the opening prices of cryptocurrencies.

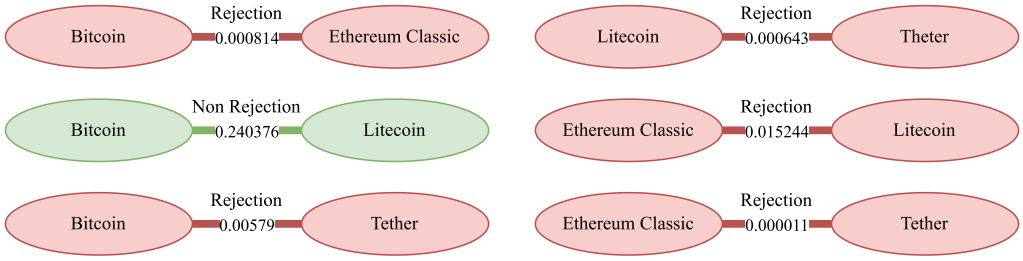

Fig. 13. Decisions at the 95 % confidence level and p-values of applying the Rényi entropy hypothesis test with $\beta = 1/3$ and $m = 3$ to the times series given by the opening prices of cryptocurrencies.

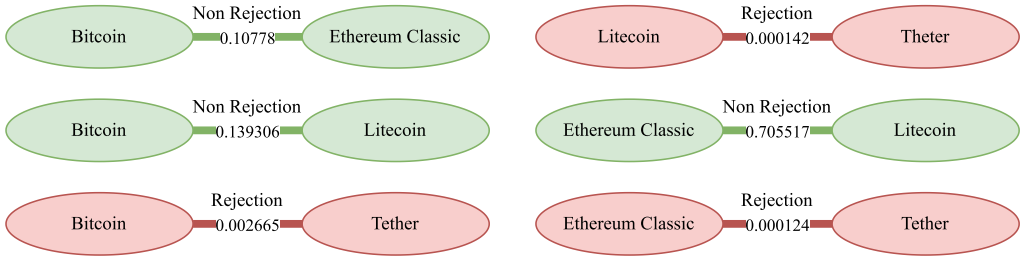

Fig. 14. Decisions at the 95 % confidence level and p-values of applying the Fisher information measure hypothesis test with $m = 3$ to the times series given by the opening prices of cryptocurrencies.

7. Conclusions

We have derived the asymptotic distribution of the Rényi and Tsallis/Havrda-Charvát entropies and the Fisher information measure of ordinal patterns embedding their serial correlation.

We verified that the convergence to the asymptotic results is slow for short time series with weak correlation structure and small embedding dimension ($m = 3, 4$). This indicates that the asymptotic results should be used with caution in such cases. The Fisher information measure showed the poorest convergence behavior. Overall, the limit distributions we provide have good performance in other situations.

Tests derived from those asymptotic results can be used to distinguish among several types of stochastic structures , with over than 90 % of correct rejection for long-term time series. We verified that the Fisher information measure achieved the best results in terms of size and power, as shown in Figs. 6, 7, 8, 9 and Table 5.

The Rényi entropy showed the desired behavior when the test was applied to real data concerning the open prices of cryptocurrencies. The Shannon and Tsallis entropies produced the same results. On the other side, the Fisher information measure could not distinguish differences between some of the time series under study.

Reproducibility and replicability

The R package that implements the functions discussed in this work, along with examples, is available at https://github.com/arey1911/StatOrdPattHxC.

CRediT authorship contribution statement

Andrea Rey: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Alejandro C. Frery:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Juliana Gambini:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Magdalena Lucini:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

The code is available through a Github repository mentioned in the text.

Acknowledgment

This work was partially supported by the Research Trust of Victoria University of Wellington, New Zealand, project 410695.

References

[1] Clausius R. The mechanical theory of heat – with its applications to the steam engine and to physical properties of bodies. London: John van Voorst; 1867.
[2] Boltzmann L. Lectures on gas theory. University of California Press; 1964.
[3] Planck M. On the law of the energy distribution in the normal spectrum. Ann Phys 1901;4:1–11.
[4] Gibbs JW. Elementary principles in statistical mechanics: developed with especial reference to the rational foundations of thermodynamics. C. Scribner's sons; 1902.
[5] Shannon CE. A mathematical theory of communications. Bell Syst Technol J 1948;27:379–423. http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x.
[6] Frery AC, Alpala J, Nascimento ADC. Identifying heterogeneity in SAR data with new test statistics. Remote Sens 2024;16. http://dx.doi.org/10.3390/rs16111973.
[7] Amigó JM, Keller K, Unakafova V. On entropy, entropy-like quantities, and applications. Contemporary mathematics and its applications: monographs, expositions and lecture notes, vol. 4, World Scientific; 2023, p. 197–231. http://dx.doi.org/10.1142/12920.
[8] Ribeiro M, Henriques T, Castro L, Souto A, Antunes L, Costa-Santos C, et al. The entropy universe. Entropy 2021;23:222.
[9] Amigó JM, Balogh SG, Hernández S. A brief review of generalized entropies. Entropy 2018;20:813.
[10] Namdari A, Li Z. A review of entropy measures for uncertainty quantification of stochastic processes. Adv Mech Eng 2019;11:1687814019857350.
[11] Sabirov DS, Shepelevich IS. Information entropy in chemistry: An overview. Entropy 2021;23:1240.
[12] Jakimowicz A. The role of entropy in the development of economics. Entropy 2020;22:452.
[13] Keum J, Kornelsen KC, Leach JM, Coulibaly P. Entropy applications to water monitoring network design: A review. Entropy 2017;19:613.
[14] Acharya UR, Fujita H, Sudarshan VK, Bhat S, Koh JE. Application of entropies for automated diagnosis of epilepsy using EEG signals: A review. Knowl-Based Syst 2015;88:85–96.
[15] Zhou R, Cai R, Tong G. Applications of entropy in finance: A review. Entropy 2013;15:4909–31.
[16] Bandt C, Pompe B. Permutation entropy: a natural complexity measure for time series. Phys Rev Lett 2002;88:174102. http://dx.doi.org/10.1103/PhysRevLett.88.174102.
[17] Zanin M, Zunino L, Rosso OA, Papo D. Permutation entropy and its main biomedical and econophysics applications: A review. Entropy 2012;14:1553–77. http://dx.doi.org/10.3390/e14081553.
[18] Henry M, Judge G. Permutation entropy and information recovery in nonlinear dynamic economic time series. Econometrics 2019;7:10–7. http://dx.doi.org/10.3390/econometrics7010010.
[19] Cuesta-Frau D, Miró-Martínez S, Oltra-Crespo P, Jordán-Núñez J, Vargas B, González P, et al. Model selection for body temperature signal classification using both amplitude and ordinality-based entropy measures. Entropy 2018;20:853–63. http://dx.doi.org/10.3390/e20110853.
[20] Xia Y, Yang L, Zunino L, Shi H, Zhuang Y, Liu C. Application of permutation entropy and permutation min-entropy in multiple emotional states analysis of RRI time series. Entropy 2018;20:148–59. http://dx.doi.org/10.3390/e20030148.
[21] Deng B, Liang L, Li S, Wang R, Yu H, Wang J, et al. Complexity extraction of electroencephalograms in alzheimer's disease with weighted-permutation entropy. Chaos 2015;25:043105. http://dx.doi.org/10.1063/1.4917013.
[22] Diaz JM, Mateos DM, Boyallian C. Complexity-entropy maps as a tool for the characterization of the clinical electrophysiological evolution of patients under pharmacological treatment with psychotropic drugs. Entropy 2017;19:540–9. http://dx.doi.org/10.3390/e19100540.
[23] Gaudêncio A, Hilal M, Cardoso J, Humeau-Heurtier A, Vaz P. Texture analysis using two-dimensional permutation entropy and amplitude-aware permutation entropy. Pattern Recognit Lett 2022;159:150–6. http://dx.doi.org/10.1016/j.patrec.2022.05.017.
[24] Chagas E, Frery AC, Rosso OA, Ramos HS. Analysis and classification of SAR textures using Information Theory. IEEE J Sel Top Appl Earth Obs Remote Sens 2021;14:663–75. http://dx.doi.org/10.1109/JSTARS.2020.3031918.
[25] Chagas ETC, Queiroz-Oliveira M, Rosso OA, Ramos HS, Freitas CGS, Frery AC. White noise test from ordinal patterns in the entropy-complexity plane. Internat Statist Rev 2022. http://dx.doi.org/10.1111/insr.12487.
[26] Amigó JM, Rosso OA. Ordinal methods: Concepts, applications, new developments, and challenges—in memory of Karsten Keller (1961–2022). Chaos 2023;33. http://dx.doi.org/10.1063/5.0167263.
[27] Zubkov AM. Limit distributions for a statistical estimate of the entropy. Theory Probab Appl 1974;18:611–8. http://dx.doi.org/10.1137/1118080.
[28] Kontoyiannis I, Skoularidou M. Estimating the directed information and testing for causality. IEEE Trans Inform Theory 2016;62:6053–67. http://dx.doi.org/10.1109/TIT.2016.2604842.
[29] Tsallis C. Possible generalization of Boltzmann-Gibbs statistics. J Stat Phys 1988;52:479–87. http://dx.doi.org/10.1007/BF01016429.
[30] Rényi A. On measures of entropy and information. In: 4th Berkeley symposium on mathematical statistics and probability, vol. 1. 1961, p. 547–61.
[31] Sánchez-Moreno P, Yáñez RJ, Dehesa JS. Discrete densities and Fisher information. In: Proceedings of the 14th international conference on difference equations and applications. Istanbul, Turkey: Uğur–Bahçeşehir University Publishing Company; 2009, p. 291–8.
[32] Salicru M, Menendez M, Morales D, Pardo L. Asymptotic distribution of (h, φ)-entropies. Comm Statist Theory Methods 1993;22:2015–31.
[33] Esteban MD, Morales D. A summary on entropy statistics. Kybernetika 1995;31:337–46.
[34] Matilla-García M. A non-parametric test for independence based on symbolic dynamics. J Econom Dynam Control 2007;31:3889–903. http://dx.doi.org/10.1016/j.jedc.2007.01.018.
[35] Matilla-García M, Marín MR. A non-parametric independence test using permutation entropy. J Econometrics 2008;144:139–55. http://dx.doi.org/10.1016/j.jeconom.2007.12.005.
[36] Matilla-García M, Rodríguez JM, Marín MR. A symbolic test for testing independence between time series. J Time Series Anal 2010;31:76–85. http://dx.doi.org/10.1111/j.1467-9892.2009.00645.x.
[37] Elsinger H. Independence tests based on symbolic dynamics. Working Papers, Oesterreichische Nationalbank (Austrian Central Bank; 2010, URL https://EconPapers.repec.org/RePEc:onb:oenbwp:165.
[38] Weiß CH, Marín MR, Keller K, Matilla-García M. Non-parametric analysis of serial dependence in time series using ordinal patterns. Comput Statist Data Anal 2022;168:107381. http://dx.doi.org/10.1016/j.csda.2021.107381.
[39] Weiß CH, Schnurr A. Generalized ordinal patterns in discrete-valued time series: nonparametric testing for serial dependence. J Nonparametr Stat 2023;1–27. http://dx.doi.org/10.1080/10485252.2023.2231565.
[40] Chagas ETC, Frery AC, Gambini J, Lucini MM, Ramos HS, Rey AA. Statistical properties of the entropy from ordinal patterns. Chaos 2022;32:113118. http://dx.doi.org/10.1063/5.0118706.
[41] Rey AA, Frery AC, Lucini M, Gambini J, Chagas ETC, Ramos HS. Asymptotic distribution of certain types of entropy under the multinomial law. Entropy 2023a;25. http://dx.doi.org/10.3390/e25050734.
[42] Rey AA, Frery AC, Gambini J, Lucini MM. The asymptotic distribution of the permutation entropy. Chaos Interdiscipl J Non Linear Sci 2023b;3:113108–33. http://dx.doi.org/10.1063/5.0171508.
[43] Spichak D, Aragoneses A. Exploiting the impact of ordering patterns in the Fisher-Shannon complexity plane. Chaos Solitons Fractals 2022;154:111620. http://dx.doi.org/10.1016/j.chaos.2021.111620.
[44] Havrda J, Charvát F. Quantification method of classification processes: concept of structural α-entropy. Kybernetika 1967;3:30–5.
[45] Frieden BR. Science from Fisher information. Cambridge: Cambridge University Press; 2004.
[46] Yamashita Rios de Sousa AM, Hlinka J. Assessing serial dependence in ordinal patterns processes using chi-squared tests with application to EEG data analysis. Chaos 2022;32:073126. http://dx.doi.org/10.1063/5.0096954.
[47] Lehmann EL, Casella G. Theory of point estimation. Springer Science & Business Media; 2006.
[48] Härdle WK, Simar L. Applied multivariate statistical analysis. 4th ed.. Heidelberg: Springer Berlin; 2015.
[49] Amigó JM, Zambrano S, Sanjuán MA. True and false forbidden patterns in deterministic and random dynamics. Europhys Lett 2007;79:50001.
[50] Freedman D, Diaconis P. On the histogram as a density estimator: L 2 theory. Z Wahrscheinlichkeitstheor Verwandte Geb 1981;57:453–76.
[51] Ricci L. Asymptotic distribution of sample shannon entropy in the case of an underlying finite, regular Markov chain. Phys Rev E 2021;103:022215.
[52] Basharin G. On a statistical estimate for the entropy of a sequence of independent random variables. Theory Probab Appl 1959;4:333–6. http://dx.doi.org/10.1137/1104033.
[53] Leslie JR, Stephens MA, Fotopoulos S. Asymptotic distribution of the shapiro-wilk W for testing for normality. Ann Statist 1986;14:1497–506. http://dx.doi.org/10.1214/aos/1176350172.
[54] Rosso OA, Larrondo HA, Martín MT, Plastino A, Fuentes MA. Distinguishing noise from chaos. Phys Rev Lett 2007;99:154102. http://dx.doi.org/10.1103/PhysRevLett.99.154102, URL http://link.aps.org/doi/10.1103/PhysRevLett.99.154102.
[55] Pandya M. Cryptocurrency prices data. 2022, http://dx.doi.org/10.34740/KAGGLE/DSV/4396186, URL https://www.kaggle.com/dsv/4396186.