Statistical Tests of Symbolic Dynamics [†]

Fernando López [1,‡], Mariano Matilla-García [2,*,‡], Jesús Mur [3,‡] and Manuel Ruiz Marín [1,‡]

1   Departamento de Métodos Cuantitativos, Universidad Politécnica de Cartagena, 30202 Cartagena, Spain;
    fernando.lopez@upct.es (F.L.); manuel.ruiz@upct.es (M.R.M.)
2   Departamento de Economía Aplicada y Estadística, Universidad Nacional de Educación a Distancia (UNED),
    28040 Madrid, Spain
3   Departamento de Análisis Económico, Universidad de Zaragoza, 50018 Zaragoza, Spain; jmur@unizar.es
*   Correspondence: mmatilla@cee.uned.es
†   This paper is an extended and revised version of our Conference manuscript presented (but not published) in
    a session at the Econometric Society World Congress 2010 under the title Symbolic Dynamics to Test Basic
    Null Hypotheses.
‡   These authors contributed equally to this work.

Abstract: A novel general method for constructing nonparametric hypotheses tests based on the field of symbolic analysis is introduced in this paper. Several existing tests based on symbolic entropy that have been used for testing central hypotheses in several branches of science (particularly in economics and statistics) are particular cases of this general approach. This family of symbolic tests uses few assumptions, which increases the general applicability of any symbolic-based test. Additionally, as a theoretical application of this method, we construct and put forward four new statistics to test for the null hypothesis of spatiotemporal independence. There are very few tests in the specialized literature in this regard. The new tests were evaluated with the mean of several Monte Carlo experiments. The results highlight the outstanding performance of the proposed test.

1. Introduction

The construction and design of powerful statistical tests are crucial elements for both theoretical and applied scientists. The utility of a test generally depends on its degree of applicability, which is usually related to the assumptions contained in the design of the test, and the restrictions of the scientific field in which the test will be used. Nowadays, the utility of statistical tests also depends on efficiency: reducing the need for computational resources and speed, which are vital for real-time monitoring and control applications. Taking applicability and efficiency into account, in this paper we propose a new general, flexible statistical methodology to design and test central hypotheses, and we establish an asymptotic distribution theory for a wide range of tests by using the new proposed approach.

The new framework is based on symbolic analysis, which is a field of increasing interest for several scientific disciplines (see [1]) Symbolic analysis studies dynamical systems on the basis of the sequences of symbols which are obtained for a suitable (and generally selected by the user) partition of the state space. In other words, the idea behind the symbolic approach is to split the phase space into a finite number of regions, and then each region is labeled with a symbol. From this point of view, the symbolic approach is a coarse-grained description of dynamics. As coarse-grained methods, which are usually used to provide some description of the data generating process, symbolic analysis focuses on some essential features of the generating dynamics which are frequently of interest to the researcher, for example, (in)dependence, cycles and nonlinear structure. In general terms, it can be said that symbolic analysis allows for designing tests that only focus on the relevant information required for the problem at hand.

This approach is not new in science. In the particular case of time series analysis, the symbolic approach implies transforming raw time series into a sequence of symbols.

Although seeminglycounter-intuitive, symbolic analysis is rooted in information theory and also in dynamics theory. For example, properties of symbols or codes are central to the theory of communication [2]. Not in vein, there is a well-established mathematical discipline, namely, symbolic dynamics, that studies the behavior of dynamical systems. The name of "symbolic dynamics" was firstly coined by [3], although the discipline started in 1898 with the pioneering work by Hadamard, who developed a symbolic description of sequences of geodesic flows. Interestingly, ref. [4] highlighted the power of the symbolic approach by showing that a complete description of the behavior of a dynamical system can be captured in terms of symbols. Notice that this property is crucial for the understanding of this paper as long as important characteristics of a random variable can also be captured by studying the symbols derived from it.

The symbolic approach has been useful in many areas of scientific research. In the experimentalist realm, relevant contributions have been made in several fields: astrophysics; biology and medicine; chemistry, mechanical systems and fluid flow; artificial intelligence, control and communication; and data mining, classification and rule discovery ([5–7], for an overview). In the non-experimentalist realm, symbolic analysis has been interestingly used. In economics and finance, data are transformed and analyzed in terms of particular symbols [8]. Two examples are recession indicators utilized to study and to determine the business cycle, and the indicators used to characterize the stock market bull and bear market periods. In geography, works like that of [9] show how qualitative variables (symbolic analysis) can be used to map descriptions. In spatial econometrics, economic spatial dependence has recently been studied by transforming data into symbols [10,11]. Other interesting applications are [12,13].

Despite all these interesting applications and the scientifically founded roots of the symbolic approach, there is no systematic body of statistical tools for conducting inference based on symbolic sequences. There are some notable exceptions: [14–21]. A common factor to all of these statistical approaches is that they are centered on ordinal patterns, which is one type of symbol. In this paper we present a novel, systematic and general framework for any potential symbol in order to test for wide range of potential null hypotheses that include, as particular cases, most of the previously indicated multidisciplinary situations, namely, ordinal patterns. We also provide a general asymptotic distribution theory for symbolic analysis. Particularly, this paper shows how, by means of symbols, it is possible to design nonparametric tests for a wide class of null hypotheses with special attention to limitations (restrictions) that typically appear in economics and finance. Therefore, this paper aims also to provide the theoretical basis for hypothesis testing by means of symbols.

An appealing advantage to symbolic analysis is that it requires very few assumptions about the data generating process in order to conduct statistical inference. This advantage is promising as the tools based on this method will share the model-free property, which avoids making unnecessary assumptions and provides more general results. Most of the econometric and statistical tests typically used in some of the mentioned disciplines cannot deal with potential nonlinear forms of dependence. By construction, nonlinear structures are not a limitation for symbolic analysis.

The capability of this approach is clearly illustrated by the scope of what we label "the symbolic main theorem" (SMT). Given a null hypothesis H_0, for example, the null of serial independence, the SMT will give us four nonparametric asymptotic tests for that null, which are distribution free. The transformation of data into symbols is done by means of a symbolization map. Some of its properties are also studied in this paper. These symbolic-based tests have to deal with ordinary statistical problems that usually appear in economics and finance, such as data scarcity and suboptimal empirical power of the test. Given the flexibility of the symbols, we provide theoretical results and strategies to overcome such difficulties.

A clear example of the power of the new tool is illustrated by the spatio-temporal data modeling issues occupying a prominent role in spatial econometrics, geography and regional science, about which we can find a vast amount of literature ([3], and references

therein). We constructed several symbolic-based tests by using the SMT. These tests also constitute an added-value of the paper, because there are currently very few available tests designed to deal with spatiotemporal dependence. The problem becomes more difficult if potential nonlinear dependence is considered. A notable exception is [9] who has treated nonlinearity in a spatial framework.

Finally, the results of this paper might be of interest to fields of research where information theory plays a relevant role. Particularly, nonparametric entropy measures and tests for serial dependence have drawn the attention of econometricians (see [22] and references therein). The clearest link between our results and information theory is through the concept of symbolic entropy. In the context of time series analysis, permutation entropy, which is a type of symbolic entropy, uses the probabilities of length-m ordinal patterns in the definition of Shannon entropy. An ordinal pattern is a particular type of symbolization map. Given the characteristics of this map, the SMT allows us to obtain an asymptotic distribution theory for a permutation entropy-based test. Providing the statistical foundation for permutation entropy is specially relevant because: (a) there are very few asymptotic distribution theories available for entropy, in general; and (b) permutation entropy is currently used in computer science due to its relation to "incompressibility", and is also useful in the study of dynamical systems because of its connection to complexity.

From another point of view, some well-established nonparametric tests can be understood as particular types of symbolic analysis. For example, the nonparametric runs test for randomness by Wald-Wolfowitz (see [23]); joint-counting procedures for spatial association [24]; and in general, categorical data techniques [25] are simple examples that use the very general procedures of translating information into symbols. In this regard, symbolic analysis can be understood as a method related to this literature.

The paper is organized as follows: In Section 2, we provide the main notation and relevant concepts that will be used in the paper. Among them we highlight: symbolization maps, standard or non-standard maps and decomposable maps. Due to the generality of the method, we require the potential tests to be adaptable to different contexts that are to be able to deal with a wide range of null hypotheses. To this end we introduce the notion of perfect and non-perfect set on subindexes in Section 3. This allows us to give general theoretical results to tackle practical situations that might otherwise be intractable because of the problem and/or of the type of hypothesis. Therefore we distinguish between two main classes of theoretical situations that lead us to different statistical solutions. In Section 4 we show how to construct symbolic-based tests via likelihood ratio statistics and via asymptotic normality. Section 5 considers the theoretical case that the null hypothesis cannot be treated under perfect situations, and hence other results are applicable. Section 6 puts forward the main theorem of this paper. Under the general conditions of this theorem, we introduce four tests for serial independence, four tests for spatial independence and four new tests for spatiotemporal independence, in Section 7. These tests are based on different symbolization maps, according to those given in Sections 2 and 3. Finally, in Section 8, we outline a Monte Carlo simulation experiment to show the capabilities of the spatiotemporal test for independence under linear and nonlinear settings. The paper ends with some conclusions.

2. Notation and Definitions

As indicated in the previous section, we give some definitions and introduce the basic notation that will be used throughout the rest of the paper.

Let $\{X_i\}_{i \in I}$ be a stationary real-valued process, where I is a set of indexes.

Let $\Gamma = \{\eta_1, \eta_2, \ldots, \eta_n\}$ be a set of $n > 1$ elements that we label as symbols. Now assume that there exists a map

$$f : \{X_i\}_{i \in I'} \to \Gamma$$

for some subset of indexes $I' \subseteq I$. We will say that $i \in I$ is of η-type if and only if $f(X_i) = \eta$. We will call the map f a symbolization map for $\{X_i\}_{i \in I}$.

Notice that it is possible to expand the definition of a symbolization map to the $k-$dimensional case by introducing the concept of decomposable maps: if $f_j : \{X_{ij}\}_{i \in I'} \to \Gamma_j, j = 1, 2, \ldots, k$, are k symbolization maps, then the product

$$F = \prod_{j=1}^{k} f_j : \prod_{j=1}^{k} \{X_{ij}\}_{i \in I'} \to \prod_{j=1}^{k} \Gamma_j$$

is a symbolization map for the k-dimensional variable $\{(X_{i1}, X_{i2}, \ldots, X_{ik})\}_{i \in I'}$. We will call F a k-decomposable symbolization map.

Given a symbolization map $f : \{X_i\}_{i \in I'} \to \Gamma$ and a symbol $\eta \in \Gamma$, we denote by p_η the probability of occurrence of symbol η. Symbolization maps can be classified according to their behavior under the null hypothesis. If the symbolization map f is such that under a given null hypothesis (H$_0$) all the symbols have the same probability to occur, we will say that f is a standard symbolization map. On the contrary, we will refer to f as a non-standard symbolization map.

The symbolic entropy of a process $\{X_i\}_{i \in I'}$ is defined as the Shannon's entropy of the n distinct symbols as follows:

$$h(\Gamma) = - \sum_{\eta \in \Gamma} p_\eta \ln(p_\eta), \tag{1}$$

with the convention $0 \times \ln 0 = \lim_{x \to 0^+} x \ln x = 0$.

Symbolic entropy, $h(\Gamma)$, can be understood as the information in terms of symbols $\eta \in \Gamma$ of the process $\{X_i\}_{i \in I'}$. Notice that $0 \leq h(\Gamma) \leq \ln(n)$. Notice also that the lower bound is attained when only one symbol occurs, and the upper bound when all n possible symbols appear with the same probability.

Consider the following index $i \in I'$ that we define an indicator random variable $Z_{\eta i}$ as follows:

$$Z_{\eta i} = \begin{cases} 1 & \text{if } f(X_i) = \eta \\ 0 & \text{otherwise,} \end{cases}$$

that is, we have that $Z_{\eta i} = 1$ if and only if i is of η-type, and $Z_{\eta i} = 0$ otherwise.

Then $Z_{\eta i}$ is a Bernoulli variable with probability of "success" p_η, where "success" means that i is of η-type. It is straightforward to see that

$$\sum_{\eta \in \Gamma} p_\eta = 1$$

Our interest is in knowing how many is are of η-type for all symbol $\eta \in \Gamma$. In order to answer the question, we construct the following counting variable:

$$Y_\eta = \sum_{i \in I'} Z_{\eta i}$$

The variable Y_η can take the values $\{0, 1, 2, \ldots, R\}$, where $|I'| = R$.

To complete with notation, we will denote by

$$n_\eta = \left| \{i \in I' | f(X_i) = \eta\} \right|,$$

the cardinality of the subset of symbolized indexes I' formed by all the elements of η-type.

Then, under the conditions above, one could easily compute the relative frequency of a symbol $\eta \in \Gamma$ by:

$$\widehat{p}_\eta := \frac{\left| \{i \in I' \mid i \text{ is of } \eta - \text{type}\} \right|}{R}$$

which is the maximum likelihood estimator of p_η.

3. On the Independence of Variables $\{Z_{\eta\,i}\}_{i \in I'}$

If the subset of symbolized indexes I' is chosen such that the $Z_{\eta i}$ variables are independent for all $i \in I'$, then Y_η is a binomial random variable

$$Y_\eta \sim B(R, p_\eta).$$

Moreover, if the subset of indexes I' is such that the variables Z_{η_s} and Z_{σ_t} are independent for all symbol $\eta_s, \sigma_t \in \Gamma$, $s \neq t$, where $s, t \in \{1, 2, \ldots, n\}$, then the joint probability density function of the n variables $(Y_{\eta_1}, Y_{\eta_2}, \ldots, Y_{\eta_n})$ is:

$$P(Y_{\eta_1} = a_1, Y_{\eta_2} = a_2, \ldots, Y_{\eta_n} = a_n) = \frac{(a_1 + a_2 + \cdots + a_n)!}{a_1! a_2! \cdot \cdots \cdot a_n!} p_{\eta_1}^{a_1} p_{\eta_2}^{a_2} \cdots p_{\eta_n}^{a_n}$$

where $a_1 + a_2 + \cdots + a_n = R$. Consequently the joint distribution of the n variables $(Y_{\eta_1}, Y_{\eta_2}, \ldots, Y_{\eta_n})$ is a multinomial distribution. In this case we will call the set of symbolized indexes I' a perfect subset.

It is possible to develop theoretical results that contemplate situations for which the researcher will benefit of constructing symbolization maps for which the set I' is not perfect. Given our previous definition of perfect set, the indexes subset I' can be nonperfect in the following cases:

- Case (a): $Z_{\eta i}$ and $Z_{\eta j}$ are not independent for all $i \neq j$ for some $\eta \in \Gamma$, and hence Y_η is not a binomial random variable.
- Case (b): Z_{η_s} and Z_{σ_t} are not independent for all $s \neq t$ for some $\sigma_t, \eta_s \in \Gamma$, and hence $(Y_{\eta_1}, Y_{\eta_2}, \ldots, Y_{\eta_n})$ is not a multinomial distribution.

Theoretical results for perfect and non-perfect subsets are the topics of the following two sections.

4. Constructing Tests with Perfect Subsets of I

In this section we establish a general framework for testing for a null hypothesis H_0. This is done by focusing on the symbols' distribution when the subset I' of symbolized symbols is perfect. Under this condition, we now show how to construct tests of hypothesis via likelihood ratio statistics and via asymptotic normality.

Procedure

In general, our procedure consists of proceeding systematically as follows:

1. (step 1) Fix the null hypothesis H_0 to be tested.
2. (step 2) Define the set of symbols Γ and the symbolization map f.
3. (step 3) Compute the distribution of the symbols under H_0, namely, $p_\eta^{(0)}$ for all $\eta \in \Gamma$.
4. (step 4) Finally design and compute a desired test statistic.

Steps 1 to 2 will be developed later in the paper. To accomplish the aim with steps 3 and 4, we first show how to construct a likelihood ratio test.

Recall that when the set is perfect, $(Y_{\eta_1}, \ldots, Y_{\eta_n})$ follows a multinomial distribution and its likelihood function is:

$$L(p_{\eta_1}, p_{\eta_2}, \ldots, p_{\eta_n}) = \frac{R!}{n_{\eta_1}! n_{\eta_2}! \cdot \cdots \cdot n_{\eta_n}!} p_{\eta_1}^{n_{\eta_1}} p_{\eta_2}^{n_{\eta_2}} \cdots p_{\eta_n}^{n_{\eta_n}}$$

and the likelihood ratio test statistic is

$$\lambda(Y) = \frac{\frac{R!}{n_{\eta_1}! n_{\eta_2}! \cdot \cdots \cdot n_{\eta_n}!} p_{\eta_1}^{(0)n_{\eta_1}} p_{\eta_2}^{(0)n_{\eta_2}} \cdots p_{\eta_n}^{(0)n_{\eta_n}}}{\frac{R!}{n_{\eta_1}! n_{\eta_2}! \cdot \cdots \cdot n_{\eta_n}!} \widehat{p}_{\eta_1}^{n_{\eta_1}} \widehat{p}_{\eta_2}^{n_{\eta_2}} \cdots \widehat{p}_{\eta_n}^{n_{\eta_n}}}, \tag{2}$$

where $\widehat{p}_{\eta_i}$ is the maximum likelihood estimator of p_{η_i} for all $i = 1, 2, \ldots, n$. In this case, as shown in the Appendix A, maximum likelihood estimators are $\widehat{p}_{\eta_i} = \frac{n_{\eta_i}}{R}$.

Moreover, it is possible to show (see Appendix A.2) that

$$-2[RLn(R) + \sum_{i=1}^{n} n_{\eta_i} Ln\left(\frac{p_{\eta_i}}{n_{\eta_i}^0}\right)] \sim \chi_k^2,$$

where the k degrees of freedom will depend on the set of symbols.

It is also proved that, if standard symbolization maps are considered, then

$$2R[Ln(n) - h(\Gamma)] \sim \chi_k^2, \tag{3}$$

which is an affine transformation of the symbolic entropy (1).

As we have emphasized in the Introduction that nonparametric results of entropy based measures are relevant for econometrics and for other fields of research. Given that the distribution of the affine transformation in (3) holds for standard symbolization maps under broad conditions, this result will be of general applicability. In particular, as we will show in other sections, permutation entropy, as introduced in [26], is a particular type of symbolic entropy that has drew the attention of several scholars for theoretical interests basically because stationary and ergodic processes coincide with Shannon entropy, and for an applied interest in nonlinear and complex systems or processes (see [1]). In this regard, (2) and (3) can be viewed as an initial step towards statistical inference of ordinal pattern distributions.

An alternative to likelihood ratio tests for symbolic maps can be considered by modifying step 4. Given that under a perfect subset I' the indicator variables Z_{η_i} are independent, the random variable $\frac{Y_\eta - Rp_\eta}{\sqrt{Rp_\eta(1-p_\eta)}}$ has a limiting normal distribution with zero mean and unit variance for all symbol $\eta \in \Gamma$. Moreover

$$\left(\frac{Y_{\eta_1} - Rp_{\eta_1}}{\sqrt{Rp_{\eta_1}(1 - p_{\eta_1})}}, \frac{Y_{\eta_2} - Rp_{\eta_2}}{\sqrt{Rp_{\eta_2}(1 - p_{\eta_2})}}, \cdots, \frac{Y_{\eta_n} - Rp_{\eta_n}}{\sqrt{Rp_{\eta_n}(1 - p_{\eta_n})}}\right)$$

asymptotically distributes as a multivariate normal distribution $MN(0, I)$. In this case, the asymptotic distribution holds for standard or nonstandard symbolization maps at the cost of estimating p_{η_i} that can be consistently estimated by $\widehat{p}_{\eta_i}$.

5. Constructing Tests with Non Perfect Subsets of I

In this section we establish the equivalent counterpart of the general framework (above presented) when the subset I' of symbolized symbols is non-perfect. Accordingly, we now show to what extent and under which situations the previous likelihood and asymptotically normal tests (elaborated under perfect situations) can be adapted to deal with them.

Non perfect sets of indexes I might be very useful for test design, especially for situations or scientific domains characterized by relatively scarce sample size as compared with the number of symbols. In macroeconomics, although not necessarily in finance, data scarcity can be the usual restriction. Ideally, one can be able to design perfect sets of indexes to carry out symbolic-based hypothesis testing. This section then tries to provide symbolic-based methods for constructing hypothesis tests for situations for which an ideal design is not possible because of the nature of the problem, because of the computational capabilities or because of any other potential reason.

5.1. Binomial Approximation

Let us consider that variables $Z_{\eta i}$ and $Z_{\eta j}$ are not independent for all $i \neq j$ for some $\eta \in \Gamma$. In this case, Y_η is not a binomial random variable. The interesting question is how

far is Y_η from $B(R, p_\eta)$. We are interested in studying under what assumptions the variable Y_η can be approximated to a binomial random variable:

$$Y_\eta \approx B(R, p_\eta).$$

In fact, it is possible to compute a bound for this binomial approximation.

Denote by $\mathcal{L}(Y_\eta)$ the distribution of Y_η, and we are interested in the bound of the binomial approximation of the distribution of Y_η measured in terms of total variation distance. The total variation distance d_{TV} between two probability measures P and Q is defined by

$$d_{TV}(P, Q) = \sup_A |P(A) - Q(A)|$$

where the supremum is taken over all measurable sets of the real line.

Following Theorem 1.1 in [27] and after a few calculations, a bound can be given as follows:

For each $i, j \in I'$ let $Z_{\eta j}, Z_{\eta i}$ and $J_{\eta ij}$ be defined in the same probability space where $J_{\eta ji} = (Z_{\eta j}|Z_{\eta i} = 1)$. Let

$$W_i = \sum_{j \neq i} Z_{\eta j} \qquad V_i = \sum_{j \neq i} J_{\eta ji}$$

and

$$C_{R,p} = \frac{1 - p^{R+1} - q^{R+1}}{(R+1)pq}.$$

W_i counts the number of indexes that are of η-type and $p = p_\eta$ and $q = 1 - p_\eta$. On the other hand $V_i = \sum_{j \neq i} (Z_{\eta j}|Z_{\eta i} = 1)$ counts the number of indexes that are of η-type conditioned to location i is of η-type.

Then

$$d_{TV}(\mathcal{L}(Y_\eta), B(R, p)) \leq C_{R,p} p \sum_{i=1}^{R} E(|W_i - V_i|).$$

Therefore, in order to get a bound for the binomial approximation, we have to get bound

$$\sum_{i=1}^{R} E(|W_i - V_i|).$$

On the other hand, we have that $E(|W_i - V_i|) \leq \sum_{j \neq i} E(|Z_{\eta j} - J_{\eta ji}|)$ and $pE(J_{\eta ji}) = E(Z_{\eta i} Z_{\eta j})$. Now denote by B_i^η a subset of indexes such that $Z_{\eta i}$ is independent of $\{Z_{\eta j}| j \notin B_i\}$. Therefore, we obtain that

$$
\begin{aligned}
d_{TV}(\mathcal{L}(Y_\eta), B(R, p)) &\leq C_{R,p} p \sum_{i=1}^{R} E(|W_i - V_i|) \\
&\leq C_{R,p} p \sum_{i=1}^{R} \sum_{j \in B_i \setminus \{i\}} E(|Z_{\eta j} - J_{\eta ji}|) \\
&\leq C_{R,p} p \sum_{i=1}^{R} \sum_{j \in B_i \setminus \{i\}} (p + E(J_{\eta ji})) \\
&\leq C_{R,p} \left(\sum_{i=1}^{R} p^2 (|B_i| - 1) + \sum_{i=1}^{R} \sum_{j \in B_i \setminus \{i\}} E(Z_{\eta i} Z_{\eta j}) \right)
\end{aligned}
$$

Therefore, we have shown that the sum of dependent indicators can be approximated to a binomial random variable when the following two conditions are satisfied: (1) dependencies among the indicators are weak and (2) the probabilities of the indicators occurring

under the null hypothesis are small. Notice that point (1) can be guaranteed by selecting the subset of symbolized indexes I' such that $|B_i|$ is small enough.

5.2. Normal Approximation

Additionally, when the indicator variables $Z_{\eta i}$ are not independent for all $i \in I'$, the following central limit theorem for dependent indicators ensures the convergence to a normal distribution:

Theorem 1 (Theorem 7.7.5 (Anderson, 1971)). *Let $Z_1, Z_2, \ldots$ be a stationary stochastic process such that for every integer n and integers $t_1, t_2, \ldots, t_n$ with $t_1 < \cdots < t_n$, $Z_{t_1}, \ldots, Z_{t_n}$ are distributed independently of $Z_1, \ldots, Z_{t_1-m-1}$ and $Z_{t_n+m+1}, \ldots$. If $E(Z_t) = 0$ and $E(Z_t^2) < \infty$, then*

$$\frac{\sum\limits_{t=1}^{T} Z_t}{\sqrt{T}}$$

has a limiting normal distribution with mean 0 and variance

$$E(Z_1^2) + 2E(Z_1 Z_2) + \cdots + 2E(Z_1 Z_{m+1}).$$

Theorem 1 states that, if the dependencies are weak (for instance, $|B_i|$ is small enough for all $i \in I'$) we get that $\frac{Y_\eta - R p_\eta}{\sqrt{Var(Y_\eta)}} \xrightarrow{d} N(0,1)$ as $R \to \infty$. The variance of the variable Y_η can be computed as follows:

$$
\begin{aligned}
Var(Y_\eta) &= E(Y_\eta^2) - E(Y_\eta)^2 = \sum_{i=1}^{R} E(Z_{\eta i}^2) + \sum_{i=1}^{R}\sum_{j \neq i} E(Z_{\eta i} Z_{\eta j}) - R^2 p_\eta^2 \\
&= R p_\eta - R^2 p_\eta^2 + \sum_{i=1}^{R} (R - |B_i|) p_\eta^2 + \sum_{i=1}^{R} \sum_{j \in B_i \backslash \{i\}} E(Z_{\eta i} Z_{\eta j}) \\
&= R p_\eta - R^2 p_\eta^2 + R^2 p_\eta^2 - p_\eta^2 \sum_{i=1}^{R} |B_i| + \sum_{i=1}^{R} \sum_{j \in B_i \backslash \{i\}} E(Z_{\eta i} Z_{\eta j}) \\
&= R p_\eta - p_\eta^2 \sum_{i=1}^{R} |B_i| + \sum_{i=1}^{R} \sum_{j \in B_i \backslash \{i\}} E(Z_{\eta i} Z_{\eta j}) \qquad (4)
\end{aligned}
$$

We now consider the case where Z_{η_s} and Z_{σ_t} are not independent for all $s \neq t$ for some $\sigma_t, \eta_s \in \Gamma$, and hence $(Y_{\eta_1}, Y_{\eta_2}, \ldots, Y_{\eta_n})$ is not a multinomial distribution (i.e., previous case (b)). To this end, denote by $B_i^{(\eta,\sigma)}$ the subset of indexes satisfying that $Z_{\eta i}$ is independent of $\{Z_{\sigma j} | j \notin B_i^{(\eta,\sigma)}\}$ and $X_\eta = \frac{Y_\eta - R p_\eta}{\sqrt{Var(Y_\eta)}}$. It is possible to show that $(X_{\eta_1}, X_{\eta_2}, \ldots, X_{\eta_n})$ is a multivariate normal distribution. In fact, it is equivalent to proving that any linear combination of the X_ηs is normal. In order to do so, notice that each variable X_η is a sum of indicator variables, and therefore, any linear combination of these variables is again a sum of indicator variables. Consider an arbitrary linear combination as follows:

$$M = \alpha_1 X_{\eta_1} + \alpha_2 X_{\eta_2} + \cdots + \alpha_n X_{\eta_n}$$

Notice that the variable αX_η can be written as $\alpha X_\eta = \sum\limits_{i=1}^{R} I_i^\eta$ where I_i^η is an indicator variable for all $i \in I'$. Therefore, it follows that the variable M is a sum of dependent indicator variables. Again, by Theorem 1 we get that M follows a normal distribution whenever the dependencies among the indicator variables are weak (for instance, when the cardinality of the set $B_i \cup B_i^{(\eta_s, \eta_t)}$ is small enough for all $i \in I'$ and all symbol $\eta, \eta_s, \eta_t \in \Gamma$).

Then we get that $(X_{\eta_1}, X_{\eta_2}, \ldots, X_{\eta_n})$ has a limiting multivariate normal distribution and we can compute the variance and covariance matrix for all $\eta_s, \eta_t \in \Gamma$ as follows:

$$
\begin{aligned}
Cov(Y_{\eta_s}, Y_{\eta_t}) &= \sum_{i=1}^{R} \sum_{j=1}^{R} Cov(Z_{\eta_s i}, Z_{\eta_t j}) = \sum_{i=1}^{R} \sum_{j \in B_i^{(\eta_s, \eta_t)}} Cov(Z_{\eta_s i}, Z_{\eta_t j}) \\
&= \sum_{i=1}^{R} \sum_{j \in B_i^{(\eta_s, \eta_t)}} \{E(Z_{\eta_s i} Z_{\eta_t j}) - E(Z_{\eta_s i}) E(Z_{\eta_t j})\} \\
&= p_{\eta_s} \sum_{i=1}^{R} \sum_{j \in B_i^{(\eta_s, \eta_t)}} p(Z_{\eta_t j} = 1 | Z_{\eta_s i} = 1) - p_{\eta_s} p_{\eta_t} \sum_{i=1}^{R} |B_i^{(\eta_s, \eta_t)}|.
\end{aligned}
\tag{5}
$$

6. The Symbolic Main Theorem

Previous partial results can be collected in the following main theorem.

Theorem 2. *Let $\{X_i\}_{i \in I'}$ be a stationary real valued process. Let $\Gamma = \{\eta_1, \eta_2, \ldots, \eta_n\}$ be a finite set of symbols. Let $f : \{X_i\}_{i \in I'} \to \Gamma$ be a standard symbolization map for a null hypothesis H_0. Then under the null H_0 we have that:*

- *If the set of symbolized indexes I' is perfect:*

    1. *Then $G(\Gamma) = 2R[\ln(n) - h(\Gamma)]$ is asymptotically χ_k^2 distributed where k is the difference between the number of parameters to be estimated under the alternative hypothesis H_1 and the number of parameters to be estimated under the null H_0.*
    2. *Then $\left(\frac{Y_{\eta_1} - Rp_{\eta_1}}{\sqrt{Rp_{\eta_1}(1-p_{\eta_1})}}, \frac{Y_{\eta_2} - Rp_{\eta_2}}{\sqrt{Rp_{\eta_2}(1-p_{\eta_2})}}, \ldots, \frac{Y_{\eta_n} - Rp_{\eta_n}}{\sqrt{Rp_{\eta_n}(1-p_{\eta_n})}}\right)$ is a multivariate normal distribution $MN(0, I)$.*

- *The set of symbolized indexes I' is not perfect:*

    1. *If the sets B_i and $B_i^{(\eta_s, \eta_t)}$ have good properties in the sense that for all symbol $\eta \in \Gamma$ the variables Y_η have a good approximation to a binomial distribution and $(Y_{\eta_1}, Y_{\eta_2}, \ldots, Y_{\eta_n})$ has a good approximation to a multinomial distribution, then $G(\Gamma) = 2R[\ln(n) - h(\Gamma)]$ is asymptotically χ_k^2 distributed where k is the difference between the number of parameters to be estimated under the alternative hypothesis H_1 and the number of parameters to be estimated under the null H_0.*
    2. *If I' is such that Theorem 1 holds, then $\left(\frac{Y_{\eta_1} - Rp_{\eta_1}}{\sqrt{Var(Y_{\eta_1})}}, \frac{Y_{\eta_2} - Rp_{\eta_2}}{\sqrt{Var(Y_{\eta_2})}}, \ldots, \frac{Y_{\eta_n} - Rp_{\eta_n}}{\sqrt{Var(Y_{\eta_n})}}\right)$ is a multivariate normal distribution $MN(0, \Sigma)$ where the variance and covariance matrix can be estimated using (4) and (5).*

Notice that in the case of nonstandard symbolization maps, an analogous theorem to Theorem 2 will hold. More concretely, for point 1 of the SMT, a likelihood ratio test is also available, although not in the closed form as presented here (see for example [28]). Point 2 of the SMT hold independently of whether the symbolization map is standard.

The usefulness of the theorem will become evident in the next section, where it will be applied to specific null hypotheses. Naturally, it is possible to consider an alternatively bootstrap-based test for symbolization maps, instead of asymptotic ones, although we do not follow this way in this paper, it being a subfield for further research of interest since it partially might avoid taking care of nonperfect indexes sets.

7. Different Symbolizations for Different Nulls Related with Independence

According to the general symbolic theorem, in this section we show how it is possible to test interesting null hypotheses by using symbolic analysis. To concrete, we focus on testing for different nulls of independence as it is a well-known field of research and because recently published articles can be generally understood and extended under this new theoretical framework. Given that each null hypothesis (step 1) will require a

particular symbolization map, in this section we present different symbolization procedures (step 2) to test for serial dependence, spatial dependence, and spatiotemporal dependence, respectively. Then we present the results of step 3 and step 4 depending on the statistic technique the researcher wants to use according to Theorem 2, i.e., either likelihood ratio statistics or/and asymptotically normal statistics. Given a null hypothesis, the behavior of the tests obtained from this approach will strongly depend on the expertise of the researcher in constructing the symbolization map. We emphasize that both power analysis of the class of tests, and power competition among alternative nonparametric tests were already given in previous work [10,16]; therefore, we are not going to replicate them here.

As we have indicated, the crucial component of the symbolic procedure is to choose a symbolic mapping which ensures that the distribution of the symbols can detect deviations from the null. The null hypotheses considered in this section are related to the important topic of "statistical independence". This is a very well-studied topic in time series analysis and therefore there is a generous number of available tests. On the contrary, spatial independence is not so well-known and is non-trivial how to test for it. As we will show, it is needed to use another different symbolization map for detecting spatial patterns. Similar comments can be made for spatio-temporal independence. Needless to say, there are other hypotheses of interest in econometrics, and the researcher will have to design suitable symbolic maps for testing them. For example, in [29], the authors dealt with the opposite problem: how to test for a pure deterministic chaotic process. In these and other cases, the power of the tests will centrally depend on the ability of the research to design the symbolization map for the desired null hypothesis.

7.1. Serial Independence Tests

In the case of time series, refs. [15,16] used the following symbolization procedure to test for serial dependence: Let $\{X_t\}_{t \in I}$ be a real-valued time series (in this case the subindex t refers to time) for which we are interested in testing the null of serial independence (step 1). In order to complete step 2, we denote by $\Gamma_1 = S_m$ the symmetric group of order $m!$, that is, the group formed by all the permutations of length m (for a positive integer $m \geq 2$). Let $\pi_i = (i_1, i_2, \ldots, i_m) \in S_m$. The positive integer m is usually known as the embedding dimension.

An ordinal pattern for a symbol is defined as $\pi_i = (i_1, i_2, \ldots, i_m) \in S_m$ at a given time $t \in I$. The time series can be embedded in an m-dimensional space:

$$X_m(t) = (X_{t+1}, X_{t+2}, \ldots, X_{t+m}) \text{ for } t \in I$$

It is said that t is of π_i−type if and only if $\pi_i = (i_1, i_2, \ldots, i_m)$ is the unique symbol in the group S_m satisfying the two following conditions:

$$(a) \quad X_{t+i_1} \leq X_{t+i_2} \leq \cdots \leq X_{t+i_m}, \text{ and}$$

$$(b) \quad i_{s-1} < i_s \text{ if } X_{t+i_{s-1}} = X_{t+i_s}$$

Notice that condition (b) guaranties uniqueness of the symbol π_i. This is justified if the values of X_t have a continuous distribution so that equal values are very uncommon, with a theoretical probability of occurrence of 0.

In this case, the symbolization map is defined as $f_1 : \{X_t\}_{t \in I'} \to S_m$ given by

$$f_1(X_t) = (i_1, i_2, \ldots, i_m) \tag{6}$$

where $(i_1, i_2, \ldots, i_m) \in S_m$ is such that t is of $(i_1, i_2, \ldots, i_m)$-type. Now the design of the symbolization map (step 2) is completed.

Moreover, under the null of independence the distribution of the symbols is uniform and therefore the map f_1 is a standard symbolization map. Additionally, the set of symbolized indexes is $I' = \{1, 2, \ldots, T - m + 1\}$, which is not perfect.

Notice that in order to have a perfect set and therefore ensure the independence of the indicator variables $Z_{\pi t}$, it is enough to consider as a set of symbolized indexes

$$I' = \{0, m-1, 2(m-1), \ldots, t(m-1), \ldots\}.$$

Accordingly, using this symbolization map, the next corollary straightforwardly follows from Theorem 2:

Corollary 1. *Let $f_1 : \{X_t\}_{t \in I'} \to \Gamma_1$ be the symbolization map defined in (6) with $|I'| = R$. Denote by $h(\Gamma_1)$ the permutation entropy defined in (1). If the time series $\{X_t\}_{t \in I}$ is independent, then*

- *If the set I' is perfect, then:*

    1. *The affine transformation of the permutation entropy*

    $$G(S_m) = 2R[Ln(m!) - h(S_m)]$$

    is asymptotically $\chi^2_{m!-1}$ distributed.
    2. *Then $\left(\frac{Y_{\pi_1} - Rp_{\pi_1}}{\sqrt{Rp_{\pi_1}(1-p_{\pi_1})}}, \frac{Y_{\pi_2} - Rp_{\pi_2}}{\sqrt{Rp_{\pi_2}(1-p_{\pi_2})}}, \ldots, \frac{Y_{\pi_{m!}} - Rp_{\pi_{m!}}}{\sqrt{Rp_{\pi_{m!}}(1-p_{\pi_{m!}})}}\right)$ is a multivariate normal distribution $NM(0, I)$.*

- *If the set I' is not perfect:*

    1. *Since the sets B_i^{π}'s has cardinality of at most $2m$, we can get a good approximation to the following result via [21]. The affine transformation of the permutation entropy*

    $$G(S_m) = 2R[Ln(m!) - h(S_m)]$$

    is asymptotically $\chi^2_{m!-1}$ distributed.
    2. *Then $\left(\frac{Y_{\pi_1} - Rp_{\pi_1}}{\sqrt{Var(Y_{\pi_1})}}, \frac{Y_{\pi_2} - Rp_{\pi_2}}{\sqrt{Var(Y_{\pi_2})}}, \ldots, \frac{Y_{\pi_{m!}} - Rp_{\pi_{m!}}}{\sqrt{Var(Y_{\pi_{m!}})}}\right)$ is a multivariate normal distribution $MN(0, \Sigma)$ where the variance and covariance matrix can be estimated using (4) and (5).*

These results for permutation entropy are in relation to a relatively recent line of research based on order patterns for analyzing time series. Ordinal patterns can be, per se, used for descriptive purposes, like autocorrelation, with the added advantage that the require no assumptions such as Gaussianity or linearity. On the contrary, only mild stationary conditions can exist in the underlying process. The above corollary is a further step for the development of statistical inference for ordinal time series. Naturally, it is possible to obtain other kinds of statistical results by adding more assumptions to the generating process. In fact, notorious results can be found in [4]) if Gaussianity and ergodicity are assumed. In this regard, our asymptotic results for order patterns keep assumptions at a minimum. Additionally, by maintaining general applicability at minimum cost (in terms of assumptions) for serial independence tests, some bootstrap-based statistics for ordinal patterns have been put forward in [29].

An interesting property of the symbolization procedure presented in this section is that it can be also used for discrete distributions. To do so it necessary to consider a non-standard version of the map. Under such circumstances, the likelihood ratio (2) can be directly used once the behavior of p_{η_i} is known under the null of serial independence.

7.2. Spatial Independence Tests

In the case of spatial processes, ref. [10] gave a symbolization procedure to test for spatial independence as follows: Let $\{X_s\}_{s \in S}$ be a real-valued spatial process, where S is a set of coordinates. Given a location s_0, we will denote by (ρ_i^0, θ_i^0) the polar coordinates of location s_i taking as origin s_0.

Let $m \in \mathbb{N}$ with $m \geq 2$. Consider now that the spatial process $\{X_s\}_{s \in S}$ is embedded in a different m-dimensional space as follows:

$$X_m(s_0) = (X_{s_0}, X_{s_1}, \ldots, X_{s_{m-1}}) \text{ for } s_0 \in S$$

where $s_1, s_2, \ldots, s_{m-1}$ are the $m-1$ nearest neighbors to s_0, which are ordered from lesser to higher Euclidean distance with respect to location s_0. Notice that in the case of two or more locations being equidistant to s_0, we will choose them in an anticlockwise manner. In formal terms, $s_1, s_2, \ldots, s_{m-1}$ are the $m-1$ nearest neighbors to s_0 satisfying the following two conditions:

- (a) $\rho_1^0 \leq \rho_2^0 \leq \cdots \leq \rho_{m-1}^0$;
- (b) If $\rho_i^0 = \rho_{i+1}^0$, then $\theta_i^0 < \theta_{i+1}^0$.

Notice that conditions (a) and (b) ensure the uniqueness of $X_m(s)$ for all $s \in S$.

The proposed standard symbolization map f is defined as follows: denote by *Me* the median of the spatial process $\{X_s\}_{s \in S}$ and let

$$\delta_s = \begin{cases} 0 & \text{if } X_s \leq Me \\ 1 & \text{otherwise} \end{cases}$$

Now, define the indicator function

$$\mathcal{I}_{s_1 s_2} = \begin{cases} 0 & \text{if } \delta_{s_1} \neq \delta_{s_2} \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

Then, the standard symbolization map

$$f_2 : \{X_s\}_{s \in I'} \to \Gamma_2 \tag{8}$$

is defined as:

$$f_2(X_s) = (\mathcal{I}_{ss_1}, \mathcal{I}_{ss_2}, \ldots, \mathcal{I}_{ss_{m-1}}), \tag{9}$$

where Γ_2 stands for the set of symbols defined by f_2.

Notice that under the null of spatial independence, the distribution of the symbols is uniform and therefore the map f_2 is a standard symbolization map.

Moreover, in this case $I' = S$ is not a perfect symbolized set. To construct a perfect symbolized set I', one can proceed as follows. Take a location $s_0 \in S$ at random. Let N_s be the set of nearest neighbors to s. Now select the following element in I' by taking $s_1 \in S$ such that $N_{s_1} \cap N_{s_0} = \emptyset$. Then construct recursively the set I' by taking $s_k \in S \setminus \{s_0, s_1, \ldots, s_{k-1}\}$ satisfying $N_{s_i} \cap N_{s_j} = \emptyset$ for all $i \neq j$ with $i, j = 1, 2, \ldots, k$.

As it is evident, the method is flexible enough to allow the researcher to select his own set and map of symbols for a given null. For example, if under the previous symbolization procedure, the power (or size) of the test is not satisfactory, one can always consider other possible symbolization procedures for the same null and for the same spatial process $\{X_s\}_{s \in S}$. Let $\Gamma_3 = \{1, 2, \ldots, k\} \times \{1, 2, \ldots, k\}$. Again, let N_s be the set of nearest neighbors to s and let n_s be its cardinality. Denote by $X_s^N = \frac{1}{n_s} \sum_{s' \in N_s} X_{s'}$. Denote by q_i and q_i^N the i-th quantile of the variables X and X^N respectively, for $i \in \{1, 2, \ldots, k-1\}$. We will denote by $q_0 = \min_{s \in S} X_s$ (resp $q_0^N = \min_{s \in S} X_s^N$) and $q_{k+1} = \max_{s \in S} X_s$ (resp. $q_{k+1}^N = \max_{s \in S} X_s^N$). Then we define the symbolization map

$$f_3(X_s) = (i, j) \tag{10}$$

if and only if $X_s \in [q_{i-1}, q_i]$ and $X_s^N \in [q_{j-1}^N, q_j^N]$.

Again, under the null of independence the distribution of the symbols is uniform and therefore the map $f_3 : \{X_s\}_{s \in I'} \to \{1, 2, \ldots, k\} \times \{1, 2, \ldots, k\}$ is a standard symbolization map.

Again, the same set of recursively constructed symbolized indexes S' ensures the independence of the indicator variables $Z_{\eta s}$. Accordingly, using this symbolization map, the next corollary straightforwardly follows from Theorem 2:

Corollary 2. *Let $f_i : \{X_s\}_{s \in S'} \to \Gamma_i$, $i = 2, 3$ be the symbolization maps defined in (8) and (10) with $|S'| = R$. Denote by $h(\Gamma_i)$ the symbolic entropy defined in (1). If the spatial process $\{X_s\}_{s \in S}$ is independent, it follows that:*

- *If the set S' is perfect then:*

  1. *The affine transformation of the symbolic entropy*

     $$SG(\Gamma_i) = 2R[Ln(n_i) - h(\Gamma_i)]$$

     is asymptotically $\chi^2_{d_i}$ distributed where $d_2 = 2^{m-1}$, $d_3 = (k-1)^2 + 2$, $n_2 = 2^{m-1}$ and $n_3 = k^2$.
  2. *Then $\left(\frac{Y_{\sigma_1} - Rp_{\sigma_1}}{\sqrt{Rp_{\sigma_1}(1-p_{\sigma_1})}}, \frac{Y_{\sigma_2} - Rp_{\sigma_2}}{\sqrt{Rp_{\sigma_2}(1-p_{\sigma_2})}}, \ldots, \frac{Y_{\sigma_{ni}} - Rp_{\sigma_{ni}}}{\sqrt{Rp_{\sigma_{ni}}(1-p_{ni})}} \right)$ is a multivariate normal distribution $MN(0, I)$.*

- *If the set S' is not perfect, then:*

  1. *Since the sets B_i^σ have small cardinality, we can get a good approximation to the following result of [17]. The affine transformation of the symbolic entropy*

     $$SG(\Gamma_i) = 2R[Ln(n_i) - h(\Gamma_i)]$$

     is asymptotically $\chi^2_{d_i}$ distributed where $d_2 = 2^{m-1}$, $d_3 = (k-1)^2 + 2$, $n_2 = 2^{m-1}$ and $n_3 = k^2$.
  2. *Then $\left(\frac{Y_{\sigma_1} - Rp_{\sigma_1}}{\sqrt{Var(Y_{\sigma_1})}}, \frac{Y_{\sigma_2} - Rp_{\sigma_2}}{\sqrt{Var(Y_{\sigma_2})}}, \ldots, \frac{Y_{\sigma_{ni}} - Rp_{\sigma_{ni}}}{\sqrt{Var(Y_{\sigma_{ni}})}} \right)$ is a multivariate normal distribution $MN(0, \Sigma)$ where the variance and covariance matrix can be estimated using (4) and (5).*

In Section 2 we indicate that there is a class of symbolization maps that are non-standard. Consider a situation in which a reduction in the number of possible symbols under study will benefit the behavior and properties of the test. In this, and other potential situations, non-standard maps might be useful. As an example, we now construct a non-standard symbolization map to test for independence in the spatial context. The following symbolization is an example of the most general procedure that we give in Appendix A.3.

Consider again the set Γ_2 of symbols defined in (9) for a fixed embedding dimension m. Now we will denote by $\bar{a}$ the rest of the division of a over $m - 1$.

Now define the following equivalence relation $\sim$:

$$(\mathcal{I}_{ss_1}, \mathcal{I}_{ss_2}, \ldots, \mathcal{I}_{ss_{m-1}}) \sim (\mathcal{I}_{s's'_1}, \mathcal{I}_{s's'_2}, \ldots, \mathcal{I}_{s's'_{m-1}})$$

if and only if there exists an integer k such that $\mathcal{I}_{s's'_i} = \mathcal{I}_{ss_{\overline{i+k}}}$ for all $i \in \{1, 2, \ldots, m-1\}$.

Now we consider as a set of symbols $\Gamma_4 = \Gamma_2 / \sim$ the set of classes in Γ_2 modulo, the equivalence relation $\sim$.

Notice that, in general, in this case not all the symbols in Γ_4 have the same probability of occurring, and therefore the symbolization map $f_4 : \{X_s\}_{s \in S'} \to \widetilde{\Gamma_4}$ is non-standard.

7.3. Spatiotemporal Independence Tests

The issues related to spatiotemporal data modeling occupy a prominent role in current econometrics, where we can find recent literature devoted to this topic (see [9,30]). Spatiotemporal dependence introduces considerable difficulties with respect to modeling, computation and statistical theory. If independence can be taken for granted, and likewise the common assumption of cross-sectional independence, then computations and the application of inference rules simplifies significantly. It seems reasonable therefore to test first for

spatiotemporal independence, and if the evidence for independence is strong, then proceed with the well-known methods. Unfortunately, tests for spatiotemporal independence are scarce. The aim of this section is twofold: to contribute to this rather scarce literature, and to highlight the usefulness of the novel general method presented in this paper. To this end we consider the relevant null of spatiotemporal dependence. Of particular interest for our tests is that dependence is not taken as a synonymous with correlation, and therefore nonlinearities are not restrictions for our test.

Consider the process $\{X_{ts}\}_{t\in I, s\in S}$. As in the previous cases, one can define several standard and non-standard symbolization maps. For simplicity, we adapt the previous symbolizations to the spatiotemporal case as follows:

For a fixed location $s_0 \in S'$ define $\{X_{t(s_0)}\}$ as the time series $\{X_{1s_0}, X_{2s_0}, \ldots, X_{p(s_0)}, \ldots\}$. Similarly for a fixed period $t_0 \in I'$ we define $\{X_{(t_0)s}\}$ as the spatial process

$$\{X_{t_0 s_1}, X_{t_0 s_2}, \ldots, X_{t_0 s_p}, \ldots\}$$

Let $m_t, m_s \in \mathbb{N}$ with $m_t, m_s \geq 2$ be the time and space embedding dimensions respectively. Then under this setting we define the following decomposable symbolization maps $F_{1i} : \{X_{ts}\}_{t\in I', s\in S'} \to \S_m \times \Gamma_i$ for $i = 2, 3$ and 4 defined by:

$$F_{1i}(X_{ts}) = (f_1(X_{t(s)}), f_i(X_{(t)s})) \tag{11}$$

where $f_1 : \{X_{t(s)}\} \to S_m$ and $f_i : \{X_{(t)s}\} \to \Gamma_i$ for $i = 2, 3, 4$ are defined as above.

Notice that, when testing for spatiotemporal independence, when $i = 2, 3$ the symbolization map F_{1i} is standard, while for $i = 4$ is non-standard.

It is also possible to define an extension of the symbolization map f_2 in a spatiotemporal context. Indeed, consider the following map:

$$g : \{X_{ts}\}_{t\in I', s\in S'} \to \prod_{i=1}^{m_t} \Gamma_2 \tag{12}$$

defined by

$$g(X_{ts}) = (f(X_{(t)s}), f(X_{(t+1)s}), \ldots, f(X_{(t+m_t-1)s}))$$

where $f(X_{(t+i)s}) = (\mathcal{I}_{ts,(t+i)s_1}, \mathcal{I}_{ts,(t+i)s_2}, \ldots, \mathcal{I}_{ts,(t+i)s_{m_t-1}}))$ for all $i = 0, 1, \ldots, m_t - 1$ and the indicator function $\mathcal{I}_{ts,(t+i)s_j}$ is defined as in (7).

Accordingly, using this symbolization map, the next corollary straightforwardly follows from Theorem 2:

Corollary 3. *Let $F_{1i} : \{X_{ts}\}_{t\in I', s\in S'} \to S_{m_t} \times \Gamma_i$ and $g : \{X_{ts}\}_{t\in I', s\in S'} \to \prod_{i=1}^{m_t} \Gamma_2$ be the standard symbolization maps defined in (11) with $i = 2, 3$ and in (12) respectively. Denote by $h(S_{m_t} \times \Gamma_i)$ and $h(\prod_{i=1}^{m_t} \Gamma_2)$ the symbolic entropy defined in (1). If the spatiotemporal process $\{X_{ts}\}_{t\in I, s\in S}$ is independent, then:*

- *If indexes sets are perfect then:*

  1. *The affine transformations of the symbolic entropy*

$$STG(S_{m_t} \times \Gamma_i) = 2RT[Ln(n_i) - h(S_{m_t} \times \Gamma_i)] \tag{13}$$

$$STG(\prod_{i=1}^{m_t} \Gamma_2) = 2RT[Ln(2^{m_t(m_s-1)}) - h(\prod_{i=1}^{m_t} \Gamma_2)] \tag{14}$$

  are asymptotically χ_q^2 distributed. In (13) for $i = 2$ $q = (m_t - 1)! 2^{m_s-1}$ and $n_2 = m_t! 2^{m_s-1}$, and for $i = 3$ $q = (m_t! - 1)((k-1)^2 + 2)$ and $n_3 = m_t! k^2$. In (14) $q = 2^{m_t(m_s-1)}$.

2.   Then $(\frac{Y_{\sigma_1}-Rp_{\sigma_1}}{\sqrt{Rp_{\sigma_1}(1-p_{\sigma_1})}}, \frac{Y_{\sigma_2}-Rp_{\sigma_2}}{\sqrt{Rp_{\sigma_2}(1-p_{\sigma_2})}}, \dots, \frac{Y_{\sigma_{2m-1}}-Rp_{\sigma n}}{\sqrt{Rp_{\sigma n}(1-p_{\sigma n})}})$ is a multivariate normal distribution $MN(0, I)$ where n is the cardinality of the set of symbols.

- If indexes sets are not perfect then

  1.   If the sets B_i's have small cardinality for all symbol η, then the affine transformations of the symbolic entropy

$$STG(S_m \times \Gamma_i) = 2RT[Ln(n_i) - h(S_m \times \Gamma_i)] \tag{15}$$

$$STG(\prod_{i=1}^{m} \Gamma_2) = 2RT[Ln(2^{m_t(m-1)}) - h(\prod_{i=1}^{m_t} \Gamma_2)] \tag{16}$$

are asymptotically χ_q^2 distributed. In (15) for $i = 2$ $q = (m_t - 1)!2^{m_s-1}$ and $n_2 = m_t!2^{m_s-1}$, and for $i = 3$ $q = (m_t! - 1)((k-1)^2 + 2)$ and $n_3 = m_t!k^2$. In (16) $q = 2^{m_t(m_s-1)}$.

  2.   Then $(\frac{Y_{\eta_1}-Rp_{\eta_1}}{\sqrt{Var(Y_{\eta_1})}}, \frac{Y_{\eta_2}-Rp_{\eta_2}}{\sqrt{Var(Y_{\eta_2})}}, \dots, \frac{Y_{\eta n}-Rp_{\eta n}}{\sqrt{Var(Y_{\eta n})}})$ is a multivariate normal distribution $MN(0, \Sigma)$ where the variance and covariance matrix can be estimated using (4) and (5) and n is the cardinality of the set of symbols.

8. Empirical Behavior of the Tests for Spatiotemporal Independence

In this section we evaluate the empirical behavior of the STG test with different configurations for the subset I'. The first aim of this section is to show the flexibility of Corollary 3 to cope with different scenarios. The second goal is to evaluate the empirical behavior of the new test. An the third intention of this simulation is to evaluate the incidence of the selection of I' on the empirical size of the test and on the power.

To those ends we designed a Monte Carlo experiment as follows: Firstly we consider the problem of testing for independence on regular lattices of several orders—R = 64 (8 × 8) and T = 150; R = 100 (10 × 10)—for which we consider two possible temporal scenarios, depending on data availability, T = 200 and T = 800. We also simulated richer regular lattices of order R = 400 (20 × 20), although on this occasion we only considered T = 200. The symbolization map follows from (12) with $m_s = 4$ and $m_t = 2$. The test under study was generated from Corollary 3 under Expression (13). Therefore, we used a perfect indexes subset. This subset was constructed recursively, as indicated in Section 7: $N_{s_i t_j} \cap N_{s_r t_k} = \emptyset$, where $N_{s_i t_j}$ is the set conformed with s_i, the three nearest neighbors of s_i in t_j and the four spatial locations in the next time period. The power of the test is evaluated with the following DGPs:

$$DGP1 : y_t = (I - \gamma W)^{-1}(\alpha y_{t-1} + \lambda + \varepsilon_t)$$
$$DGP2 : y_t = \sqrt[3]{(I - \gamma W)^{-1}(\alpha y_{t-1} + \lambda + \varepsilon_t)}$$

where $\varepsilon_t \sim N(0,1)$, which was also used for evaluating the empirical size of the test. Parameters α, γ intensified temporal and spatial dependencies, respectively, and λ was fixed at five in all simulations. The weighting matrix W has been specified as a binary type using a contiguity criterion and rook-type movements.

Table 1 collects the empirical size and power of STG statistical test for 1000 repetitions. It is straightforward to observe that the size is controlled, and the test is powerful. For low intensity level of parameter ($\alpha = \gamma = 0.1$) the test is absolutely powerful. We have to set $\alpha = 1/40, \gamma = 1/25$ (or below) to lose power. This occurred despite the DGP under consideration.

Table 1. Size and power of STG test under a perfect subset of indexes.

			DGPs1			DGPs2		
		Size		**Power**				
R	T		$\alpha = 0.1$ $\gamma = 0.1$	$\alpha = 2/80,$ $\gamma = 2/50$	$\alpha = 1/80,$ $\gamma = 1/50$	$\alpha = 0.1$ $\gamma = 0.1$	$\alpha = 2/60,$ $\gamma = 2/30$	$\alpha = 1/60,$ $\gamma = 1/30$
64	150	0.060	1.000	0.096	0.073	1.000	0.152	0.077
100	200	0.064	1.000	0.149	0.065	1.000	0.220	0.086
100	800	0.053	1.000	0.492	0.112	1.000	0.860	0.196
400	200	0.058	1.000	0.412	0.111	1.000	0.772	0.160

Regular spatiotemporal configurations are interesting because (1) time series posit a natural order for observations, (2) lattice data provide the simplest extension of time series and (3) some scientific methods are compatible with this spatiotemporal configuration. However, irregular patterns are of frequent occurrence with spatial data. In geographical settings, data are liable to be recorded across heterogeneously-sized administrative regions, while economic distances do not correspond to regular spacing. Therefore, it is also useful to adapt the STG symbolic test to irregular spatiotemporal settings. In terms of our general methodology (see Corollary 3) this problem in tractable by considering the symbolization map $F_{1i}, i = 2$ where we control the dependence among the indicators by controlling on average the cardinality of the sets B_i. Particularly, we will select the set of indexes I' such that $|\bar{B}_i| \leq (m_s + m_t - 1)/2$; i.e., the average of the cardinality of the sets B_i is less than half of the number of spationtemporal neighbors.

Therefore, to complete the experiment (in the case of nonperfect lattices) we evaluate the STG-version for irregular lattices where coordinates of each spatial location are drawn from a N(0,1). We have considered the three nearest neighbors for irregular lattices. Afterwards, the resulting matrix was row-standardised in the usual way.

Table 2 collects the size and power for models constructed from DGP1 and DGP2. The introduction of irregular lattices has led us to introduce non-perfect indexes, and accordingly the size of the test slightly increased, although the levels seem acceptable, particularly for generous sample data. Power is as interesting as for the case of perfect indexes, and therefore the same comments applies (similar results are obtained in the case of using the multivariate normal approximation).

Table 2. Size and power of STG test under a non-perfect subset of indexes.

			DGPs1			DGPs2		
		Size		**Power**				
R	T		$\alpha = 0.1$ $\gamma = 0.1$	$\alpha = 2/80,$ $\gamma = 2/50$	$\alpha = 1/80,$ $\gamma = 1/50$	$\alpha = 0.1$ $\gamma = 0.1$	$\alpha = 2/60,$ $\gamma = 2/30$	$\alpha = 1/60,$ $\gamma = 1/30$
64	150	0.077	1.000	0.094	0.070	1.000	0.082	0.082
100	200	0.070	1.000	0.211	0.098	1.000	0.286	0.096
100	800	0.077	1.000	0.794	0.201	1.000	0.410	0.274
400	200	0.060	1.000	0.758	0.196	1.000	0.467	0.278

Comparison with Other Spatiotemporal Test for Independence

We now face our test with an unfavorable scenario characterized by small amount of available data on irregular lattices, also in linear and nonlinear setups. To this end we consider pairs of the following sample sizes: (36×10), (64×10), (100×10), (100×30) and (200×10). According to our theoretical discussion, given data scarcity and irregular spatial configuration, we use the non-perfect subset of indexes. Additionally, we consider the symbolization map based on equivalence relations $\bar{F}_{1,2} = (f_1, \bar{f}_2)$ as depicted in Appendix A.3 for nonstandard maps.

To complete the empirical study, we compare our test with another nonparametric spatiotemporal test [31] which is described in Appendix A.4 and we refer to it as STBP.

Notice, however, that the STBP test requires one to correctly specify the weighting matrix, W; this is not a requirement for the symbolic test.

In terms of empirical size, both tests behave similarly well for linear processes (Table 3). On the contrary, for the nonlinear processes (Table 4), the size of the STBP test is poor, while the symbolic-based test performs as expected. In terms of empirical power, the STBP test outperforms the STG test, especially for low intensity levels of dependence in the case of the linear process. However, under a nonlinear spatiotemporal configuration, the STG clearly presents a better balance between size and power and outperforms the STBP in all cases.

Table 3. Size and power of STG and STBP tests under linearity.

| | | Size | | Power | | | | | |
| | | STG | STBP | STG | STBP | STG | STBP | STG | STBP |
R	T			$(\alpha = 5/80, \gamma = 5/50)$		$(\alpha = 10/80, \gamma = 10/50)$		$(\alpha = 15/80, \gamma = 15/50)$	
36	10	0.069	0.046	0.096	0.154	0.328	0.615	0.787	0.989
64	10	0.057	0.054	0.145	0.193	0.504	0.815	0.962	0.999
100	10	0.063	0.066	0.192	0.228	0.677	0.926	0.993	1.000
100	30	0.058	0.054	0.443	0.395	0.982	1.000	1.000	1.000
200	10	0.063	0.067	0.346	0.327	0.935	0.997	1.000	1.000

Table 4. Size and power of STG and STBP tests under nonlinearity.

| | | Size | | Power | | | | | |
| | | STG | STBP | STG | STBP | STG | STBP | STG | STBP |
R	T			$(\alpha = 1/10, \gamma = 1/10)$		$(\alpha = 2/10, \gamma = 2/10)$		$(\alpha = 3/10, \gamma = 3/10)$	
36	10	0.060	0.002	0.094	0.007	0.269	0.060	0.507	0.473
64	10	0.055	0.000	0.111	0.002	0.376	0.119	0.784	0.743
100	10	0.067	0.000	0.175	0.011	0.557	0.209	0.941	0.915
100	30	0.056	0.000	0.357	0.005	0.949	0.642	1.000	1.000
200	10	0.062	0.001	0.281	0.007	0.856	0.381	1.000	0.998

9. Conclusions

Central null hypotheses in experimental and non-experimental branches of science can be easily tested by means of symbolized information. This paper provides with the analytical tools to construct nonparametric hypothesis tests based on symbols. These tools are able to cope with different null hypotheses and with distinct scenarios in which some realistic limitations might be imposed to test designs.

A shared characteristic of all these symbolic test families is that few assumptions are needed to obtain asymptotic results. Therefore, general applicability of this method is guaranteed. In particular, in this paper we have shown that two well-known symbolic-based tests are particular cases of the main symbolic theorem (Theorem 2), which is stated in this paper for the first time. Furthermore, a set of new symbolic-based tests for spatiotemporal independence is put forward by using the main results of this paper collected under the main symbolic theorem (Theorem 2). Monte Carlo simulations provide evidence of the extraordinary power of the contrasted test. Currently, there are circumstances where robustness to speed, noise or computational cost are paramount, so fruitful applications of symbolic analysis are favored.

Further lines of research are worthy. We now indicate some of them on which we and other scholars are currently working: (i) One of the appealing properties of symbolic-based testing is that it requires few assumptions. In this paper we have assumed stationarity; however, it would be interesting to study whether it is possible to be less restrictive. (ii) In the context of time series analysis, most available techniques require the existence of second moments; however, by using certain symbolizations, it might be possible to waive this requirement. This will allow time series researchers to consider a wider variety of model classes. (iii) One of the main contributions of the paper is that it suggests that researchers

can design a symbolization procedure (map) to test null hypotheses. It would be interesting to study what types of null hypotheses are more suitable to analysis using symbolic maps.

Appendix A

Appendix A.1. Maximum Likelihood Estimators

Since $\sum\limits_{i=1}^{n} p_{\eta_i} = 1$ it follows that

$$L(p_{\eta_1}, p_{\eta_2}, \ldots, p_{\eta_n}) = \frac{R!}{n_{\eta_1}! n_{\eta_2}! \cdot \ldots \cdot n_{\eta_n}!} p_{\eta_1}^{n_{\eta_1}} p_{\eta_2}^{n_{\eta_2}} \cdots (1 - p_{\eta_1} - p_{\eta_2} - \cdots - p_{\eta_{n-1}})^{n_{\eta_n}}.$$

Then the logarithm of this likelihood function remains as

$$
\begin{aligned}
Ln(L(p_{\eta_1}, p_{\eta_2}, \ldots, p_{\eta_n})) &= Ln(\frac{R!}{n_{\eta_1}! n_{\eta_2}! \cdot \ldots \cdot n_{\eta_n}!}) + \sum_{i=i}^{n-1} n_{\eta_i} Ln(p_{\eta_i}) \\
&+ n_{\eta_n} Ln(1 - p_{\eta_1} - p_{\eta_2} - \cdots - p_{\eta_{n-1}}).
\end{aligned}
$$

Maximum likelihood estimators $\widehat{p}_{\eta_i} = \frac{n_{\eta_i}}{R}$ are obtained by solving the following equation

$$\frac{\partial Ln(L(p_{\eta_1}, p_{\eta_2}, \ldots, p_{\eta_n}))}{\partial p_{\eta_i}} = 0.$$

Appendix A.2. Chi-Squared Distributions for Symbolic Maps

Notice that $2Ln(\lambda(Y))$ asymptotically follows a Chi-squared distribution with k degrees of freedom, which will depend on the set of symbols. Hence

$$-2Ln(\lambda(Y)) = -2[R Ln(R) + \sum_{i=1}^{n} n_{\eta_i} Ln\left(\frac{p_{\eta_i}}{n_{\eta_i}}\right)] \sim \chi_k^2$$

Now, if the symbolization map f is standard, that is, under the null H_0 all the symbols have the same probability to occur, $p_{\eta_i} = \frac{1}{n}$ for all $i = 1, 2, \ldots, n$, then it follows that

$$
\begin{aligned}
-2Ln(\lambda(Y)) &= -2R[Ln(R) + \sum_{i=1}^{n} \frac{n_{\eta_i}}{R} Ln\left(\frac{p_{\eta_i}}{n_{\eta_i}}\right)] \\
&= -2R[Ln(R) + \sum_{i=1}^{n} \frac{n_{\eta_i}}{R}(Ln\left(\frac{1}{n}\right) - Ln(n_{\eta_i}))] \\
&= -2R[Ln(R) + \sum_{i=1}^{n} \frac{n_{\eta_i}}{R}(Ln\left(\frac{1}{n}\right) - Ln\left(\frac{n_{\eta_i}}{R}\right) - Ln(R))]
\end{aligned}
$$

Now taking into account that $h(\Gamma) = -\sum_{i=1}^{n} p_{\eta_i} Ln(p_{\eta_i}) = -\sum_{i=1}^{n} \frac{n_{\eta_i}}{R} Ln\left(\frac{n_{\eta_i}}{R}\right)$, we have that $-2Ln(\lambda(Y)) = 2R[Ln(n) - h(\Gamma)]$.

Appendix A.3. Non-Standard Symbolization Maps

This appendix gives a procedure to construct a non-standard symbolization maps.

Let $\mathcal{A} = \{A_1, A_2, \ldots, A_d\}$ be a family of nonempty subsets of the set of symbols Γ. Assume that $\mathcal{A}$ is a partition of Γ, that is, $\Gamma = \bigcup_{i=1}^{d} A_i$ and $A_1 \cap A_j = \emptyset$ for all $i \neq j$. Now we define the relation $\sim$ in Γ by $\eta \sim \sigma$ if and only if $\eta, \sigma \in A_i$ for some $i = 1, 2, \ldots, d$. Obviously the relation $\sim$ is an equivalence relation and therefore we can consider the quotient set $\overline{\Gamma} = \Gamma/\sim$ formed by all the classes of equivalence as a new set of symbols. Denote by $\overline{\sigma}$ the class o equivalence of symbol σ. Therefore, there exists a natural projection $\pi : \Gamma \to \overline{\Gamma}$ defined by $\pi(\sigma) = \overline{\sigma}$. Moreover, any symbolization map with set of symbols Γ, namely $f : \{X_i\}_{i \in I'} \to \Gamma$, can be extended to a symbolization map with set of symbols $\overline{\Gamma}$ by considering the following map

$$\overline{f} = \pi \circ f : \{X_i\}_{i \in I'} \to \overline{\Gamma}$$

defined by $\overline{f}(X_i) = \overline{f(X_i)}$.

Notice that the cardinality of the set $\overline{\Gamma}$ is d which is always smaller or equal than the cardinality of the former set Γ.

Appendix A.4. A Generalization to Spatiotemporal Data of the Brett and Pinkse Test

The test of [32] is a nonparametric test of spatial dependence based on that two variables are independent if the joint characteristic function factorizes into the product of the marginal characteristic functions. We adapt this test to its use for studying spatiotemporal data.

Let $\{y_{ts}\}_{t \in I, s \in S}$ a spatiotemporal realization of a process. The y_{ts}'s can have continuous, discrete or mixed distributions, and the distribution functions are generally unknown. Under the null hypothesis, the spatiotemporal process is stationary and independent in space and time.

The design of the test is as follows. Let g be any practitioner-chosen density function with infinite support, and denote by $h(x) = \int e^{iux} g(u) du$ the Fourier transform of g. Given a location s, N_s refers to the set of neighbors of coordinate s. Fix a positive integer m and define $N_{ts}^m = \{t's' \mid t' = t, t+1, \ldots, t+m-1; s' \in N_s\}$ as the set of nearby observations to location s in period t and $n_{ts} = \sharp N_{ts}^m$ as the number of observations. Let $y_{ts}^N = \frac{1}{n_{ts}} \sum_{rk \in N_{ts}^m} y_{rk}$ stands for the sampling average of the proximate observations to y_{ts}. The *STBP* test null hypothesis is

H$_0$. *y_{ts} and y_{ts}^N are independent for all $t \in I$ and $s \in S$.*

Define $h_{(t_1s_1,t_2s_2)} = h(y_{t_1s_1} - y_{t_2s_2})$ and $h^{NN}_{(t_1s_1,t_2s_2)} = h(y^N_{t_1s_1} - y^N_{t_2s_2})$. Introduce η_{n1}, η_{n2} and η_{n3} defined by

$$\eta_{n1} = n^{-2} \sum_{ts,rk} h_{(ts,rk)} h^{NN}_{(ts,rk)}; \quad \eta_{n2} = n^{-3} \sum_{ts,rk,uv} h_{(ts,rk)} h^{NN}_{(ts,uv)};$$
$$\eta_{n3} = n^{-4} \sum_{ts,rk,uv,pq} h_{(ts,rk)} h^{NN}_{(uv,pq)},$$

where $n = RT$ is the number of observations. Let

$$\eta_n = (\eta_{n1} - \eta_{n2})^2 + (\eta_{n2} - \eta_{n3})^2$$

and

$$\nu_n = (\gamma_n - \mu_n^2)^2 n^{-1} \sum_{ts} n_{ts}^{-1} \left(I(n_{ts} > 0) + \sum_{rk} n_{rk}^{-1} I(rk \in N_{ts}) I(ts \in N_{rk}) \right)$$

where $\mu_n = n^{-2} \sum_{t,s} h_{ts}$, $\gamma_n = n^{-3} \sum_{s,t,u} h_{ts} h_{tu}$ and $I(\cdot)$ is an indicator function.

Under the null of the test, the extension of the Brett and Pinkse statistic for a spatiotemporal context is the following:

$$STBP = \frac{n\eta_n}{2\nu_n}$$

which is asymptotically χ_1^2 distributed.

The following two sufficient conditions are required by the *STBP* test to be consistent: (1) spatiotemporal dependence of a fixed order, (2) the sequence has to be strongly mixing. Strong mixing is a weak dependence condition, while fixed ordered dependence is a restriction regarding these relationships must be produced between proximate observations. In this case, then the null hypothesis will be asymptotically rejected; however the behavior of the test will be undetermined, when the dependence involves observations that are not geographically or temporally proximate. Using spatial data, this means that an specification of the so-called spatial weighting matrix is needed and that this specification must be correct (Lopez et al, 2011, for more details).

References

1. Amigo, J.M. *Permutation Complexity in Dynamical Systems*, 1st ed.; Springer: Berlin, Germany, 2010.
2. Weaver, W.; Shannon, C.E. *The Mathematical Theory of Communication*; University of Illinois: Urbana, IL, USA, 1949.
3. Morse, M. Recurrent geodesics on a surface of negative curvature. *Trans. Am. Math. Soc.* **1921**, *22*, 84. [CrossRef]
4. Collet, P.; Eckmann, J.P. *Iterated Maps on the Interval as Dynamical Systems*; Brickhauser: Basel, Switzerland, 1980.
5. Ruiz, M.; Matilla-García, M.; García, J.A.; Susillo, J.L.; Romo, A.; Gonzalez, A.; Ruiz, A.; Gayan, J. An Entropy Test for Single-Locus Genetic Association Analysis. *BMC Genet.* **2010**, *11*, 1–15.
6. Daw, C.S.; Finney, C.E.A.; Tracy, E.R. A review of symbolic analysis of experimental data. *Rev. Sci. Instrum.* **2003**, *74*, 915–930. [CrossRef]
7. Hao, B.; Zheng, W. *Applied Symbolic Dynamics and Chaos*; World Scientific: Singapore, 1998.
8. Lunde, A.; Timmermann, A.G. Duration Dependence in Stock Prices:An Analysis of Bull and Bear Markets. *J. Bus. Econ. Stat.* **2004**, *22*, 253–273. [CrossRef]
9. Robinson, P. Large-Sample Inference on Spatial Dependence. *Econom. J.* **2009**, *12*, 68–82. [CrossRef]
10. López, F.; Matilla-García, M.; Mur, J.; Ruiz, M. Non-Parametric Spatial Independence Test Using Symbolic Entropy. *Reg. Sci. Urban Econ.* **2010**, *40*, 106–115. [CrossRef]
11. García-Córdoba, J.A.; Matilla-García, M.; Ruiz Marin, M. A test for deterministic dynamics in spatial processes. *Spat. Econ. Anal.* **2019**, *14*, 361–377. [CrossRef]
12. Conaway, M.R. Analysis of repeated categorical measurements with conditional likelihood methods. *J. Am. Stat.* **1989**, *84*, 53–62. [CrossRef]
13. Stephenson, D.B. Use of the "Odds Ratio" for Diagnosing Forecast Skill. *Weather. Forecast.* **2000**, *15*, 221–232. [CrossRef]
14. Bandt, C.; Shiha, F. Order patterns in time series. *J. Time Ser. Anal.* **2007**, *28*, 646–665. [CrossRef]
15. Matilla-García, M. A non-parametric test for independence based on symbolic dynamics. *J. Econ. Dyn. Control.* **2007**, *31*, 3889–3903. [CrossRef]

16. Matilla-García, M.; Ruiz, M. A non-parametric independence test using permutation entropy. *J. Econom.* **2008**, *144*, 139–155. [CrossRef]

17. Sinn, M.; Keller, K. Estimation of ordinal pattern probabilities in Gaussian processes with stationary increments. *Comput. Stat. Data Anal.* **2011**, *55*, 1781–1790. [CrossRef]

18. Schnurr, A. An Ordinal Pattern Approach to Detect and to Model Leverage Effects and Dependence Structures Between Financial Time Series. *Stat. Pap.* **2014**, *55*, 919–931. [CrossRef]

19. Schnurr, A.; Dehling, H. Testing for Structural Breaks via Ordinal Pattern Dependence. *J. Am. Stat. Assoc.* **2017**. [CrossRef]

20. Caballero-Pintado, M.V.; Matilla-García, M.; Marín, M.R. Symbolic correlation integral. *Econom. Rev.* **2019**, *38*, 533–556. [CrossRef]

21. Caballero-Pintado, M.V.; Matilla-García, M.; Ruiz Marín, M. Symbolic recurrence plots to analyze dynamical systems. *Chaos Interdiscip. J. Nonlinear Sci.* **2018**, *28*, 063112. [CrossRef]

22. Hong, Y.; White, H. Asymptotic distribution theory for nonparametric entropy measures of serial dependence. *Econometrica* **2005**, *73*, 837–901. [CrossRef]

23. Faura, Ú; Lafuente, M.; Matilla-García, M.; Ruiz, M. Identifying the Most Relevant Lag with Runs. *Entropy* **2015**, *17*, 2706–2722. [CrossRef]

24. Cliff, A.D.; Ord, J.K. *Spatial Processes: Models and Applications*; Pion: London, UK, 1981.

25. Agresi , A. *An Introduction to Categorical Data Analysis*; John Wiley and Sons: Brisbane, Australia, 1996.

26. Bandt, C.; Keller, G.; Pompe, G. Entropy of interval maps via permutations. *Nonlinearity* **2002**, *12*, 1595–1602. [CrossRef]

27. Soon Spario, Y.T. Binomial Approximation for dependent indicators. *Stat. Sin.* **1996**, *6*, 703–714.

28. Ruiz, M.; López, F.; Páez, A. Testing for spatial association of qualitative data using symbolic dynamics. *J. Geogr.* **2009**, *12*, 281–308. [CrossRef]

29. Matilla-García, M.; Marín, M.R. A new test for chaos and determinism based on symbolic dynamics. *J. Econ. Behav. Organ.* **2010**, *76*, 600–614. [CrossRef]

30. Baltagi, B.; Kelejian, H.; Prucha, I. Anal. Spatiallydependent Data. *J. Econom.* **2007**, *140*, 1–4. [CrossRef]

31. López, F.A.; Matilla-García, M.; Mur, J.; Marín, M.R. Four tests of independence in spatiotemporal data. *Pap. Reg.* **2011**. [CrossRef]

32. Brett, C.; Pinkse, J. Those Taxes are all over the Map A Test for Spatial Independence of Municipal Tax Rates in British Columbia. *Int. Reg. Sci. Rev.* **1997**, *20*, 131–151. [CrossRef]