Statistica & Applicazioni - Vol. VIII, n. 1, 2010, pp. 57-74

ON THE USE OF χ^2-TEST TO CHECK SERIAL INDEPENDENCE

Luca Bagnato*
Antonio Punzo**

SUMMARY

In this paper two tests of serial independence are proposed. The building block of these procedures is the definition of a component χ^2-test for testing independence between pairs of lagged variables. With reference to different component χ^2-tests, it is shown that the corresponding test statistics are independent. Taking advantage of this result, the component tests are used from both a simultaneous and a direct viewpoint to define two different test procedures denoted by SIS (Serial Independence Simultaneous) and SICS (Serial Independence Chi-Square). Simulations are used to explore the performance of these tests in terms of size and power. Our results underline that both the proposals are powerful against various types of alternatives. It is also shown, through what we call Lag Subsets Dependence Plot (LSDP), how to detect possible lag(s)-dependences graphically. Some examples are finally provided in order to evaluate the effectiveness of the LSDP.

Keywords: *Nonlinear Time Series, Serial Independence, Simultaneous Tests, χ^2-test.*

1. INTRODUCTION

In modelizing observed time series the most popular approach consists in adopting the class of linear models. In this wide class, the well-known *white noise* process – characterized only by the properties of its first two moments – represents the building block and reflects information that is not directly observable. Nevertheless, the limitations of the linear models appear already in the classical paper by Moran (1953). Nowadays we know that there are nonstandard features, commonly referred as *nonlinear features* (Jianqing and Qiwei 2003) that, by definition, cannot be captured in the linear frame. In the attempt to overcome this problem, nonlinear models have to be taken into account (for a discussion on the chief objectives which guarantee the introduction of nonlinear models see, *e.g.*, Tjøstheim, 1994). For this kind of models, as indicated by Jianqing and Qiwei (2003), a white noise process is no longer a pertinent building block as we have to look for measures beyond the second moments to characterize the nonlinear dependence structure. Thus, the white noise has to be replaced by a noise process composed by independent and identically distributed ($i.i.d.$) variables.

---

* Dipartimento di Metodi Quantitativi per le Scienze Economiche e Aziendali - Università Milano-Bicocca - via Bicocca degli Arcimboldi, 8 - 20126 MILANO (e-mail: luca.bagnato@unimib.it).

** Dipartimento di Impresa, Culture e Società - Università di Catania - Corso Italia, 55 - 95129 CATANIA (e-mail: antonio.punzo@unict.it).

This paper results from the common work of the Authors. Luca Bagnato has written Sections 1, 2, and 5 while Antonio Punzo has written Sections 3 and 4.

As well as for serial uncorrelatedness tests in the linear case, tests for validating serial independence are fundamental in the nonlinear approach. On one hand, testing serial independence is often a preliminary step carried out before modeling the data generating process (*model selection*). On the other hand, tests for serial independence are useful statistical tools, once a nonlinear model is fitted to the observed data, to check possible dependences among estimated residuals (*model validation*).

The most commonly used test, both for linear and nonlinear models, is the Box-Ljung statistic. This test is not a real test of serial independence but rather of serial uncorrelatedness; it is clear, therefore, it will have very low power against residuals that are uncorrelated but dependent. In order to solve this problem, McLeod and Li (1983) suggest inserting squared residuals in the Box-Ljung statistic. By doing, however, the authors solve the problem of checking a possible quadratic correlation but, at the same time, their method may clearly fail to be consistent against different dependence structures.

Motivated by these considerations, various authors have looked at other tests of serial independence (see, *e.g.*, Hoeffding, 1948; Baek and Brock, 1992; Skaug and Tjøstheim, 1996; Tjøstheim, 1996; Johnson and McClelland, 1998; Pinkse, 1998, 1999; Diks and Panchenko, 2007; Diks, 2008). Note that, in principle, much of these tests were been conceived to check serial independence of time series. Nevertheless, according to Tjøstheim (1994), we believe that a lot of these tests may be appropriate to be applied to tests of independence of estimated residuals by applying similar reasoning to that used for the BDS test of Brock, Dechert and Scheinkman (1996).

This paper proposes two tests of serial independence that take advantage of the well-known χ^2-test applied to pairs of lagged variables. A graphic device is also presented as a natural extension of both test procedures. Obviously, being based on the χ^2-test, the proposals are valid asymptotically. Natural fields of application, where a large number of observations is typically available, are: environment, finance, quality control, and so on. Another important step in introducing the statistical procedures is the demonstration that, when considering two different component χ^2-tests, the corresponding test statistics are independent.

Proposals and paper can be so schematically summarized. Section 2 begins by presenting the common setting for both proposed tests. Although the treatment is carried out directly on the observed data, the proposed methods can be also easily applied – in the model validation phase – to the estimated residuals from a nonlinear model. The first test, presented in Section 2 A), exploits the union-intersection (UI) principle subdividing the ''overall'' null hypothesis of serial independence in l simpler component null hypotheses of serial independence of lag r, $r = 1, \dots, l$. Here, the independence among test statistics, shown in Theorem 1, is used in order to improve the classical *Bonferroni correction* (see Shaffer, 1995). In the second test, presented in Section 2 B), the component statistics are summed among them in order to define a new asymptotically χ^2-distributed statistic; this result is obtained by taking advantage of the independence among component test statistics. The performance of these tests, in terms of size and power, is faced in Section 3 by a wide simulation study. In Section 4 a graphic device connected with the SIS and SICS tests, which we have called Lag Subsets Dependence Plot (LSDP), is also presented. In a more informative way, this tool allows to evaluate dependence among all possible combinations of lags taken into account. In this section some simulations confirming such considerations are also provided. Finally, in Section 5, some conclusions are given.

2. THE PROPOSED SERIAL INDEPENDENCE TESTS

Let $\{X_t\}_{t\in\mathbb{N}}$ be a stationary and ergodic stochastic process. Let $(X_{t-n+1}, X_{t-n+2}, \ldots, X_t)$ be a random vector, of length n, which describes the realization of the process from time $t-n+1$ to time t. From now on with the lower case ''x'' we will indicate the observed counterpart of ''X''.

Suppose that we are interested in testing the null hypothesis of *serial independence* of the process by using the observed time series $(x_{t-n+1}, x_{t-n+2}, \ldots, x_t)$. In order to do this, choose an integer l, with $0 < l < n$, and consider Table 1 where, for each lag r, $r = 1, \ldots, l$, the observed information up to time $t-r$ represents the realizations of X_{t-r}. Moreover, considering $L = \{1, \ldots, l\}$, with l big enough, the null hypothesis of interest can be expressed as

$$H_0 : X_t \perp\!\!\!\perp X_{t-r} \quad \text{for all } r \in L, \tag{1}$$

versus the alternative

$$H_1 : X_t \not\perp\!\!\!\perp X_{t-r} \quad \text{for some } r \in L. \tag{2}$$

Obviously, by increasing l, the number of possible comparisons increases as well. Before going on, for comprehension's sake, consider $r = 1$: accepting $X_t \perp\!\!\!\perp X_{t-1}$, one also accepts implicitly that $X_{t-1} \perp\!\!\!\perp X_{t-2}$, $X_{t-2} \perp\!\!\!\perp X_{t-3}$, and so on. This issue can be easily understood by considering adjacent columns in Table 1.

The ''overall'' hypotheses in (1) and (2) can be easily re-expressed as follows

$$H_0 : \bigcap_{r=1}^{l} H_0^{(r)} \quad \text{versus} \quad H_1 : \bigcup_{r=1}^{l} H_1^{(r)}, \tag{3}$$

where

$$H_0^{(r)} : X_t \perp\!\!\!\perp X_{t-r} \tag{4}$$

TABLE 1. - *Lagged variables* $X_{t-r}, r = 1, \ldots, l, l < n$

X_t	Lagged variables				
	X_{t-1}	$\cdots$	X_{t-r}	$\cdots$	X_{t-l}
x_{t-n+1}					
x_{t-n+2}	x_{t-n+1}				
$\vdots$	$\vdots$	$\ddots$			
$x_{t-n+r+1}$	x_{t-n+r}	$\cdots$	x_{t-n+1}		
$\vdots$	$\vdots$		$\vdots$	$\ddots$	
$x_{t-n+l+1}$	x_{t-n+l}	$\cdots$	$x_{t-n+l+1-r}$	$\cdots$	x_{t-n+1}
$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_t	x_{t-1}	$\cdots$	x_{t-r}	$\cdots$	x_{t-l}

is the component null hypothesis concerning a sort of independence of lag r, while $H_1^{(r)}$ is the negation of (4), $r \in L$. The superscript "(r)" underlines that we are considering the comparison of lag r, that is, the variables X_t and X_{t-r}. In Table 1 the elements considered for this comparison are highlighted. In all comparisons we have decided to fix X_t which, from a practical point of view, it is equivalent to fixing the first column in Table 1. In analogy to the construction of the correlogram, this choice allows us to have the maximum number of elements in each comparison.

In order to test $H_0^{(r)}$, the first step is to create the bivariate joint distribution as shown in Table 2, where k is the number of classes for each variable and $a_0^{(r)}, a_1^{(r)}, \ldots, a_k^{(r)}$, $a = c, d$, are cutoff points. Thus, the simple and well-known χ^2-test of independence can be adopted to test the null hypothesis (4) with reference to Table 2. In particular we have denoted the test statistic for $H_0^{(r)}$ by

$$T_r = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\left[n_{ij}^{(r)} - \hat{n}_{ij}^{(r)}\right]^2}{\hat{n}_{ij}^{(r)}}, \tag{5}$$

where $n_{ij}^{(r)}$ are the observed frequencies and $\hat{n}_{ij}^{(r)}$ are the theoretical frequencies under the null hypothesis, $i, j = 1, \ldots, k$. It is well-known that under $H_0^{(r)}$ the test statistic (5) is asymptotically distributed as a χ^2 with $(k-1)^2$ degrees of freedom.

In order to compute T_r in (5), both the number of classes k and the cutoff points $a_0^{(r)}, a_1^{(r)}, \ldots, a_k^{(r)}$, $a = c, d$, have to be defined. We fix the extreme cutoff points as

$$\begin{aligned} c_0^{(r)} &= \min\{x_{t-n+r+1}, \ldots, x_t\} \\ c_k^{(r)} &= \max\{x_{t-n+r+1}, \ldots, x_t\} \\ d_0^{(r)} &= \min\{x_{t-n+1}, \ldots, x_{t-r}\} \\ d_k^{(r)} &= \max\{x_{t-n+1}, \ldots, x_{t-r}\}, \end{aligned} \tag{6}$$

where the arguments of the functions min and max are highlighted in Table 1. *Equifrequency classes* are adopted since simulations confirm that they preserve and highlight the dependence structure between pairs of variables. Moreover, such a choice allows us to define automatically the interior cutoff points once fixed k. Thus, k re-

TABLE 2. - *Bivariate frequencies table related to* $H_0^{(r)}$

X_t \ X_{t-r}	$d_0^{(r)} \dashv d_1^{(r)}$	$\cdots$	$d_{j-1}^{(r)} \dashv d_j^{(r)}$	$\cdots$	$d_{k-1}^{(r)} \dashv d_k^{(r)}$	
$c_0^{(r)} \dashv c_1^{(r)}$	$n_{11}^{(r)}$	$\cdots$	$n_{1j}^{(r)}$	$\cdots$	$n_{1k}^{(r)}$	
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$c_{i-1}^{(r)} \dashv c_i^{(r)}$	$n_{i1}^{(r)}$	$\cdots$	$n_{ij}^{(r)}$	$\cdots$	$n_{ik}^{(r)}$	
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$c_{k-1}^{(r)} \dashv c_k^{(r)}$	$n_{k1}^{(r)}$	$\cdots$	$n_{kj}^{(r)}$	$\cdots$	$n_{kk}^{(r)}$	
						$n^{(r)}$

mains the only parameter to be chosen in order to construct Table 2. We have decided to apply the same value of k to each comparison; naturally, the procedure can be generalized by setting k as a function of r.

In choosing k, the number of observations has to be taken into account. From Table 1 it is possible to note that we have $n^{(r)} = n - r$ available observations in constructing Table 2. In order to make valid the r-th component χ^2-test, the well-known rule of thumb $\widehat{n}_{ij}^{(r)} \geq 5$, for each $i, j = 1, \ldots, k$ and $r \in L$, has to hold. This means that the number of observations, in comparison of lag r, has to satisfy the inequality $n^{(r)} \geq 5k^2$. Being k equal for each comparison, and being l the maximum value of r, we have

$$k \leq \sqrt{\frac{n^{(l)}}{5}}. \tag{7}$$

It is clear that, if choosing a too small k such as $k = 2$ or $k = 3$, the dependence structure between pairs of lagged variables could be obscured. However, in most of simulations illustrated in the next section, we will show that there is no more difference in choosing k between its possible values. Finally, note that both joint frequencies and classes vary by changing compared variables.

Before presenting our proposals, we provide the following theorem showing useful relations between test statistics referred to different lags.

Theorem 1

Suppose $\{X_t\}_{t \in \mathbb{N}}$ be an i.i.d. stochastic process. Let $(x_{t-n+1}, x_{t-n+2}, \ldots, x_t)$ be an observed time series, of length n, arising from the process. As in Table 1, let $(x_{t-n+r+1}, \ldots, x_t)$ be the realizations of the variable X_t and $(x_{t-n+1}, \ldots, x_{t-r})$ be the realizations of the generic variable X_{t-r}, $r = u, v$. Suppose to build the two bivariate joint distributions in Table 3(a) and Table 3(b) where the extreme cutoff points are fixed as in (6). *Moreover, adopt equifrequency classes for the two marginal distributions in each table. Then, the random variables T_u and T_v, derived as in* (5) *using respectively Table 3(a) and Table 3(b), are independent for any $u \neq v$.*

Proof:

Table 3. - *Bivariate frequency tables*

(a) Variables X_t and X_{t-u}

X_t \ X_{t-u}	$d_0^{(u)} \dashv d_1^{(u)}$	$\cdots$	$d_{j-1}^{(u)} \dashv d_j^{(u)}$	$\cdots$	$d_{k-1}^{(u)} \dashv d_k^{(u)}$	
$c_0^{(u)} \dashv c_1^{(u)}$	$n_{11}^{(u)}$	$\cdots$	$n_{1j}^{(u)}$	$\cdots$	$n_{1k}^{(u)}$	
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$c_{i-1}^{(u)} \dashv c_i^{(u)}$	$n_{i1}^{(u)}$	$\cdots$	$n_{ij}^{(u)}$	$\cdots$	$n_{ik}^{(u)}$	
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$c_{k-1}^{(u)} \dashv c_k^{(u)}$	$n_{k1}^{(u)}$	$\cdots$	$n_{kj}^{(u)}$	$\cdots$	$n_{kk}^{(u)}$	
						$n^{(u)}$

(b) Variables X_t and X_{t-v}

X_t \ X_{t-v}	$d_0^{(v)} \dashv d_1^{(v)}$	$\cdots$	$d_{j-1}^{(v)} \dashv d_j^{(v)}$	$\cdots$	$d_{k-1}^{(v)} \dashv d_k^{(v)}$	
$c_0^{(v)} \dashv c_1^{(v)}$	$n_{11}^{(v)}$	$\cdots$	$n_{1j}^{(v)}$	$\cdots$	$n_{1k}^{(v)}$	
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$c_{i-1}^{(v)} \dashv c_i^{(v)}$	$n_{i1}^{(v)}$	$\cdots$	$n_{ij}^{(v)}$	$\cdots$	$n_{ik}^{(v)}$	
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$c_{k-1}^{(v)} \dashv c_k^{(v)}$	$n_{k1}^{(v)}$	$\cdots$	$n_{kj}^{(v)}$	$\cdots$	$n_{kk}^{(v)}$	
						$n^{(v)}$

Without loss of generality, for $u \neq v$, consider the elements $n_{11}^{(u)}$ and $n_{11}^{(v)}$ in position (1,1) of Table 3(a) and Table 3(b), respectively. Since, by definition, $X_t \perp\!\!\!\perp X_{t-r}$, $r = u, v$, these elements can be expressed in the following way

$$n_{11}^{(u)} = \sum^{n^{(u)}} I_{\left[c_0^{(u)},\, c_1^{(u)}\right]}(X_t) \cdot I_{\left[d_0^{(u)},\, d_1^{(u)}\right]}(X_{t-u}), \tag{8}$$

$$n_{11}^{(v)} = \sum^{n^{(v)}} I_{\left[c_0^{(v)},\, c_1^{(v)}\right]}(X_t) \cdot I_{\left[d_0^{(v)},\, d_1^{(v)}\right]}(X_{t-v}), \tag{9}$$

where

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, since the process is supposed to be *i.i.d.*, we have $X_{t-u} \perp\!\!\!\perp X_{t-v}$. Consequently, also (8) and (9) are independent as well. The same reasoning can be applied for the other elements $n_{ij}^{(u)}$ and $n_{ij}^{(v)}$ of Table 3(a) and Table 3(b), respectively. Since the two single χ^2-test statistics T_u and T_v have been built using sums of independent random variables, then they are independent for any $u \neq v$.

A) Serial Independence Simultaneous Test By construction, a statistical test for (3) must adhere to the union-intersection (UI) principle introduced by Roy (1953) as a heuristic method for test construction. Under the UI principle, the overall null hypothesis H_0 is accepted against the alternative if, and only if, all $H_0^{(r)}, r = 1, \ldots, l$, are simultaneously accepted. In other words, taking p-values into account, the null hypothesis is accepted if, and only if, the component test with the smallest p-value is accepted.

Differently from a possible ''direct'' test of serial independence, the present approach shows its profitability whenever the null hypothesis in (3) is rejected. Indeed, the l individual (component) tests for $H_0^{(r)}$ can still be employed to investigate the nature of the departure from H_0. This aspect will be taken into account in Section 4.

Aiming at completing this ''road'', we need to specify the significance levels at which each individual test of independence of lag r has to be performed in order to preserve the significance level α (chosen in advance) for the overall null hypothesis in (3).

Several classes of procedures have been proposed in literature (see Shaffer, 1995, for a survey). The first and foremost reason for choosing among them relies on the dependence/independence between test statistics referred to each individual hypothesis. This motivates and explains the use of the Bonferroni method (Seber, 1984) that is based on a probability inequality which does not require any assumption about the joint distribution of the individual test statistics or any specific dependence structure among them. Nevertheless, since we have proved in Theorem 1 that the χ^2-test statistics are independent, the classical Bonferroni method can be slightly improved by replacing the Bonferroni component significance levels α/l, for any $r = 1, \ldots, l$, by $1 - (1 - \alpha)^{1/l}$, always greater than α/l, although the difference is small for small values of α (Shaffer, 1995). For example, when $l = 10$ and $\alpha = 0.05$, each component significance level drops to 0.0051. Note that this method is as simple to use as the Bonferroni one. Alternative procedures with improved power have been proposed (see, *e.g.*, Holm, 1979 and Hommel, 1983) but it is worth noting that they are difficult to apply when the number of the individual tests increases.

We will call the entire described procedure, comprehensive of the above suggested correction, Serial Independence Simultaneous (SIS) test.

B) Serial Independence Chi-Square Test Alternatively, starting from (5), we propose to use the (positive) test statistic

$$T = \sum_{r=1}^{l} T_r = \sum_{r=1}^{l} \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\left[n_{ij}^{(r)} - \hat{n}_{ij}^{(r)}\right]^2}{\hat{n}_{ij}^{(r)}}, \tag{10}$$

assuming small values under H_0. Theorem 1 shows that under the null hypothesis of independence, the random variables T_r, $r \in L$, defining T, are independent too. Thus, the asymptotic distribution of T, used for testing independence, results to be a χ^2 with $l(k-1)^2$ degrees of freedom. We will call the resulting procedure Serial Independence Chi-Square (SICS) test.

3. SIMULATION STUDY

In this section a simulation study, by using the R environment, is described. The aim is twofold: at first the performance of the proposed tests of serial independence in terms of their empirical size and power is evaluated and then such performance is investigate at the varying of the number of classes k. In order to do this, some well-known models in the time series literature have been adopted. Details on these models, used also in Diks (2008), are contained in the second column of Table 4.

The first three *i.i.d.* models, which allow us to evaluate the performance of the tests in terms of size, are of type: standard normal (A), Student's t with 2 degrees of freedom (B), normal-mixture with density $6/10\phi(x; -5/3, 1) + 4/10\phi(x; 5/2, 1)$ that

TABLE 4. - *Simulated rejection rates at nominal size* $\alpha = 0.05$ *for the two proposed tests at the varying of* l *in* $\{2, \ldots, 10\}$. *Sample size* $n = 1,000$ *and number of simulations equal to* 1,000 *for each independently realized time series from each process are adopted*

Model		Specification	Test	Lag l: 2	3	4	5	6	7	8	9	10
size	(A)	$X_t = \varepsilon_t$	SIS	0.061	0.048	0.044	0.053	0.054	0.053	0.056	0.052	0.051
			SICS	0.053	0.051	0.052	0.059	0.054	0.062	0.064	0.054	0.068
	(B)	$X_t = \tau_t$	SIS	0.053	0.053	0.049	0.053	0.048	0.046	0.045	0.042	0.039
			SICS	0.048	0.048	0.043	0.038	0.032	0.031	0.036	0.039	0.036
	(C)	$X_t = \eta_t$	SIS	0.052	0.059	0.064	0.060	0.058	0.054	0.054	0.055	0.054
			SICS	0.049	0.058	0.058	0.057	0.050	0.055	0.054	0.050	0.047
power	(D)	$X_t = \varepsilon_t + 0.8\varepsilon_{t-1}^2$	SIS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			SICS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(E)	$X_t = \varepsilon_t + 0.6\varepsilon_{t-1}^2 + 0.6\varepsilon_{t-2}^2$	SIS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			SICS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(F)	$X_t = \varepsilon_t + 0.8\varepsilon_{t-1}\varepsilon_{t-2}$	SIS	0.566	0.513	0.471	0.445	0.412	0.400	0.383	0.359	0.348
			SICS	0.676	0.575	0.502	0.456	0.417	0.389	0.357	0.335	0.312
	(G)	$X_t = 0.3X_{t-1} + \varepsilon_t$	SIS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			SICS	1.000	1.000	1.000	0.998	0.996	0.993	0.994	0.986	0.985
	(H)	$X_t = 0.8\lvert X_{t-1}\rvert^{0.5} + \varepsilon_t$	SIS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			SICS	1.000	0.998	0.994	0.993	0.984	0.975	0.962	0.951	0.941
	(I)	$X_t = \text{sign}(X_{t-1}) + \varepsilon_t$	SIS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			SICS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(J)	$X_t = 0.6\varepsilon_{t-1}X_{t-2} + \varepsilon_t$	SIS	0.784	0.751	0.726	0.702	0.691	0.675	0.653	0.630	0.626
			SICS	0.824	0.774	0.725	0.669	0.630	0.592	0.552	0.515	0.483
	(K)	$X_t = \sqrt{h_t}\varepsilon_t,\ h_t = 1 + 0.4X_{t-1}^2$	SIS	0.855	0.817	0.799	0.781	0.761	0.751	0.735	0.719	0.711
			SICS	0.814	0.731	0.659	0.617	0.550	0.511	0.470	0.439	0.415
	(L)	$X_t = \sqrt{h_t}\varepsilon_t,\ h_t = 0.01 + 0.8h_{t-1} + 0.15X_{t-1}^2$	SIS	0.583	0.608	0.612	0.616	0.606	0.597	0.586	0.581	0.574
			SICS	0.670	0.734	0.766	0.789	0.797	0.801	0.807	0.803	0.799
	(M)	$X_t = \left[-0.5 + 0.9I_{[0,\infty]}(X_{t-1})\right]X_{t-1} + \varepsilon_t$	SIS	0.994	0.993	0.991	0.987	0.987	0.985	0.985	0.982	0.980
			SICS	0.986	0.957	0.939	0.919	0.887	0.854	0.834	0.806	0.784
	(N)	$X_t = 4X_{t-1}(1 - X_{t-1}),\ 0 < X_t < 1$	SIS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			SICS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(O)	$X_t = 1 + 0.3X_{t-2} - 1.4X_{t-1}^2$	SIS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			SICS	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(P)	$X_t = Z_t + \sigma\varepsilon_t,\ Z_t = 1 + 0.3Z_{t-2} - 1.4Z_{t-1}^2$	SIS	0.775	0.880	0.841	0.820	0.800	0.782	0.769	0.762	0.752
			SICS	0.817	0.945	0.921	0.920	0.892	0.877	0.849	0.836	0.812

ensures: zero mean, bimodality and skewness (C). With $\phi(\cdot; \mu, \sigma)$ a normal density with mean μ and standard deviation σ is indicated. Density plots of the variables related to these models are displayed in Figure 1. To evaluate the power, the models used to generate the series are of type: nonlinear moving average (D)-(F), linear autoregressive (G), nonlinear autoregressive (H), sign autoregressive (I), bilinear (J), ARCH(1) (K), GARCH(1,1) (L), threshold autoregressive (M), logistic map (N), Hénon map (O) and Hénon map with dynamic noise (P). For each of these sixteen models, the proposed tests to 1,000 simulated series, each of length $n = 1{,}000$, have been applied. The number of classes has been fixed to $k = 5$ and the nominal size has been set to $\alpha = 0.05$. Table 4 shows the (simulated) rejection rates for each model and both the proposed tests. The study has also the objective to evaluate the performance of the two proposals at the varying of l in $\{2, \ldots, 10\}$.

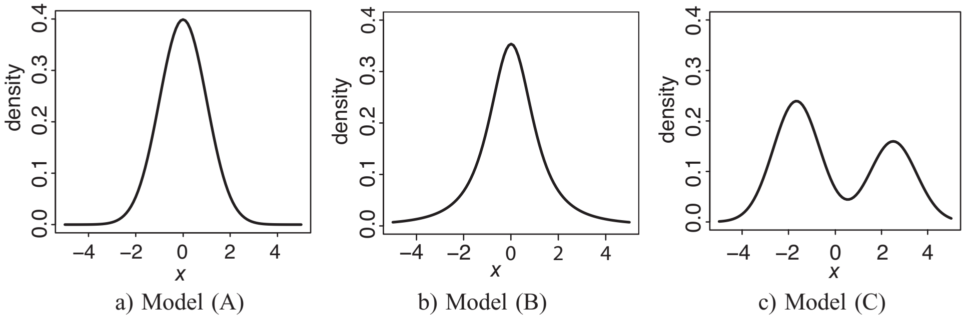


FIGURE 1. - *Density of variables related to models (A)-(C)*

Rejection rates for the *i.i.d.* processes – models (A)-(C) – are displayed in the first three rows of Table 4. These rates are very close to the nominal size 0.05 for both proposals and for all considered values of l. Note that the nominal sizes are well ap-

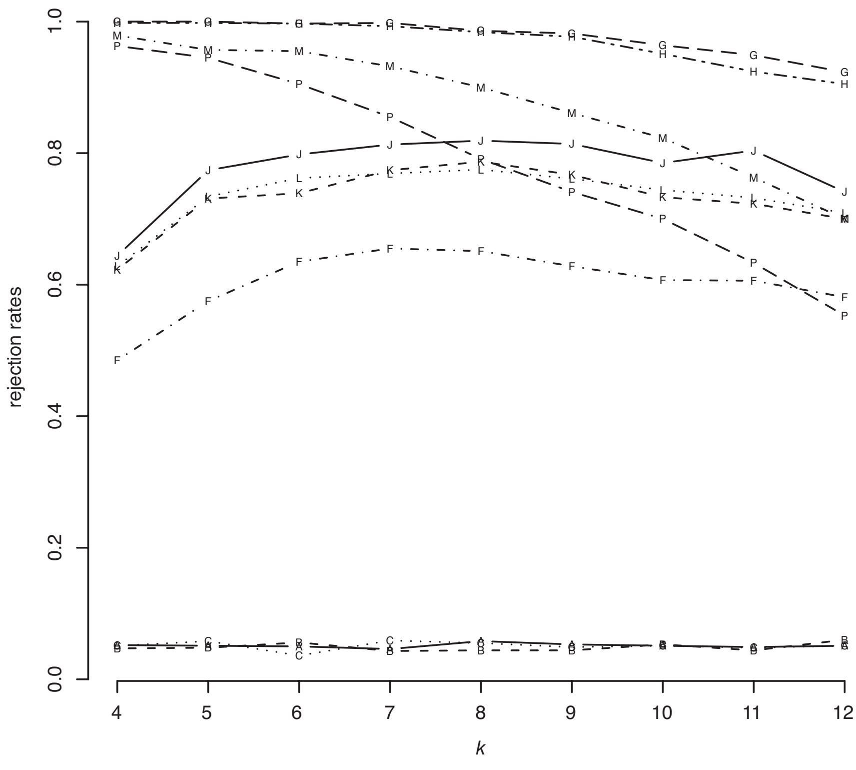


FIGURE 2. - *Rejection rates of the SICS test at the varying of k* ($n = 1{,}000$ *and* $l = 3$). *The numbers on each piecewise line highlight the model, of Table 4, taken into account*

proximated even if the true distribution is heavy tailed, model (B), and bimodal-skewed, model (C). In the remaining rows of Table 4, rejection rates for the serial dependent models (D)-(P) are shown. Both the proposals reveal to be powerful against these types of alternatives.

As just mentioned, in the illustrated simulations a number of classes equal to $k = 5$ has been adopted. Obviously, this is only one of the possible values satisfying the condition (7). Motivated by this consideration, in Figure 2, under the same above-mentioned simulation conditions, a study is reported regarding the behavior of the SICS test at the varying of k between 4 and 12, fixing $l = 3$. Our analysis has been narrowed to the SICS test, and only for the value $l = 3$, after realizing that the SIS test, and other values of l, provide similar results. For graphical comprehension's sake, models (D), (E), (I), (N) and (O), for which rejection rates results to be constantly equal to 1 for any choice of k, have not been added in Figure 2. It is easy to note that, for models (A)-(C) referred to the study of the size, the choice of k does not affect the performance of the test. On the other hand, among the thirteen considered models for power analysis, (M) and (P) are the only ones showing dependence between performance of the test and value of k; in these cases, indeed, the rejection rates (simulated power) decrease as k increases.

4. THE LAG SUBSETS DEPENDENCE PLOT

Whenever the null hypothesis of serial independence is rejected, it could be useful to discover the causes. Indeed, it is well-known that choosing the correct dependence structure is an important step in time series modeling. In these terms, following the idea of Genest and Rémillard (2004), the procedures described in Section 2 lead to a naturally graphical representation of p-values as function of all the possible significant subsets of lags $S \subseteq L$, once fixed l. So, let $\mathcal{S}$ be the set containing all these $S \subseteq L$. The cardinality of $\mathcal{S}$ results: card $(\mathcal{S}) = \sum_{r=1}^{l} \binom{l}{r}$.

As for the procedures proposed in Section 2, X_t is the reference variable used for the construction of this graph. Thus, for example, fixing $l = 4$ and considering the subset $S = \{2, 4\} \subset \mathcal{S}$, the simultaneous serial dependence between X_t and X_{t-2} and between X_t and X_{t-4} is investigated. Similarly, considering the subset $S = \{1, 2, 3\} \subset \mathcal{S}$, the simultaneous serial dependence between X_t and X_{t-1}, between X_t and X_{t-2}, and between X_t and X_{t-3} is taken into account.

On the x-axis of what can be called a Lag Subsets Dependence Plot (LSDP), all the subsets S of $\mathcal{S}$ are ordered lexicographically by size, beginning with cardinality one. The corresponding p-values are then represented by vertical bars. Test statistics could be also taken into account; we have chosen to adopt p-values in order to avoid bars referred to χ^2-distributions with different numbers of degrees of freedom. Naturally, the LSDP can be provided for both tests SIS and SICS. A black point is also placed at height equal to the chosen significance level; this makes easier to identify subsets of $S \in \mathcal{S}$ leading to the rejection of the serial independence hypothesis.

Two considerations should to be given before going on. Let s be an element of S.

The first remark regards the SIS test; here, as previously said, the term ''p-value'' is considered as

$$\min_{s\in S}\{p\text{-value}_s\},$$

where p-value$_s$ is the component p-value referred to the test of independence of lag s between X_t and X_{t-s}. The second remark regards the SICS test; here, the p-value for each bar is referred to a test statistic $T = \sum_{s\in S} T_s$ in which the sum runs across all elements $s \in S$.

In order to illustrate the LSDP, for both the tests, the class of nonlinear moving average models, also considered in Section 3, has been adopted. In detail, the series here analyzed, of length $n = 1{,}000$, arise from two nonlinear moving average models of equation

$$X_t = \varepsilon_t + 0.8\varepsilon_{t-2}\varepsilon_{t-3} \quad (11)$$

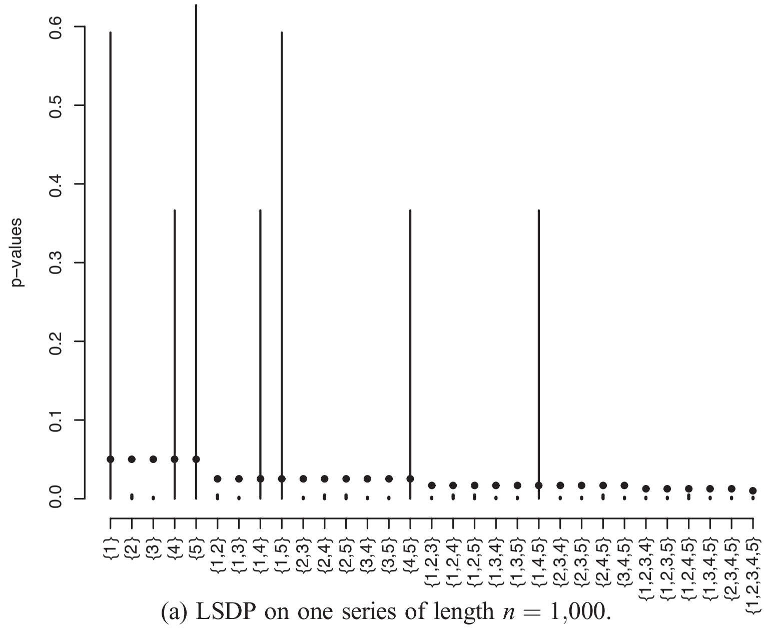
$$X_t = \varepsilon_t + 0.6\varepsilon_{t-1}^2 + 0.6\varepsilon_{t-2}^2, \quad (12)$$

with $\varepsilon_t \sim N(0,1)$, that ensure stationarity. The asymptotic nominal level $\alpha = 0.05$, and a value $k = 5$, will be considered from now on. While equation (12) coincide with model (E) in Table 4, equation (11) is similar to model (F) which is the one working worse in the previous simulations; by doing, we would confirm that the LSDP allows both a valuable visual inspection, and a formal test of serial independence, in various situations.

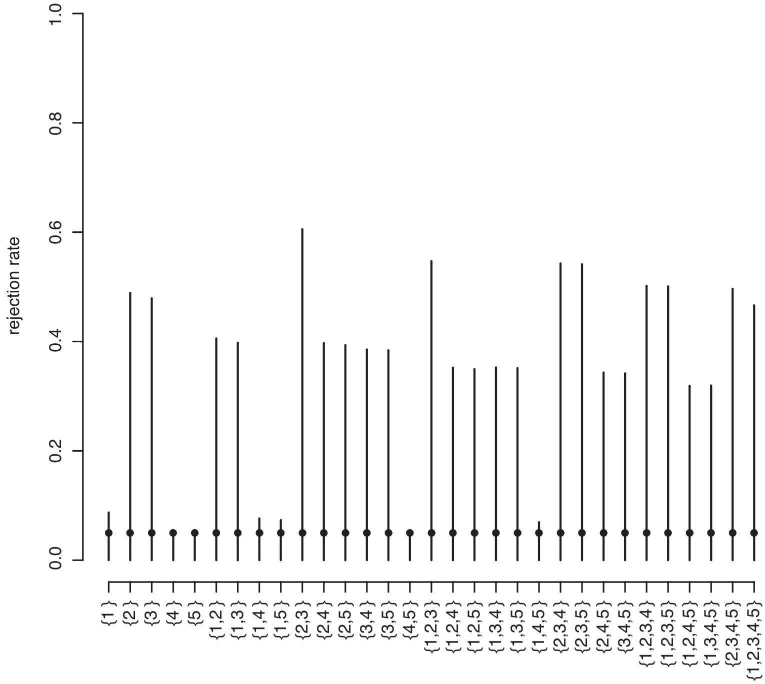
Figure 3 and Figure 4 display, respectively, the results related to SIS and SICS tests applied to model (11), fixing $l = 5$. In detail, from the LSDPs of Figure 3(a) and Figure 4(a), it is easy to note that all bars referred to sets of $\mathcal{S}$ including at least one element in $\{2,3\}$ are seen to be highly significant (p-values under the threshold), as expected from the dependence structure characterizing the adopted model. Note that the significance level for each set $S \subseteq \mathcal{S}$ is represented by a black point added on each bar. To make these results stabler, the same analysis has been replicated 10,000 times so to compute rejection rates. These rates substitute the p-values of Figure 3(a) and Figure 4(a) producing, respectively, the graphical displays in Figure 3(b) and Figure 4(b). Even in this case, the nominal level $\alpha = 0.05$ is represented by a black point on each bar. In particular, Figure 3(b) and Figure 4(b) show that:

- the higher bar is in correspondence to the set $S = \{2,3\}$;
- for the subsets of $\mathcal{S}$ that do not include at least one element of $\{2,3\}$ the corresponding bar is near the nominal level. This is true above all for the subsets $\{1\}$, $\{4\}$, $\{5\}$, $\{1,4\}$, $\{1,5\}$, $\{4,5\}$ and $\{1,4,5\}$;
- for the other subsets (which contain at least an element belonging to $\{2,3\}$) the bars are high enough to consider the power acceptable.

Similarly, Figure 5 and Figure 6 display, respectively, the results of the SIS and SICS tests applied to model (12), fixing $l = 5$. From Figure 5(a) and Figure 6(a) it is easy to note that all bars referred to sets of $\mathcal{S}$ including at least one element in $\{1,2\}$

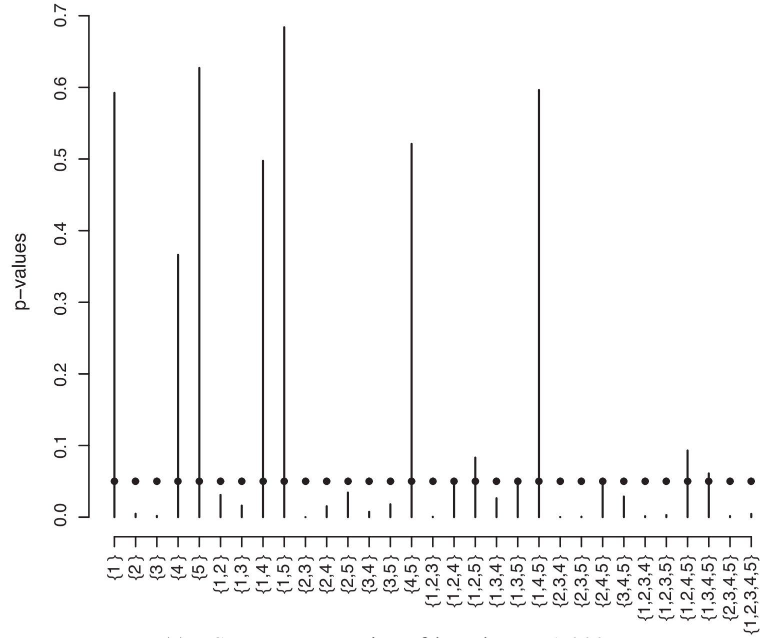


(a) LSDP on one series of length $n = 1{,}000$.

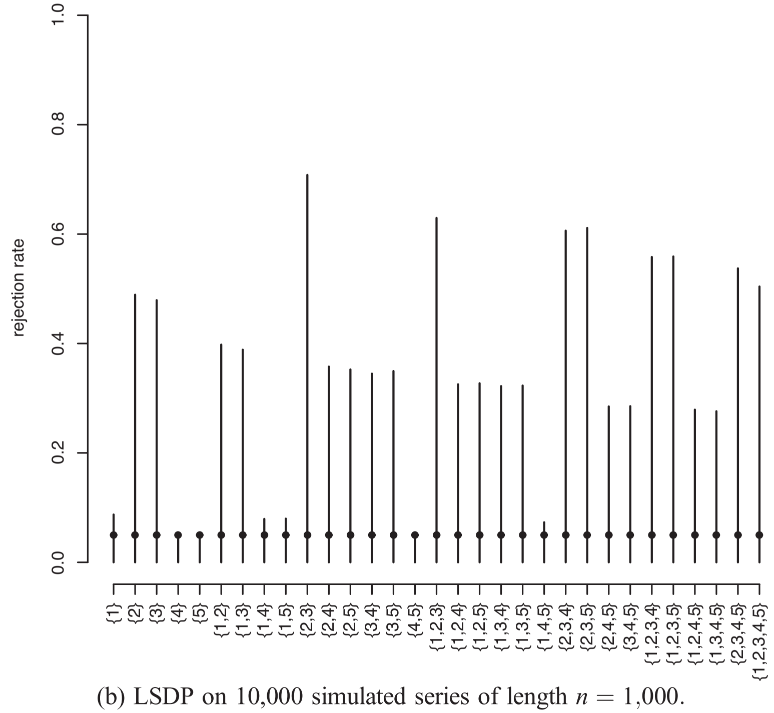


(b) LSDP on 10,000 simulated series of length $n = 1{,}000$.

FIGURE 3. - *Lag Subsets Dependence Plots, of asymptotic level* $\alpha = 0.05$*, constructed for the SIS with reference to a nonlinear moving average model* (11)

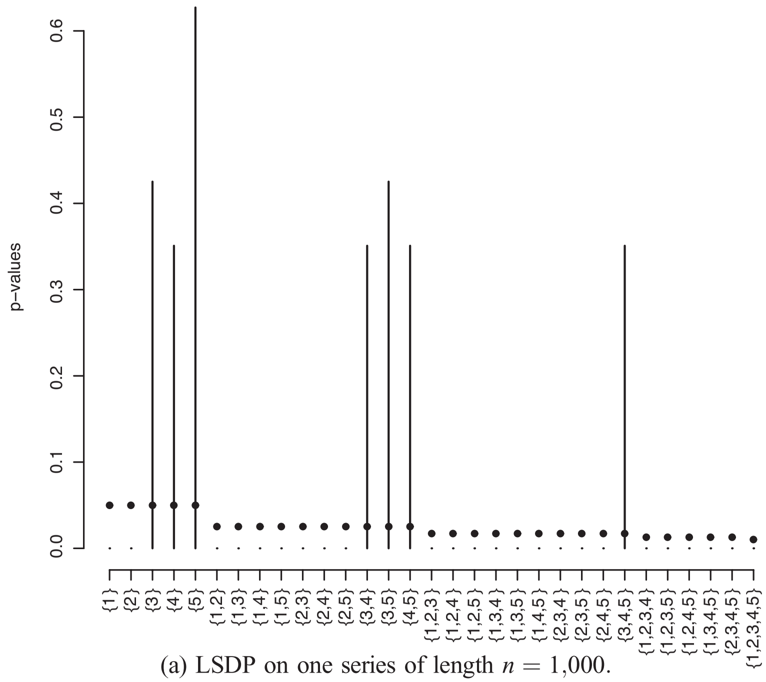


(a) LSDP on one series of length $n = 1{,}000$.

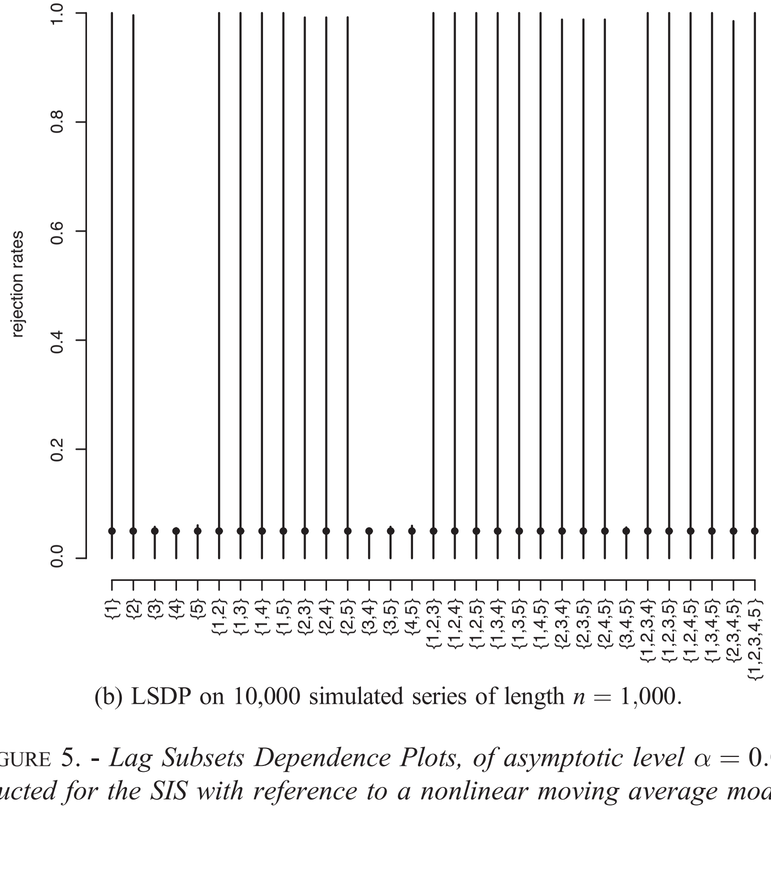


(b) LSDP on 10,000 simulated series of length $n = 1{,}000$.

FIGURE 4. - *Lag Subsets Dependence Plots, of asymptotic level* $\alpha = 0.05$, *constructed for the SICS with reference to a nonlinear moving average model* (11)

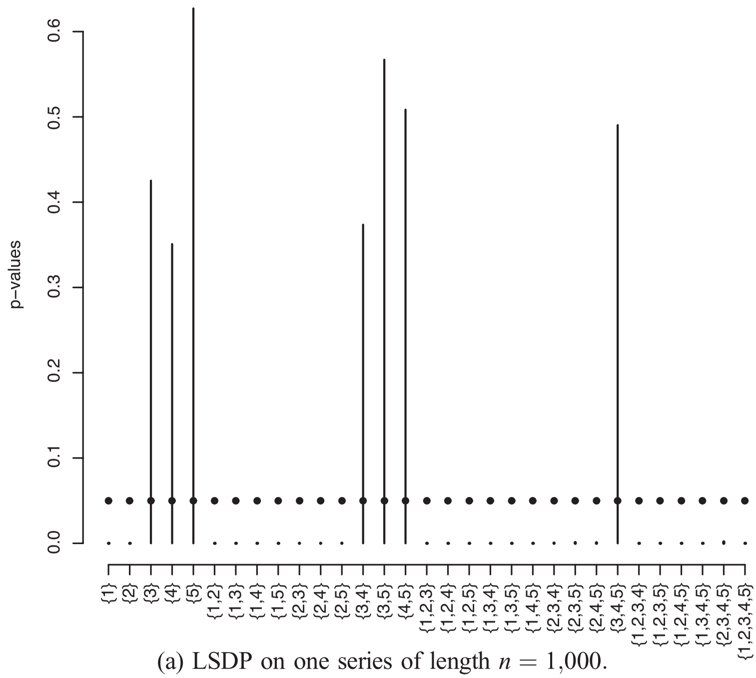


(a) LSDP on one series of length $n = 1{,}000$.

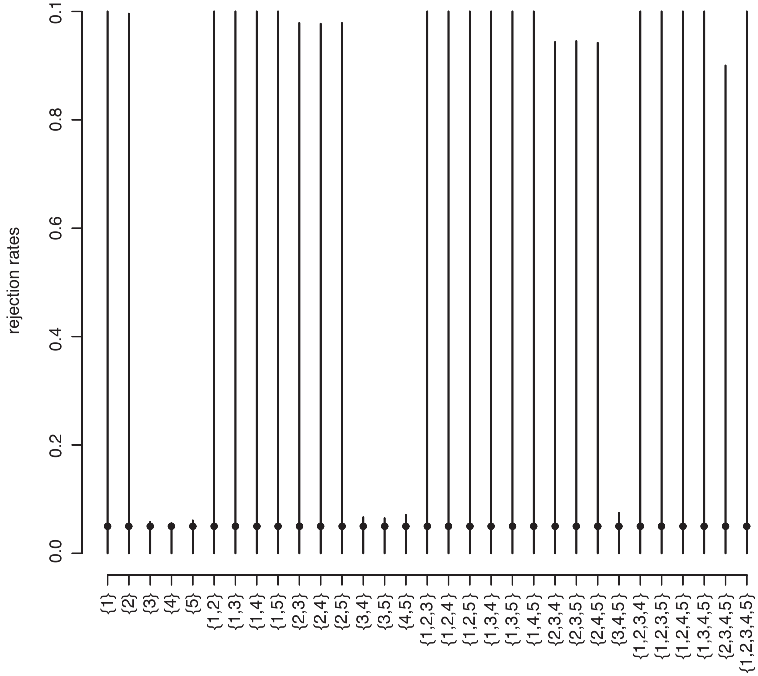


(b) LSDP on 10,000 simulated series of length $n = 1{,}000$.

FIGURE 5. - *Lag Subsets Dependence Plots, of asymptotic level* $\alpha = 0.05$, *constructed for the SIS with reference to a nonlinear moving average model* (12)



(a) LSDP on one series of length $n = 1{,}000$.



(b) LSDP on 10,000 simulated series of length $n = 1{,}000$.

FIGURE 6. - *Lag Subsets Dependence Plots, of asymptotic level* $\alpha = 0.05$, *constructed for the SICS with reference to a nonlinear moving average model* (12)

are seen to be highly significant (p-values under the threshold), as expected from the dependence structure characterizing the adopted model. Even in this case, to make these results stabler, the same analysis was replicated 10,000 times so to compute rejection rates. In particular, Figure 5(b) and Figure 6(b) show that:

- for all the subsets of $\mathcal{S}$ that do not include at least one element of $\{1, 2\}$, the corresponding bar is very close to the nominal level;
- for the other subsets (which contain at least one element belonging to $\{1, 2\}$) the bars are practically equal to one, denoting a good performance in terms of power.

These simulations, along with other trials not presented here, confirm the capability of LSDP to detect lag(s)-dependences.

5. CONCLUSIONS

In this paper, after a short introduction on how important the condition of serial independence is, above all, in the nonlinear context (both in model specification and model validation), two asymptotic methodologies to test it have been proposed. They have in common the well-known χ^2-test that is used to define single component tests to verify independence between pairs of lagged variables. The first, called SIS test (Serial Independence Simultaneous), takes advantage of the union-intersection principle while the second, denoted as SICS test (Serial Independence Chi-Square), is obtained as the sum of the component test statistics. In order to do this, the independence between component χ^2-test statistics is demonstrated and adopted. Both the tests are also used to investigate possible dependences. In particular, a graphical representation, named Lag Subsets Dependence Plot (LSDP), has been presented.

In brief, the proposed tests offer the following advantages:

- being based on the simple χ^2-test, both the procedures are easily interpretable and implementable;
- the rejection rates under null hypothesis, studied through a wide simulation study, result very close to the nominal size even if the variables are not normals. In particular, the size is also preserved with non-standard features as bimodality, heavy tails and skewness.
- simulations confirm that both the procedures result to be powerful against various types of alternatives;
- the component tests, relying on the SIS and the SICS procedures, can be used to suggest the nature of the departure from the null hypothesis of independence. In these terms, the LSDP reveals to be a useful graphical tool to detect the lag, or the set of lags, causing dependence. Obviously, this obtained information can be used for both model selection and model validation.

This work may open the door to more sophisticated studies. For example, it would be of interest to provide an automatic procedure for choosing a fixed or, more generally, a variable value of k (number of classes used in the construction of the bivariate tables for the components χ^2-test). In addition, also the problem of selecting a convenient value for l (maximum considered lag) deserves further studies and developments. Finally, it may be interesting to evaluate the performance of the proposed tests both in parametric and nonparametric model validation.

RIASSUNTO

Nel contesto delle serie storiche, soprattutto per quelle non lineari, il concetto di indipendenza seriale gioca un ruolo fondamentale. Valutare la presenza o meno di dipendenza seriale può essere infatti utile sia in fase di specificazione che di validazione del modello. Nel presente lavoro vengono proposti due test per la verifica dell'indipendenza seriale: il SIS test (Serial Independence Simultaneous) ed il SICS test (Serial Independence Chi-Square). Entrambi sfruttano il ben noto test χ^2 *per valutare l'indipendenza distributiva tra diverse coppie di variabili ritardo. Nel lavoro si fornisce inoltre uno strumento grafico, denominato LSDP (acronimo di Lag Subsets Dependence Plot), utile per individuare possibili dipendenze esistenti fra le variabili ritardo. La performance dei test proposti, in termini di ampiezza e potenza, viene valutata mediante simulazioni. Alcuni esempi evidenziano come l'LSDP permetta di individuare quali siano i ritardi che causano l'allontanamento dall'ipotesi di indipendenza seriale.*

REFERENCES

Baek E., Brock W. (1992). A nonparametric test for independence of a multivariate time series. *Statistica Sinica*, **2**(3), 137-156.

Brock W., Dechert W., Scheinkman J. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, **15**(3), 197-235.

Diks C. (2008). Nonparametric tests for independence. Technical report, University of Amsterdam.

Diks C., Panchenko V. (2007). Nonparametric tests for serial independence based on quadratic forms. *Statistica Sinica*, **17**, 81-98.

Genest C., Rémillard B. (2004). Test of independence and randomness based on the empirical copula process. *Test*, **13**(2), 335-369.

Hoeffding W. (1948). A non-parametric test of independence. *The Annals of Mathematical Statistics*, **19**(4), 546-557.

Holm S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**(2), 65-70.

Hommel G. (1983). Tests of the overall hypothesis for arbitrary dependence structures. *Biometrical Journal*, **25**(5), 423-430.

Jianqing F., Qiwei Y. (2003). *Nonlinear time series: nonparametric and parametric methods*. Springer, New York.

Johnson D., McClelland R. (1998). A general dependence test and applications. *Journal of Applied Econometrics*, **13**(6), 627-644.

McLeod A., Li W. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, **4**(4), 269-273.

Moran P.A.P. (1953). The statistical analysis of the Canadian lynx cycle, I: Structure and prediction. *Australian Journal of Zoology*, **1**, 163-173.

Pinkse J. (1998). A consistent nonparametric test for serial independence. *Journal of Econometrics*, **84**(2), 205-231.

Pinkse J. (1999). Nonparametric misspecification testing. *Technical report, University of British Columbia*.

Roy S. (1953). On a heuristic method of test construction and its use in multivariate analysis. *The Annals of Mathematical Statistics*, **24**(2), 220-238.

Seber G.A.F. (1984). *Multivariate Observations*. John Wiley & Sons, New York.

Shaffer J.P. (1995). Multiple hypothesis testing. *Annual Reviews in Psychology*, **46**(1), 561-584.

Skaug H., Tjøstheim D. (1996). Measures of distance between densities with application to testing for serial independence. In P. Robinson and M. Rosenblatt (Eds.), *Time Series Analysis in Memory of EJ Hannan* (pp. 363-377). Springer, New York.

Tjøstheim D. (1994). Non-linear time series: A selective review. *Scandinavian Journal of Statistics*, **21**(2), 97-130.

Tjøstheim D. (1996). Measures of dependence and tests of independence. *Statistics*, **28**(3), 249-284.