Journal of Statistical Software

Home | Instructions for Authors | JSS Style Guide | Editorial Board | Register

Current Volume | Archives | Search

Authors: Luca Bagnato, Lucio De Capitani, Angelo Mazza, Antonio Punzo

Title: **SDD: An R Package for Serial Dependence Diagrams**

Abstract: Detecting and measuring lag-dependencies is very important in time-series analysis. This study is commonly carried out by focusing on the linear lag-dependencies via the well-known autocorrelogram. However, in practice, there are many situations in which the autocorrelogram fails because of the nonlinear structure of the serial dependence. To cope with this problem, in this paper the R package SDD is introduced. Among the available approaches to analyze the lag-dependencies in an omnibus way, the SDD package considers the autodependogram and some of its variants. The autodependogram, defined by computing the classical Pearson χ^2-statistic at various lags, is a graphical device recently proposed in the literature to analyze lag-dependencies. The concept of reproducibility probability, and several density-based measures of divergence, are considered to define the variants of the autodependogram. An application to daily returns of the Swiss Market Index is also presented to exemplify the use of the package.

Page views:: 1091. **Submitted:** 2013-05-23. **Published:** 2015-03-20.

Paper: SDD: An R Package for Serial Dependence Diagrams   Download PDF (Downloads: 1029)

Supplements:

SDD_1.2.tar.gz: R source package   Download (Downloads: 220; 15KB)

v64c02.R: R example code from the paper   Download (Downloads: 185; 635B)

DOI: 10.18637/jss.v064.c02

ARTICLE TOOLS

Print this article

Indexing metadata

How to cite item

Supplementary files

Finding References

Email this article (Login required)

Email the author (Login required)

RELATED ITEMS

Author's work

Book searches

Databases

Math theory

Related studies

Pay-per-view

Online forums

Teaching files

Government policy

Media reports

Web search

Hide

SDD: An R Package for Serial Dependence Diagrams

Luca Bagnato
Università Cattolica di Milano

Lucio De Capitani
Università di Milano-Bicocca

Angelo Mazza
Università di Catania

Antonio Punzo
Università di Catania

Abstract

Detecting and measuring lag-dependencies is very important in time-series analysis. This study is commonly carried out by focusing on the linear lag-dependencies via the well-known autocorrelogram. However, in practice, there are many situations in which the autocorrelogram fails because of the nonlinear structure of the serial dependence.

To cope with this problem, in this paper the R package **SDD** is introduced. Among the available approaches to analyze the lag-dependencies in an *omnibus* way, the **SDD** package considers the autodependogram and some of its variants. The autodependogram, defined by computing the classical Pearson χ^2-statistic at various lags, is a graphical device recently proposed in the literature to analyze lag-dependencies. The concept of reproducibility probability, and several density-based measures of divergence, are considered to define the variants of the autodependogram. An application to daily returns of the Swiss Market Index is also presented to exemplify the use of the package.

Keywords: serial dependence, autocorrelogram, autodependogram, reproducibility probability, divergence functional.

1. Introduction

Investigating the temporal dependence structure is of fundamental importance in time-series analysis. The autocorrelogram, which measures the strength of linear dependencies (autocorrelations) as a function of the time lags, has been one of the primary tools for exploring and testing serial dependence for many decades. However, in practice, there are many situations in which the autocorrelogram fails because of the nonlinear structure of the serial dependence. A typical example is given by the GARCH model, whose components are uncorrelated but

dependent (see Diks 2009, p. 6253, for another example). Despite the fast advancement in nonlinear time series models, there are few tools which can explore the complex dependence structures in nonlinear time series, like the autocorrelogram does for the linear ones. Starting from these considerations, several diagrams have been recently proposed which are very similar, in aspect and intent, to the autocorrelogram but they are widely applicable to both linear and nonlinear time series (see Anderson and Vahid 2005, Bagnato, Punzo, and Nicolis 2012, Bagnato, De Capitani, and Punzo 2014, 2013a, and Zhou 2012).

In this paper we present the R (R Core Team 2014) package **SDD** (Bagnato, De Capitani, Mazza, and Punzo 2015) which allows for the calculation and display of some of the dependence measures discussed above. Note that, although R is well-provided with functions (such as `acf()`) and packages (such as **tseries**, Trapletti and Hornik 2013), to analyze time series (see also McLeod, Yu, and Krougly 2007, Hyndman and Khandakar 2008, Gasparrini 2011, Cowpertwait and Metcalfe 2009, and Hyndman 2014 for an overview), it does not offer graphical tools of the type described above.

The paper is organized as follows. Sections 2, 3, and 4, review the theoretical foundations of the dependence measures implemented in the **SDD** package. The relevance of these diagrams is shown, via a real data set, in Section 5, and conclusions are finally given in Section 6.

2. The autodependogram

Let $\{X_t\}_{t\in\mathbb{N}}$ represent a strictly stationary and ergodic stochastic process. Moreover, let $(X_1, \ldots, X_n)$ be an observed time series of length n from $\{X_t\}_{t\in\mathbb{N}}$. To study the generic dependence of lag r, $r < n$, let us consider the $n_r = n - r$ couples $\{(X_i, X_{i+r})\}_{i=1}^{n_r}$. Bagnato and Punzo (2010) propose to group these couples in a $k \times k$ contingency table, with marginal sets of adjacent intervals $\left\{C_u^{(r)}\right\}_{u=1}^k$ and $\left\{D_v^{(r)}\right\}_{v=1}^k$. In detail, assume that $p_{u+}^{(r)} = \mathsf{P}\left[X_t \in C_u^{(r)}\right] > 0$ for all u, $p_{+v}^{(r)} = \mathsf{P}\left[X_t \in D_v^{(r)}\right] > 0$ for all v, and $\mathsf{P}\left[X_t \in \bigcup_{u=1}^k C_u^{(r)}\right] = \mathsf{P}\left[X_t \in \bigcup_{v=1}^k D_v^{(r)}\right] = 1$. Moreover, define $p_{uv}^{(r)} = \mathsf{P}\left[(X_t, X_{t+r}) \in C_u^{(r)} \times D_v^{(r)}\right]$. The statistical hypothesis of interest is

$$H_0^r: \ p_{uv}^{(r)} = \theta_{uv}^{(r)} \qquad \forall\, (u,v) \in \{(i,j), i,j = 1, \ldots, k\}, \tag{1}$$

with $\theta_{uv}^{(r)} = p_{u+}^{(r)} p_{+v}^{(r)}$. This hypothesis can be tested using the Pearson χ^2 statistic

$$\widehat{\delta}_r = n_r \sum_{u=1}^k \sum_{v=1}^k \frac{\left(\widehat{p}_{uv}^{(r)} - \widehat{\theta}_{uv}^{(r)}\right)^2}{\widehat{\theta}_{uv}^{(r)}}, \tag{2}$$

where $\widehat{p}_{uv}^{(r)}$ and $\widehat{\theta}_{uv}^{(r)}$ denote the sample counterparts of $p_{uv}^{(r)}$ and $\theta_{uv}^{(r)}$, respectively. In particular, it results that $\widehat{p}_{uv}^{(r)} = n_{uv}^{(r)}/n_r$, $\widehat{p}_{u+}^{(r)} = n_{u+}^{(r)}/n_r$, $\widehat{p}_{+v}^{(r)} = n_{+v}^{(r)}/n_r$, and $\widehat{\theta}_{uv}^{(r)} = \widehat{p}_{u+}^{(r)}\widehat{p}_{+v}^{(r)}$, where $n_{uv}^{(r)}$ denotes the number of couples in $\{(X_i, X_{i+r})\}_{i=1}^{n_r}$ belonging to $C_u^{(r)} \times D_v^{(r)}$ and $n_{u+}^{(r)}$ $(n_{+v}^{(r)})$ denotes the number of values in $\{X_i\}_{i=1}^{n_r}$ $(\{X_{i+r}\}_{i=1}^{n_r})$ belonging to $C_u^{(r)}$ $(D_v^{(r)})$. In Bagnato *et al.* (2012) it is shown by simulations that, also in the serial context, the large sample null distribution of $\widehat{\delta}_r$ is well-approximated by the χ^2 with $(k-1)^2$ degrees of freedom. This fact allows to test the null hypothesis of independence for lag r using $\widehat{\delta}_r$ as test statistic: denoting with $\chi^2_{[\eta;q]}$ the q-quantile of the χ^2 distribution with η degrees of freedom, the null hypothesis is rejected at level α if $\widehat{\delta}_r > \chi^2_{[(k-1)^2;1-\alpha]}$.

In conformity with the *autocorrelogram* (ACF), the diagram obtained by plotting $\widehat{\delta}_r$ as a function of the time lags r, $r = 1, \dots, l$, is called *autodependogram* (ADF) by Bagnato *et al.* (2012). The autodependogram may be applied to time series with missing data and can be used for model diagnostic checking of nonlinear models (Bagnato and Punzo 2013).

To completely specify the test statistic $\widehat{\delta}_r$ it is necessary to define the partitions $\{C_u\}_{u=1}^k$ and $\{D_v\}_{v=1}^k$. Following Bagnato *et al.* (2012), the so-called *equifrequency interval*, which assigns equal frequencies to each interval C_u and to each interval D_v, will be adopted. Using the equifrequency intervals, only the value of k must be selected. As a default in the **SDD** package, k is chosen such that

$$ k = \min\{k_s, k_p\} \qquad \text{with} \qquad k_s = \left\lfloor \left(\frac{n_l}{5}\right)^{\frac{1}{2}} \right\rfloor \quad \text{and} \quad k_p = \left\lfloor 2^{\frac{11}{10}} \left(\frac{n_l - 1}{|z_{1-\alpha}|}\right)^{\frac{1}{5}} \right\rfloor, \qquad (3) $$

where $\lfloor \cdot \rfloor$ denotes the floor function while $z_{1-\alpha}$ stands for the $(1-\alpha)$-quantile of the standard normal distribution (see Bagnato *et al.* 2012 to find out more on the motivation of this rule and for simulation results confirming its validity). Note that, based on (3), k is a function of n. In particular, if $n \to \infty$, then also $k \to \infty$ (although with a lower rate of divergence) and the limiting null distribution of $\widehat{\delta}_r$, conveniently standardized, will be a standard normal (Morris 1975). As suggested by the discussion in Mann and Wald (1942), this asymptotic normality can be heuristically justified by the fact that the χ^2 distribution tends to the normal when the degrees of freedom diverge. However, for low values of k, the normal approximation of the χ^2 distribution is poor and, since rule (3) provides small values of k even when n is large (as an example, for $n = 100$ and $n = 1000$ we obtain $k = 4$ and $k = 7$, respectively), the **SSD** package always uses the χ^2 distribution.

Being k fixed and equal in each lag r, the same *level-α* critical value $\chi^2_{[(k-1)^2;1-\alpha]}$ can be used for each of the l tests of lag-independence. This allows a horizontal line at height $\chi^2_{[(k-1)^2;1-\alpha]}$ to be added to the autodependogram. Analogously with reference lines used in autocorrelograms, this line (hereafter level-α *critical line* or, simply, *critical line*) demarcates acceptance and rejection regions for each lag.

2.1. Some considerations

As documented in Bagnato *et al.* (2014), although the autodependogram is very similar in aspect to the autocorrelogram, the latter graphically represents the strength of the linear lag-dependencies through the autocorrelations, while the former displays the evidence of the presence of linear/nonlinear lag-dependencies. To clarify this distinction, it is sufficient to observe the different behavior of these diagrams when n diverges. In this case, the level-α critical lines of the autocorrelogram collapse to 0 while its bars tend toward the true values of the autocorrelations; on the other hand, the level-α critical line of the autodependogram remains fixed and the bars diverge with n under the alternative, and converge to zero under the null. Another important difference is that, the autodependogram is able to capture various and general serial dependence structures while the autocorrelogram points out only linear relationships. This feature of the autodependogram motivates the adjective *omnibus* used in Bagnato *et al.* (2012). Obviously, this generality is paid for in terms of power, and hence in the resulting descriptive ability under certain dependence structures. For example, the autodependogram will be less informative with respect to the autocorrelogram when the lag-dependencies are linear, or with respect to a representation using rank correlation statistics when the lag-dependencies are monotonic.

2.2. Normalizations

Lags may be ranked according to the evidence of their dependence, as measured by the autodependogram bars. Nevertheless, being the not normalized χ^2 statistic, these bars do not reveal clearly the strength of this evidence and do not allow comparisons between different time series. A better alternative, proposed in Bagnato *et al.* (2012, Section 6), consists in substituting $\widehat{\delta}_r$ with the Cramer contingency coefficient (C-ADF)

$$\widehat{\nu}_r = \sqrt{\frac{\widehat{\delta}_r}{n_r\,(k-1)}}, \qquad r = 1, \dots, l. \tag{4}$$

Because the denominator in (4) is equal to the maximum value that the numerator can assume, this statistic can be considered as a "normalized dependence measure" assuming values between 0 (independence) and 1 (maximum dependence). It is interesting to note that, through the application of the transformation in (4) to the critical value $\chi^2_{[(k-1)^2;1-\alpha]}$, the (different) critical values for the tests based on $\widehat{\nu}_r$ are obtained. These values allow to superimpose on the diagram of $\widehat{\nu}_r$, $r = 1, \dots, l$, a level-α critical line which, in this case, will not be horizontal but increasing in r. However, as it is well-known, $\widehat{\nu}_r$ attains values near to its maximum sporadically (see Bagnato *et al.* 2012, Section 6).

A further normalized diagram representing the evidence of the presence of dependence can be obtained substituting $\widehat{\delta}_r$ with

$$\widetilde{p}_r = 1 - \widehat{p}_r, \qquad r = 1, \dots, l, \tag{5}$$

where $\widehat{p}_r = 1 - F_{(k-1)^2}\left(\widehat{\delta}_r\right)$ denotes the p value associated with $\widehat{\delta}_r$ and F_p is the cumulative χ^2 distribution with p degrees of freedom. Also in this case a level-α critical line at height $1 - \alpha$ can be added to the resulting diagram. Nevertheless, $\widetilde{p}_r$ unbalances the two decision regions by mapping the rejection one only on $(1 - \alpha, 1]$. To avoid the last problem, it is possible to introduce a particular monotonic decreasing transformation $p_r^* = g\left(\widehat{p}_r\right)$, on $[0, 1]$, such that $g\left(\alpha\right) = 1/2$. In this way the acceptance and the rejection regions are mapped on $[0, 1/2]$ and $(1/2, 1]$, respectively. For example:

$$p_r^* = \begin{cases} \frac{2\alpha - \widehat{p}_r}{2\alpha} & \text{if } \widehat{p}_r < \alpha \\ \frac{1 - \widehat{p}_r}{2(1-\alpha)} & \text{if } \widehat{p}_r \geq \alpha \end{cases}, \qquad r = 1, \dots, l. \tag{6}$$

The transformation in (6) preserves the autodependogram's ability to detect autodependencies.

3. The reproducibility probability autodependogram

A particularly appealing alternative diagram, proposed by Bagnato *et al.* (2014), stems from the reproducibility probability (RP; Goodman 1992) representation of the χ^2-test of independence.

Under local alternatives of dependence of lag r, Bagnato *et al.* (2014) show that the large sample distribution of $\widehat{\delta}_r$ is generally well-approximated by a noncentral χ^2 with $(k-1)^2$

degrees of freedom and noncentrality parameter

$$\delta_r = n_r \sum_{u=1}^{k} \sum_{v=1}^{k} \frac{\left(p_{uv}^{(r)} - \theta_{uv}^{(r)}\right)^2}{\theta_{uv}^{(r)}}.$$

This result allows the definition of the approximated *power function*

$$\pi_{\alpha,n}(\delta_r) = 1 - F_{(k-1)^2,\delta_r}\left(\chi^2_{[(k-1)^2;1-\alpha]}\right), \tag{7}$$

where $F_{\eta,q}$ denotes the cumulative noncentral χ^2 distribution with η degrees of freedom and noncentrality parameter q. The function $\pi_{\alpha,n}(\cdot)$ associates the probability of rejecting the null hypothesis of lag-independence to each possible value of δ_r. It is strictly increasing and continuous since $F_{\eta,q}(x)$ is strictly decreasing in q for each $x > 0$ and $\eta \in \mathbb{N}$ (see Johnson, Kotz, and Balakrishnan 1995, p. 444). Then, the χ^2-test of lag-independence is strictly unbiased and it can be re-defined using the RP-testing technique, according to Martini (2008). To understand this technique (see also De Capitani 2013, for a simple introduction) it is useful to represent the test by the so called *critical function*:

$$\Psi_{\alpha}^{(r)}\left[\{X_i\}_{i=1}^{n}\right] = \begin{cases} 1 & \text{if} \quad \widehat{\delta}_r > \chi^2_{[(k-1)^2;1-\alpha]} \\ 0 & \text{if} \quad \widehat{\delta}_r \leq \chi^2_{[(k-1)^2;1-\alpha]} \end{cases}. \tag{8}$$

Representation (8) highlights that the test is a Bernoulli random variable and so the random nature of the statistical test results.

Given n, denote with δ_r^* the true but unknown value of the noncentrality parameter. The value assumed by the power function in correspondence to δ_r^*, that is $\pi_{\alpha,n}(\delta_r^*)$, is the *true power* of the test and it coincides with the only unknown parameter of the Bernoulli random variable $\Psi_{\alpha}^{(r)}$. Thus, the randomness of the test is completely described by $\pi_{\alpha,n}(\delta_r^*)$ and this makes the true power the key element to evaluate the reliability of the test result. The true power is also known as *reproducibility probability* (RP; Goodman 1992) since it is the probability of obtaining a rejection of the null hypothesis in subsequent and identical replications of the test.

From a practical point of view, the RP is unknown since δ_r^* is unknown too. The most natural way to obtain an RP estimator is to plug $\widehat{\delta}_r$ (which is an estimator of δ_r) into the power function obtaining $\widetilde{\pi}_r = \pi_{\alpha,n}\left(\widehat{\delta}_r\right)$. A further possible solution considers the estimator $\widehat{\pi}_r = \pi_{\alpha,n}\left(\widehat{\delta}_r^{\bullet}\right)$, where

$$\widehat{\delta}_r^{\bullet} \; : \; F_{(k-1)^2,\widehat{\delta}_r^{\bullet}}\left(\widehat{\delta}_r\right) = 1/2, \tag{9}$$

which satisfies $\mathsf{P}\left(\widehat{\delta}_r^{\bullet} \leq \delta_r^*\right) = 1/2$. Martini (2008) shows that $\widehat{\pi}_r$, differently from $\widetilde{\pi}_r$, allows the replication of the test (8) with the following critical function:

$$\Psi_{\alpha}\left[\{(X_i, X_{i+r})\}_{i=1}^{n_r}\right] = \begin{cases} 1 & \text{if} \quad \widehat{\pi}_r > 1/2 \\ 0 & \text{if} \quad \widehat{\pi}_r \leq 1/2 \end{cases}. \tag{10}$$

In practice: "when the probability to reject the null hypothesis is estimated to be greater than that to accept it, then the null is rejected" (Martini 2008, Remark 2). The usefulness of

the RP-based decision rule is also due to the fact that the RP estimate provides a rational indicator of the reproducibility of the test outcome. It is worth noting that, the equivalence between the RP-based decision rule and the classical one holds for all the values of α and n since their effect is implicit in the definition of the power function (Martini 2008 and De Capitani and Martini 2011).

The *RP-autodependogram* (RP-ADF) of Bagnato *et al.* (2014) is obtained by substituting $\widehat{\delta}_r$ with

$$\widehat{\pi}_r = \pi_{\alpha,n_r}\left(\widehat{\delta}_r^{\bullet}\right), \qquad r = 1,\ldots,l, \tag{11}$$

where $\widehat{\pi}_r$ can be interpreted as reproducibility probability of the level-α independence test at lag r. This additional information is very precious since it clearly quantifies the evidence of the presence of lag dependence. Moreover, for the values of α commonly used, the RP-autodependogram bars are almost normalized since the values of $\widehat{\pi}_r$ range in $[\alpha, 1]$. Thanks to the RP-testing result in (10), this diagram is also endowed with a level-α critical line, at height $1/2$, which balances the acceptance and the rejection regions by mapping them on $[\alpha, 1/2]$ and $(1/2, 1]$, respectively. Note that the sample size n_r and the level α do not influence the height of the critical line but they have a great impact on the bars of the RP-autodependogram. In more detail: the greater the value of α, the higher the bars of the RP-autodependogram (*ceteris paribus*); the greater the sample sizes n_r, the higher the bars of the RP-autodependogram (*ceteris paribus*). Note that this behavior is in agreement with the influence of n_r and α on the power of the test (Bagnato *et al.* 2014).

4. The divergence-based autopedendograms

In addition to the requirements introduced at the beginning of Section 2, suppose that X_1 has an absolutely continuous density g with support $\mathbb{R}$. Moreover, assume that (X_1, X_{1+r}) has an absolutely continuous joint density f_r with support $\mathbb{R}^2$. For a motivation about these additional assumptions, see Bagnato *et al.* (2013a). The presence of dependence for lag r can be checked by testing the statistical hypothesis

$$H_0^r: \ f_r(x,y) = g(x)g(y) \qquad \text{almost surely.} \tag{12}$$

To evaluate the discrepancy between $f_r(x,y)$ and $g(x)g(y)$, with the aim to obtain a test statistic for the testing problem (12), several divergence functionals can be considered (see, e.g., Diks 2009). By analogy with Bagnato, De Capitani, and Punzo (2013b), all the functionals considered in **SDD** have the form

$$\Delta^{(r)} = \int_{\mathbb{R}^2} D\{f_r(x,y), g(x), g(y)\} f_r(x,y)\, dxdy, \tag{13}$$

where D is a real-valued function. An example of (13) is the generalized entropy of Tsallis (1988)

$$\Delta_\gamma^{(r)} = \begin{cases} \dfrac{1}{1-\gamma} \displaystyle\int_{\mathbb{R}^2} \left[1 - \left(\dfrac{g(x)g(y)}{f_r(x,y)} \right)^{1-\gamma} \right] f_r(x,y)\, dxdy & \gamma \neq 1, \\[4mm] \displaystyle\int_{\mathbb{R}^2} \log\left(\dfrac{f_r(x,y)}{g(x)g(y)} \right) f_r(x,y)\, dxdy & \gamma = 1. \end{cases} \tag{14}$$

It is easy to note that: $\Delta_{1/2}^{(r)}$ coincides with the Hellinger metric, $\Delta_1^{(r)}$ is the Kullback-Leibler (KL) divergence, while $\Delta_2^{(r)}$ can be interpreted as the "continuous counterpart" of the Pearson χ^2 statistic adopted in Section 2. A further intuitive dependence measure is the L_1-distance:

$$\Delta_{L_1}^{(r)} = \int_{\mathbb{R}^2} |f_r(x,y) - g(x)g(y)| \, dxdy. \tag{15}$$

It can be interpreted as the "continuous counterpart" of the well-known Mortara dependence index (Mortara 1922). Similarly, in line with Rosenblatt (1975) and Skaug and Tjøstheim (1993), the following functionals can be respectively introduced:

$$\Delta_{SD}^{(r)} = \int_{\mathbb{R}^2} [f_r(x,y) - g(x)g(y)]^2 \, dxdy, \tag{16}$$

$$\Delta_{ST}^{(r)} = \int_{\mathbb{R}^2} [f_r(x,y) - g(x)g(y)] f_r(x,y) \, dxdy. \tag{17}$$

Naturally, all the aforementioned functionals are sensitive to departures from independence and, consequently, the testing problem (12) can be solved using an estimator $\widehat{\Delta}^{(r)}$ of $\Delta^{(r)}$ as test statistic with the null hypothesis rejected for large values of $\widehat{\Delta}^{(r)}$.

Once chosen Δ, the corresponding autodependogram, hereafter simply denoted as Δ-*autodependogram* (Δ-ADF), is obtained by substituting $\widehat{p}_r$ in (6) with the p value $\widehat{q}_r$ related to $\widehat{\Delta}^{(r)}$:

$$q_r^* = \begin{cases} \frac{2\alpha - \widehat{q}_r}{2\alpha} & \text{if } \widehat{q}_r < \alpha \\ \frac{1 - \widehat{q}_r}{2(1-\alpha)} & \text{if } \widehat{q}_r \geq \alpha \end{cases}, \qquad r = 1, \ldots, l. \tag{18}$$

The **SDD** package implements the Δ-ADF based on the eight functionals $\Delta_{1/2}^{(r)}$, $\Delta_1^{(r)}$, $\Delta_2^{(r)}$, $\Delta_3^{(r)}$, $\Delta_4^{(r)}$, $\Delta_{L_1}^{(r)}$, $\Delta_{SD}^{(r)}$, and $\Delta_{ST}^{(r)}$. Among them, the Δ_1-ADF proposed in Bagnato *et al.* (2013a) is considered as default. This choice follows from the results of a wide simulation study on several data generating processes presented by Bagnato *et al.* (2013b) which shows that $\Delta_1^{(r)}$ seems to be, among the eight aforementioned functionals, the best performer.

Several methodologies are proposed in the literature to implement $\widehat{\Delta}^{(r)}$; the differences among the various approaches stem from the way: (i) the densities f_r and g are estimated, (ii) the integral in (13) is computed, (iii) the p values $\widehat{q}_r$, $r = 1, \ldots, l$, are obtained. Among the alternatives considered by Bagnato *et al.* (2013b), the Gaussian kernel density estimator, an approximated numerical integration (similarly to Granger, Maasoumi, and Racine 2004), and the permutation approach, appear to be the best solutions for (i), (ii) and (iii), respectively. Details on these settings are given below.

4.1. Gaussian kernel density estimator

The Gaussian kernel (GK) estimator for the univariate density g is

$$\widehat{g}(x) = \frac{1}{n} \sum_{i=1}^n K_h(x; X_i), \tag{19}$$

where $K_h(x; X_i) = (2\pi h^2)^{-1/2} \exp\left\{ -\frac{1}{2} \left[h^{-1}(x - X_i) \right]^2 \right\}$ and $h > 0$ is the bandwidth. Simi-

larly, the GK (product) estimator for the bivariate density f_r is

$$\widehat{f}_r(x, y) = \frac{1}{n - r} \sum_{i=1}^{n-r} K_h\left(x; X_i\right) K_h\left(y; X_{i+r}\right),$$ (20)

where K_h is defined as in (19).

To apply the GK density estimators (19) and (20), a value for the bandwidth h needs to be chosen. With this aim, in the literature, several data dependent procedures have been proposed. Examples are the Silverman's rule of thumb and the likelihood cross-validation method (see Silverman 1986, p. 45 and Section 3.4.4, respectively). The bandwidth obtained with the latter procedure, denoted with h_{LCV}, is particularly useful in this context because, as observed in Granger *et al.* (2004, p. 654), it produces optimal density estimators according to the KL-criterion. In detail, working on the estimator in (19), h_{LCV} is obtained by minimizing the cross-validation statistic

$$\mathrm{CV}\left(h\right) = n^{-1} \sum_{i=1}^{n} \ln\left[\widehat{g}_{-i}\left(x\right)\right],$$

where

$$\widehat{g}_{-i}\left(x\right) = (n - 1)^{-1} \sum_{j \neq i} K_h\left(x; X_j\right)$$

is the estimated density constructed from all the data points except X_i. The obtained value of h_{LCV} is then used also in (20).

However, it is well-known that a bandwidth which is optimal for estimation is usually sub-optimal for testing. In particular, although h_{LCV} suffices to establish consistency of the test statistics, this choice could not be optimal in terms of the power of the resulting tests. As observed by Anderson, Hall, and Titterington (1994), in testing procedures a relative over-smoothing may be appropriate for some dependence functionals (test statistics). Nevertheless, the simulation results of Bagnato *et al.* (2013b) highlight that when the GK density estimator is adopted to define the estimator of $\Delta^{(r)}$, the use of h_{LCV} is appropriate and, then, no oversmoothing is applied in **SDD**.

4.2. Estimation of the dependence functional

Simulation results in Bagnato *et al.* (2013b) indicate that a good solution for the estimation of $\Delta^{(r)}$ is

$$\widehat{\Delta}^{(r)} = 10^{-4} \sum_{i=1}^{100} \sum_{j=1}^{100} D\left\{\widehat{f}_r\left(\tilde{x}_i, \tilde{y}_j\right), \widehat{g}\left(\tilde{x}_i\right), \widehat{g}\left(\tilde{y}_j\right)\right\} \widehat{f}_r\left(\tilde{x}_i, \tilde{y}_j\right),$$ (21)

where $\widehat{f}_r$ and $\widehat{g}$ are the GK estimates of f and g, and where, following the default setting of the **sm** package (Bowman and Azzalini 2014),

$$\tilde{x}_i = \left(x_{(1)} - a\right) + (i - 1)\frac{x_{(n)} - x_{(1)} + 2a}{99},$$

with $a = \left(x_{(n)} - x_{(1)}\right)/4$ and $x_{(1)}$ $\left(x_{(n)}\right)$ denoting the minimum (maximum) observed value. The values $\tilde{y}_j$ are defined in the same way. Note that (21) is an approximation of

$$\widehat{\Delta}^{(r)} = \int_{\mathbb{R}^2} D\left\{\widehat{f}_r\left(x, y\right), \widehat{g}\left(x\right), \widehat{g}\left(y\right)\right\} \widehat{f}_r\left(x, y\right) dx dy$$ (22)

based on the 100×100 grid of equally spaced values $\{(\tilde{x}_i, \tilde{y}_j) : i, j = 1, \ldots, 100\}$.

4.3. Computing p values

Among the various proposals to compute the p value $\widehat{q}_r$, the simulation studies of Bagnato *et al.* (2013b,a) suggest that a permutation approach represents a good compromise between simplicity and performance. It exploits the fact that, conditionally on the observed data $x_1, \ldots, x_n$, each of the possible $n!$ permutations is equally likely under the assumption of serial independence (see Diks 2009). In detail, let $\widehat{\Delta}^{(r,0)}$ denote the value assumed by $\widehat{\Delta}^{(r)}$ for the observed data. Analogously, let $\widehat{\Delta}^{(r,b)}$ be the dependence functional estimate obtained from a random permutation of the original data, with $b = 1, \ldots, B$. Under the assumption of serial independence, $\widehat{\Delta}^{(r,1)}, \ldots, \widehat{\Delta}^{(r,B)}$ are equally likely and the p value can be defined as in Diks and Panchenko (2007):

$$\widehat{q}_r = \frac{\# \left\{ \widehat{\Delta}^{(r,s)} : \widehat{\Delta}^{(r,s)} > \widehat{\Delta}^{(r,0)} ; s = 0, 1, \ldots, B \right\} + L}{B + 1}, \qquad r = 1, \ldots, l, \qquad (23)$$

where L is defined as follows. Let $Z = \# \left\{ \widehat{\Delta}^{(r,s)} : \widehat{\Delta}^{(r,s)} = \widehat{\Delta}^{(r,0)} ; s = 0, 1, \ldots, B \right\} \geq 1$ denote the number of ties plus one. If $Z = 1$ then $L = 1$, while if $Z > 1$ then L is drawn from the discrete uniform distribution on $\{1, \ldots, Z\}$. The above procedure for computing the p value leads to a randomized test having an exact level α if the null is rejected whenever $\widehat{q}_r \leq \alpha$ and $0 < \alpha = c/(B + 1) < 1$ for some integer c.

4.4. Testing independence on more than one lag simultaneously

The **SDD** package also allows to test independence on a set of lags, say $\mathcal{R}$, specified by the user. Such a procedure is typically used to test the more general hypothesis of serial independence (cf. Diks 2009). Among the available procedures (see Bagnato *et al.* 2013b, for some examples), the **SDD** package handles the "multiple-lag testing problem"

$$H_0^{\mathcal{R}} : \; f_r(x, y) = g(x) g(y) \qquad \text{almost surely } \forall \, r \in \mathcal{R}$$

in two simple ways.

A first way, the most common, consists in building a "Portmanteau test" (P-test); when the Δ-ADF is considered, the test statistic is $\widehat{Q}_{\mathcal{R}} = \sum_{r \in \mathcal{R}} \widehat{\Delta}^{(r)}$. In this case, recalling (23), the p value is computed as

$$\widehat{q}_{\mathcal{R}} = \frac{\# \left\{ \widehat{Q}_{\mathcal{R}}^{(s)} : \widehat{Q}_{\mathcal{R}}^{(s)} > \widehat{Q}_{\mathcal{R}}^{(0)} ; s = 0, 1, \ldots, B \right\} + L}{B + 1}, \qquad (24)$$

where $\widehat{Q}_{\mathcal{R}}^{(s)} = \sum_{r \in \mathcal{R}} \widehat{\Delta}^{(r,s)}$. If the user adopts $\widehat{\delta}_r$ instead of $\widehat{\Delta}^{(r)}$ (similar reasoning holds for the C-ADF), the asymptotic distribution of $\widehat{Q}_{\mathcal{R}} = \sum_{r \in \mathcal{R}} \widehat{\delta}_r$, used to compute the p value, is a χ^2 with $(k-1)^2 \operatorname{card}(\mathcal{R})$ degrees of freedom (see Bagnato and Punzo 2010, 2012, for details), where $\operatorname{card}(\mathcal{R})$ denotes the cardinality of $\mathcal{R}$.

A second simpler way consists in using a "Simultaneous test" (S-test); it is based on the adjustment of the p values $\widehat{q}_r$, $r \in \mathcal{R}$, in order to take a decision about the simultaneous independence for the lags in $\mathcal{R}$. A lot of methods exist for adjustment (see, e.g., Wright

1992); the package **SDD** implements all the methods allowed by the function `p.adjust()` of the **stats** package. As an example, the famous Bonferroni adjustment leads to the following decision rule: "do not reject $H_0^{\mathcal{R}}$ if all the single tests, for lags belonging to $\mathcal{R}$, do not reject at level $\alpha/\mathrm{card}\,(\mathcal{R})$".

As can be inferred from the simulations of Bagnato *et al.* (2013b), the P-test seems to be a better choice. A further alternative, whose performance is comparable to the P-test, would be the use of the multiple-lag procedure described in Bagnato *et al.* (2013b). However, because the computational burden required is substantial, this technique is not considered here.

5. Package description and illustrative example

Package **SDD** is developed in an object-oriented design, using the standard S3 paradigm. Its main function, `ADF()`, allows for the computation and the plot of all the different types of serial dependence diagrams illustrated in the paper and it returns an object of class 'SDD'. Its arguments, along with their description, are listed in Table 1.

To illustrate the use of the package, the `SMI` dataset included in the **SDD** package and already analyzed in Bagnato *et al.* (2014) and Bagnato and Punzo (2013) is considered. Data consist of $n = 660$ daily returns of the Swiss Market Index spanning the period from 2009-08-12 to 2012-03-06 (the share prices used to compute the daily returns are downloadable from http://finance.yahoo.com/).

To begin the analysis, data are loaded by

```
R> library("SDD")
R> data("SMI", package = "SDD")
```

In the financial context, a first glance to the serial dependence can be obtained by considering the autocorrelogram of the squared raw series. The corresponding plot is obtained via the command

```
R> res1 <- ADF(SMI^2, dtype = "ACF", main = "")
```

It produces the plot in Figure 1(a), which shows linear dependencies, on the squared raw data, for the majority of the considered lags. The function `acf()`, of the **stats** package, is used and the autocorrelogram is displayed without the first bar referred to lag zero. The presence of serial dependence on the squared raw data is also corroborated by the autocorrelation-based tests, whose results are printed via the `print()` method

```
R> print(res1)


 ACF bars for series 'SMI^2'
     1       2       3       4       5       6       7       8       9      10      11
0.3739 0.3132 0.1842 0.3274 0.2460 0.2329 0.1606 0.1988 0.0948 0.1064 0.1265
    12      13      14      15      16      17      18      19      20      21      22
0.2346 0.1566 0.0655 0.1125 0.1651 0.2498 0.1912 0.1470 0.0641 0.1472 0.1774
    23      24      25      26      27      28
0.1344 0.1107 0.1240 0.0132 0.0522 0.0537
```

Arguments	Description
x	Univariate time series $(x_1, \ldots, x_n)$.
dtype	Autodependence function to be computed: `"ACF"` for the ACF; `"ADF"` (default) for the ADF; `"CADF"` for the C-ADF; `"RPADF"` for the RP-ADF; `"DeltaADF"` for the Δ-ADFs.
lag.max	Maximum lag l considered.
alpha	Significance level α for the critical line.
num.clas	Value of k in the contingency table used for the χ^2-based ADFs. If not specified, it is determined internally using the rule of thumb (3).
bandwidth	Bandwidth h in the GK density estimator. If not specified, h_{LCV} is used.
B	Number of permutations B used to compute p values for the Δ-ADF.
delta	Divergence functional considered in the Δ-ADF as defined in equations from (14) to (17); options are: `"Delta_1"` (default), `"Delta_0.5"`, `"Delta_2"`, `"Delta_3"`, `"Delta_4"`, `"Delta_SD"`, `"Delta_L1"`, `"Delta_ST"`, and `"Delta_fdiv"` when the user wants to provide an alternative external divergence function (see `fdiv`).
fres	Alternative external function specifying the resampling method from the raw series; it is used to compute p values.
fdenest	Alternative external function to perform estimation of $g(x)$ and $f_r(x, y)$.
fdiv	Alternative external divergence function used when `delta = "Delta_fdiv"`.
plot	If `TRUE` (default), the diagram specified by `dtype` (and eventually by `delta`) is plotted.
R	The set $\mathcal{R}$ of Section 4.4.
p.adjust.method	Adjustment method in the simultaneous test. It must be one of `p.adjust.methods` (see the **stats** package).

Table 1: List of arguments for the function `ADF()`.

```
Simultaneous Test
  adjustment method: holm
  adjusted p-value: 0

Portmanteau Test (Box-Pierce test)
  p-value: 0

Tested lags
  1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28
```

(a) ACF on squared raw data
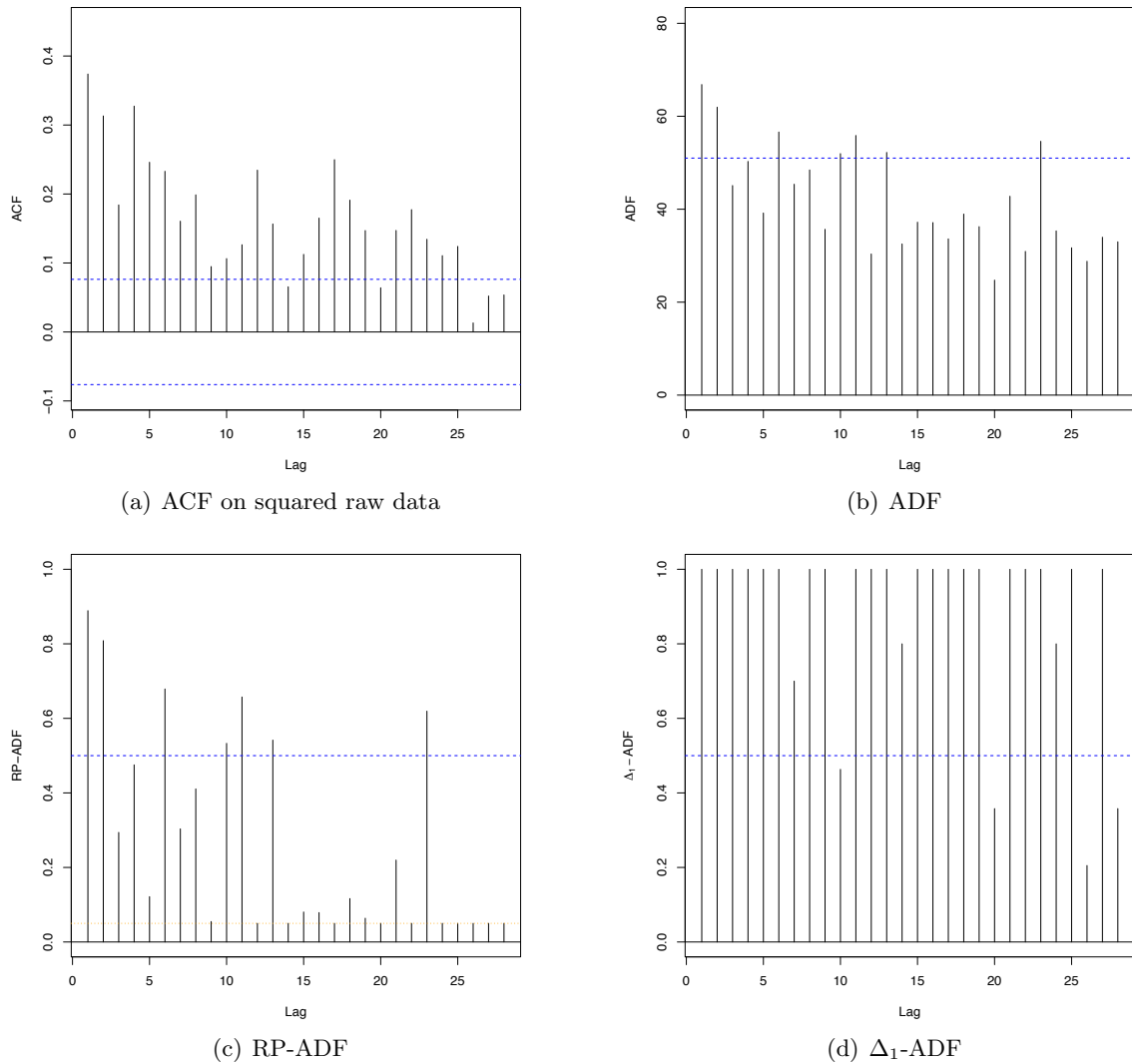

(b) ADF

(c) RP-ADF

(d) Δ_1-ADF

Figure 1: Dependence diagrams on the SMI data.

With both the simultaneous and Portmanteau tests, the null hypothesis $H_0^{\mathcal{R}}$, with $\mathcal{R} = \{1, \ldots, 28\}$, is rejected with any $\alpha > 0$.

Instead of considering directional (specific) types of lag-dependencies, we can act in an *omnibus* way by displaying the autodependogram variants. The following command

```
R> res2 <- ADF(SMI, main = "")
```

produces the ADF in Figure 1(b); its first two bars show evidence in favor of the alternative hypothesis of lag-dependence. As before, to check serial independence, we can run the command

```
R> print(res2)
```

```
  ADF bars for series 'SMI'
   1    2    3    4    5    6    7    8    9   10   11   12   13   14   15
66.8 62.0 45.1 50.3 39.2 56.6 45.4 48.5 35.7 52.0 55.9 30.4 52.2 32.5 37.2
  16   17   18   19   20   21   22   23   24   25   26   27   28
37.1 33.6 39.0 36.3 24.7 42.8 30.9 54.6 35.3 31.7 28.8 34.0 33.0

Simultaneous Test
  adjustment method: holm
  adjusted p-value: 0.038

Portmanteau Test
  p-value: 0

Tested lags
  1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28
```

Also in this case, serial independence is rejected at any significance level by the Portmanteau test and at levels greater than 0.038 by the simultaneous test.

The RP-ADF and the Δ_1-ADF, shown respectively in Figure 1(c) and Figure 1(d), are obtained with the commands

```
R> res3 <- ADF(SMI, dtype = "RPADF", main = "")
R> set.seed(1)
R> ADF(SMI, dtype = "DeltaADF", main = "")
```

In particular, the RP-ADF allows for additional interpretation; by focusing on the first two bars $\widehat{\pi}_1$ and $\widehat{\pi}_2$, using the following command

```
R> res3$res$vbar[1:2]
```

```
[1] 0.8890191 0.8085375
```

we can conclude that the null hypothesis of lag-1 (lag-2) independence is rejected and we estimate that this hypothesis will be rejected with probability 0.88901 (0.80853) if the same test is performed on a different time series of the same length from the same generating process. On the other hand, the Δ_1-ADF in Figure 1(d) detects more lag-dependencies. From this point of view, it seems more similar to the ACF on the squared residuals. As stated in Bagnato *et al.* (2013a), it should be due to the higher power of the Δ_1-ADF among the implemented autodependograms.

In order to capture the underlying dependence structure, in line with the common practice for financial time series (see, e.g., Bollerslev, Chou, and Kroner 1992), we consider a GARCH model. In particular, a GARCH$(1, 1)$ with Gaussian innovations is estimated, with the maximum likelihood approach, using the `garch()` function of the R package **tseries**, and the residuals are obtained as

```
R> library("tseries")
R> residuals <- garch(SMI, order = c(1, 1))$residuals[-1]
```
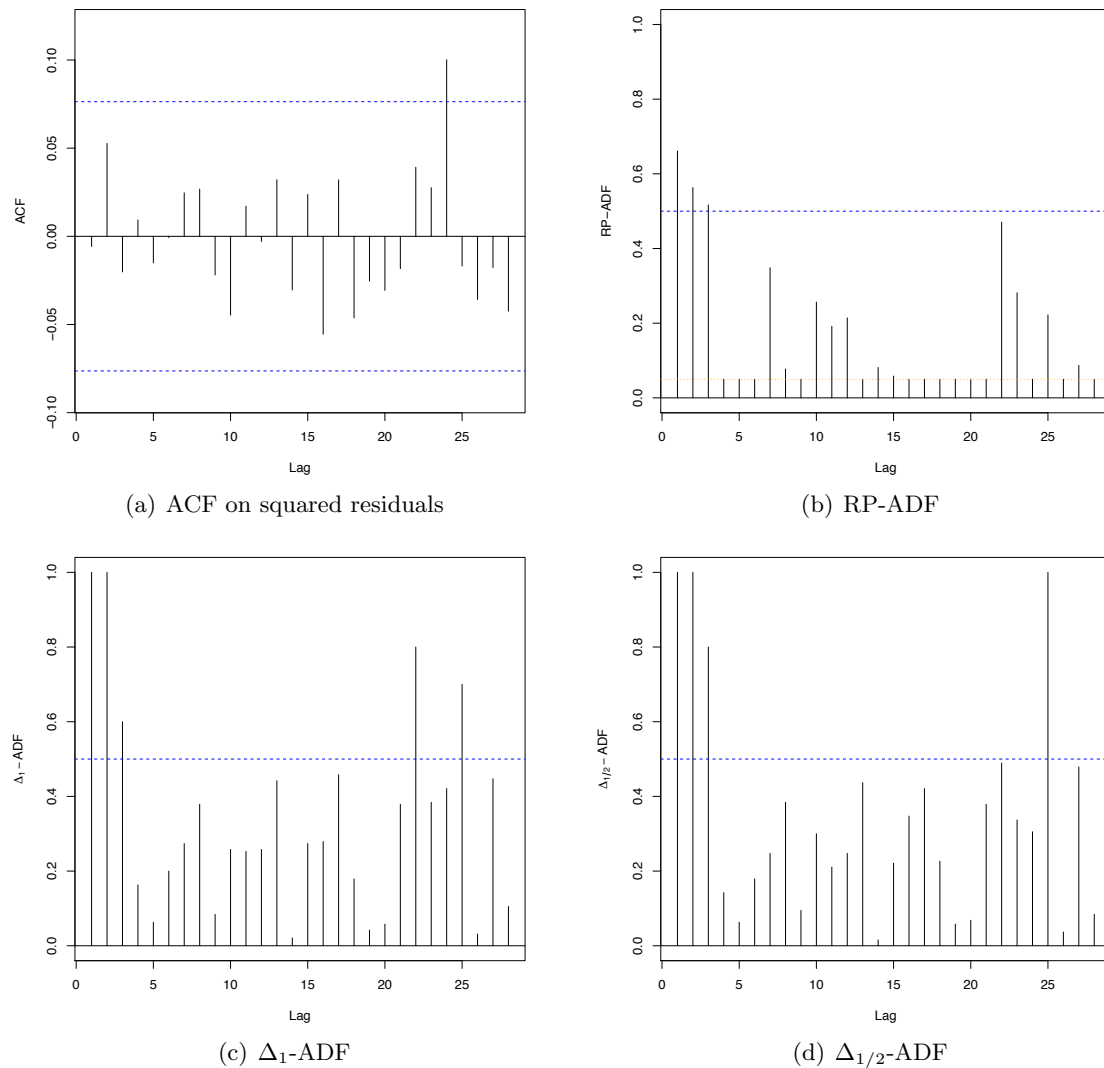
(a) ACF on squared residuals

(b) RP-ADF

(c) Δ_1-ADF

(d) $\Delta_{1/2}$-ADF

Figure 2:   Dependence diagrams on the residuals obtained fitting the GARCH$(1, 1)$ on the `SMI` data.

Figure 2 shows the ACF on the squared residuals, the RP-ADF, the Δ_1-ADF, and the $\Delta_{1/2}$-ADF on the residuals. These plots are obtained via the commands

```
R> res4 <- ADF(residuals^2, dtype = "ACF", main = "")
R> res5 <- ADF(residuals, dtype = "RPADF", main = "")
R> set.seed(1)
R> ADF(residuals, dtype = "DeltaADF", main = "")
R> set.seed(1)
R> ADF(residuals, dtype = "DeltaADF", delta = "Delta_0.5", main = "")
```

As it can be seen from Figure 2(a), the autocorrelogram does not display significant linear dependencies among the squared residuals. Serial independence is also confirmed by the simultaneous and the Portmanteau tests obtained via the command

```
R> print(res4)

  ACF bars for series 'residuals^2'
         1          2          3          4          5          6          7
 -0.005790   0.052679  -0.020194   0.009306  -0.015073  -0.000694   0.024682
         8          9         10         11         12         13         14
  0.026763  -0.021902  -0.044663   0.017074  -0.002904   0.032077  -0.030425
        15         16         17         18         19         20         21
  0.023796  -0.055524   0.032010  -0.046304  -0.025401  -0.030701  -0.018324
        22         23         24         25         26         27         28
  0.039210   0.027600   0.100133  -0.016851  -0.035821  -0.017693  -0.042427

Simultaneous Test
  adjustment method: holm
  adjusted p-value: 0.284

Portmanteau Test (Box-Pierce test)
  p-value: 0.752

Tested lags
  1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28
```

On the contrary, the RP-ADF in Figure 2(b) highlights that a dependence structure involving the first three lags still exists. We can evaluate the entity of this decaying memory dependence via the commands

```
R> res5$res$vbar[1:3]

[1] 0.6611932 0.5633196 0.5167406
```

The other two diagrams in Figure 2(c) and Figure 2(d) detect such lag-dependencies too.

6. Conclusions

In this paper we have presented the **SDD** package for the R environment. It provides several diagrams to analyze linear/nonlinear lag-dependencies for time series. Differently from other implemented and existing diagrams, like the autocorrelogram, this analysis is carried out in an *omnibus* way, without focusing on specific (directional) forms of lag-dependence. In other words, we can detect lag-dependencies that are more general than, for example, the linear or monotonic ones. The package also includes an illustrative data set on daily returns of the Swiss Market Index (SMI). Through the application of variants of the autodependogram to this data set, we have presented the usefulness of the **SDD** package for detecting and evaluating the lag-dependencies present in both raw data and residuals of a fitted model. We believe the availability of such diagnostic checks will be appreciated by other R users as well. Extensions to either "partial" or "cross" variants of the considered diagrams will be implemented in future versions of the package.

References

Anderson HM, Vahid F (2005). "Nonlinear Correlograms and Partial Autocorrelograms." *Oxford Bulletin of Economics and Statistics*, **67**(Supplement s1), 957–982.

Anderson NH, Hall P, Titterington DM (1994). "Two-Sample Test Statistics for Measuring Discrepancies between Two Multivariate Probability Density Functions Using Kernel-Based Density Estimates." *Journal of Multivariate Analysis*, **50**(1), 41–54.

Bagnato L, De Capitani L, Mazza A, Punzo A (2015). **SDD***: Serial Dependence Diagrams*. R package version 1.2, URL http://CRAN.R-project.org/package=SDD.

Bagnato L, De Capitani L, Punzo A (2013a). "Improving the Autodependogram Using the Kullback-Leibler Divergence." arXiv:1306.5006 [stat.ME], URL http://arxiv.org/pdf/1306.5006v1.pdf.

Bagnato L, De Capitani L, Punzo A (2013b). "Testing Serial Independence via Density-Based Measures of Divergence." *Methodology and Computing in Applied Probability*, **16**(3), 627–641.

Bagnato L, De Capitani L, Punzo A (2014). "Detecting Serial Dependencies with the Reproducibility Probability Autodependogram." *Advances in Statistical Analysis*, **98**(1), 35–61.

Bagnato L, Punzo A (2010). "On the Use of χ^2-Test to Check Serial Independence." *Statistica & Applicazioni*, **VIII**(1), 57–74.

Bagnato L, Punzo A (2012). "Checking Serial Independence of Residuals from a Nonlinear Model." In W Gaul, A Geyer-Shulz, L Schmidt-Thieme, J Kunze (eds.), *Challenges at the Interface of Data Analysis, Computer Science, and Optimization*, volume XIV of *Studies in Classification, Data Analysis and Knowledge Organization*, pp. 203–211. Springer-Verlag, Berlin Heidelberg.

Bagnato L, Punzo A (2013). "Using the Autodependogram in Model Diagnostic Checking." In N Torelli, F Pesarin, A Bar-Hen (eds.), *Advances in Theoretical and Applied Statistics*, volume XIX of *Studies in Theoretical and Applied Statistics*, pp. 129–139. Springer-Verlag, Berlin Heidelberg.

Bagnato L, Punzo A, Nicolis O (2012). "The Autodependogram: A Graphical Device to Investigate Serial Dependences." *Journal of Time Series Analysis*, **33**(2), 233–254.

Bollerslev T, Chou RY, Kroner KF (1992). "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence." *Journal of Econometrics*, **52**(1), 5–59.

Bowman A, Azzalini A (2014). **sm***: Smoothing Methods for Nonparametric Regression and Density Estimation*. Version 2.2-5.4, URL http://CRAN.R-project.org/package=sm.

Cowpertwait PSP, Metcalfe AV (2009). *Introductory Time Series with* R. Springer-Verlag, Dordrecht.

De Capitani L (2013). "An Introduction to RP-Testing." *Epidemiology, Biostatistics, and Public Health*, **10**(1), 1–16.

De Capitani L, Martini DD (2011). "On Stochastic Orderings of the Wilcoxon Rank Sum Test Statistic – With Applications to Reproducibility Probability Estimation Testing." *Statistics & Probability Letters*, **81**(8), 937–946.

Diks C (2009). "Nonparametric Tests for Independence." In RA Meyers (ed.), *Encyclopedia of Complexity and Systems Science*, pp. 6252–6271. Springer-Verlag, New York.

Diks C, Panchenko V (2007). "Nonparametric Tests for Serial Independence Based on Quadratic Forms." *Statistica Sinica*, **17**(1), 81–89.

Gasparrini A (2011). "Distributed Lag Linear and Non-Linear Models in R: The Package **dlnm**." *Journal of Statistical Software*, **43**(8), 1–20. URL http://www.jstatsoft.org/v43/i08/.

Goodman SN (1992). "A Comment on Replication, p-Values and Evidence." *Statistics in Medicine*, **11**(7), 875–879.

Granger CW, Maasoumi E, Racine J (2004). "A Dependence Metric for Possibly Nonlinear Processes." *Journal of Time Series Analysis*, **25**(5), 649–669.

Hyndman RJ (2014). "CRAN Task View: Time Series Analysis." Version 2014-12-23, URL http://CRAN.R-project.org/view=TimeSeries.

Hyndman RJ, Khandakar Y (2008). "Automatic Time Series Forecasting: The **forecast** Package for R." *Journal of Statistical Software*, **27**(3), 1–22. URL http://www.jstatsoft.org/v27/i03/.

Johnson N, Kotz S, Balakrishnan N (1995). *Continuous Univariate Distributions*, volume 2. John Wiley & Sons, New York.

Mann HB, Wald A (1942). "On the Choice of the Number of Intervals in the Application of the Chi-Square Test." *The Annals of Mathematical Statistics*, **13**(3), 306–317.

Martini DD (2008). "Reproducibility Probability Estimation for Testing Statistical Hypotheses." *Statistics & Probability Letters*, **78**(9), 1056–1061.

McLeod AI, Yu H, Krougly ZL (2007). "Algorithms for Linear Time Series Analysis: With R Package." *Journal of Statistical Software*, **23**(5), 1–26. URL http://www.jstatsoft.org/v23/i05/.

Morris C (1975). "Central Limit Theorems for Multinomial Sums." *The Annals of Statistics*, **3**(1), 165–188.

Mortara G (1922). *Lezioni di Statistica*. Società Tipografica "Leonardo da Vinci", Città di Castello.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rosenblatt M (1975). "A Quadratic Measure of Deviation of Two-Dimensional Density Estimates and a Test of Independence." *The Annals of Statistics*, **3**(1), 1–14.

Silverman BW (1986). *Density Estimation for Statistics and Data Analysis.* Monographs on Statistics and Applied Probability. Chapman & Hall, London.

Skaug HJ, Tjøstheim D (1993). "Nonparametric Tests for Serial Independence." In TS Rao (ed.), *Developments in Time Series Analysis: The Priestley Birthday Volume*, pp. 207–229. Chapman & Hall, London.

Trapletti A, Hornik K (2013). **tseries**: *Time Series Analysis and Computational Finance.* R package version 0.10-32, URL http://CRAN.R-project.org/package=tseries.

Tsallis C (1988). "Possible Generalization of Boltzmann-Gibbs Statistics." *Journal of Statistical Physics*, **52**(1), 479–487.

Wright SP (1992). "Adjusted *p*-Values for Simultaneous Inference." *Biometrics*, **48**(4), 1005–1013.

Zhou Z (2012). "Measuring Nonlinear Dependence in Time-Series, a Distance Correlation Approach." *Journal of Time Series Analysis*, **33**(3), 438–457.

Affiliation:

Luca Bagnato
Dipartimento di Scienze Matematiche, Finanza Matematica e Econometria
Università Cattolica del Sacro Cuore
Via Lodovico Necchi, 9, 20123 Milano, Italia
Telephone: +39/02/72342466
Fax: +39/02/72342324
E-mail: luca.bagnato@unicatt.it
URL: http://docenti.unicatt.it/ita/luca_bagnato/

Lucio De Capitani
Dipartimento di Statistica e Metodi Quantitativi
Università degli Studi di Milano-Bicocca
Via Bicocca degli Arcimboldi, 8, 20126 Milano, Italia
Telephone: +39/02/64483186
Fax: +39/02/64483105
E-mail: lucio.decapitani1@unimib.it
URL: http://www.economia.unimib.it/DECAPITANI/

Angelo Mazza, Antonio Punzo
Dipartimento di Economia e Impresa
Università di Catania

Corso Italia, 55, 95129 Catania, Italia
Telephone: +39/095/7537736, +39/095/7537640
Fax: +39/095/7537610
E-mail: a.mazza@unict.it, antonio.punzo@unict.it
URL: http://docenti.unict.it/a.mazza/, http://www.economia.unict.it/punzo/

"SDD: An R package for serial dependence diagrams"

SUDDIVIONE DEI CONTRIBUTI.
Il presente articolo è il frutto della stretta collaborazione tra gli autori L. Bagnato, L. De Capitani, A. Mazza e A. Punzo che hanno contribuito in egual misura alla stesura dello stesso.

25/07/2016

Luca Bagnato

Antonio Punzo

Angelo Mazza

Lucio De Capitani