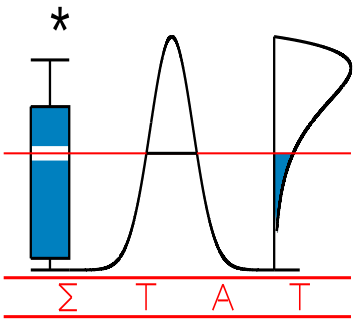
T E C H N I C A L

R E P O R T

0461

ON THE DISTRIBUTION OF THE SAMPLE AUTOCORRELATIONS OF A WHITE NOISE PROCESS

AYADI, A. and G. MELARD



I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

ON THE DISTRIBUTION OF THE SAMPLE AUTOCORRELATIONS OF A WHITE NOISE PROCESS

Abdelghafour AYADI* and Guy MÉLARD†

December 15, 2004

Abstract

For testing a time series of length n for randomness, the sample lag-k autocorrelations $r_k = c_k/c_0$, $k = 1, 2, ...$are generally considered, where the c_k are the sample autocovariances. Under the null hypothesis that the series is generated by a Gaussian white noise process, the large sample distribution of the statistic is $r_k \sim N(0, 1/n)$. Therefore, it is a standard to look for significant autocorrelations at the 5% level, i.e. being larger than $1.96/\sqrt{n}$ in absolute value. The objective of this paper is to obtain the exact distribution of r_k so that exact critical values or exact p-values can be computed. The method is based on Imhof (1961) and requires computing the eigenvalues of a $n \times n$ matrix and numerical integration over an infinite interval but where the truncation error is controlled. The method is valid not only for Gaussian white noise but under a more general spherically symmetric white noise process. We show an improvement with respect to alternative methods and compare it with the above mentioned (Box-Pierce) limits and other approximations (Ljung-Box, Dufour-Roy) plus a new one based on a result of Provost and Rudiuk (1995). We also consider the noncentered autocorrelations for which we derive new results.

*Service Informatique, Facultés Universitaires Saint-Louis, 43, bd du Jardin Botanique, B-1000 Bruxelles, Belgium, e-mail: ayadi@fusl.ac.be.

†ECARES & ISRO CP114, Université Libre de Bruxelles, Avenue Franklin Roosevelt 50, B-1050 Bruxelles, Belgium, e-mail: gmelard@ulb.ac.be. This paper has benefited from an IAP-network in Statistics, contract P5/24, financially supported by the Belgian Federal Office for Scientific, Technical and Cultural Affairs (OSTC). All correspondance to be sent to second author. Interested readers can obtain a copy of the programs upon request.

1 Introduction

Sample autocorrelations are one of the main instruments of time series analysis at the specification stage, see Box *et al.* (1994) and Choi (1992). They are essentially used to test that a series is a realization of a white noise process and to assess dependence at various lags. Among the many definitions that have been proposed in the literature, we consider here the most standard one: given n observations $x^T = (x_1, x_2, ..., x_n)$, where $"^T"$ denotes transposition, the sample autocorrelation at lag k is defined as

$$r_k(\bar{x}) = \frac{\sum_{i=1}^{n-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \quad 1 \leq k \leq n-1, \tag{1}$$

where $\bar{x} = \sum_{i=1}^{n} x_i/n$ is the sample mean.

The developments of the current paper are based on the assumption that the observations $x_1, x_2, ..., x_n$ are a realization of an uncorrelated spherically symmetric process, which we will call a spherically symmetric white noise process. A particular case is a Gaussian white noise process having a joint normal distribution with vector mean μ and covariance matrix $\sigma^2 I_n$. A white noise process with no other qualification will be implicitly spherically symmetric.

In the Gaussian case, Pan Jie-Jian (1968) obtained an integral representation for the distribution of $r_1(\bar{x})$. Provost and Rudiuk (1995), also in the Gaussian case, obtained an explicit representation (see Equation (15)) for the distribution of $r_k(\bar{x})$ for any lag $k = 1, ..., n-1$. Their representation is quite complex and, although it involves three infinite sums, they give no indication on how to deal with the truncation errors.

Using the approach of Imhof (1961) for inverting the characteristic function of quadratic forms in normal variables, through numerical integration, Ali (1984) obtained the exact distribution of $r_k(\bar{x})$ for a Gaussian stationary autoregressive moving average process, and investigated the adequacy of the normal, Edgeworth-type expansions and Pearson distributions. Recent papers with related interests are due to Provost and Rudiuk (1996), Butler and Paolella (1998), Forchini (2002), Paolella (2003)

Our work overlaps with that of Ali (1984) in the sense that we also use the approach of Imhof (1961) to treat the case of a white noise process, but we have dealt with the more general case of a spherically symmetric white noise and the matrix for which eigenvalues need to be computed is much simpler. Further, given the importance of this particular case, we have sought to refine his analysis. More precisely, we paid more attention to the numerical aspects of the implementation, and investigated new and classical

approximate distributions which have not been investigated by Ali (1984). Indeed,

i) we have been able to control more simply for numerical errors when approximating the integral in (13) (see Subsection 2.3),

ii) we have compared two numerical methods for computing the eigenvalues involved in the computation of the exact distribution with the aim of using the most efficient (see Section 2.3.3),

iii) we have proposed a new approximation to the exact distribution and investigated its adequacy and the adequacy of four other approximate distributions of which only one was investigated by Ali (1984).

The following simpler alternative to (1) is sometimes considered, see e.g. Anderson, 1970, p. 254:

$$r_k(0) = \frac{\sum_{i=1}^{n-k} x_i x_{i+k}}{\sum_{i=1}^{n} x_i^2}, \quad 1 \leq k \leq n-1. \tag{2}$$

The above definition can be considered as (1) applied to the centered series $\{x_1 - \bar{x}, x_2 - \bar{x}, ..., x_n - \bar{x}\}$. It can be used when there is some reason to believe that the mean is zero, as for model residuals.

Most theoretical results are concerned with the noncentered sample autocorrelation $r_k(0)$. The computation of its exact distribution is rendered computationally less demanding ; its symmetry about the origin is proved and new results on the moments are established. In particular, the expressions of the first four moments about the origin of $r_k(0)$ given by Equations (34) of Provost and Rudiuk (1995) are simplified. On the other hand, since the true mean of a process is generally unknown; we have investigated the question *"what is the risk incurred if* $r_k(0)$ *is used instead of* $r_k(\bar{x})$ *when the true mean is unknown?"*. A tentative answer is given in Section 5.

The sequel of this paper is organized as follows. The exact distribution is derived and the numerical aspects of its implementation are discussed in Section 2; it is shown that, apart from rounding-off errors, we can have control over all sources of error. In Section 3.1 the classical asymptotic approximations to the exact distribution of $r_k(\bar{x})$ are reviewed. In Section 3.2 a new approximation to the exact distribution is proposed. In Section 4 the exact distribution of $r_k(0)$, which may be used as a new approximation to the distribution of $r_k(\bar{x})$, is derived and new results are stated. In Section 5, the accuracy of all the approximations is investigated by examining the

deviations of the approximate critical values from the exact ones, for some levels and for some sample sizes n and lags k. General conclusions about the tests based on those approximations are drawn and recommendations on their use are given. In Section 6, the various tests are applied to examples taken from the literature. Numerical implementation of the different algorithms and simulation experiments throughout our work has been conducted using Fortran.

2 The exact distribution of centered autocorrelations

2.1 A representation of $\Pr(r_k(\bar{x}) \leq r)$

In matrix notation, $r_k(\bar{x})$ can be written as

$$r_k(\bar{x}) = \frac{x^T B_k x}{x^T B_0 x}, \tag{3}$$

where

$$\begin{array}{l} B_k = V A_k V, \\ V = I_n - \frac{1}{n}\varepsilon^T \varepsilon, \text{ where } \varepsilon = (1, ..., 1), \\ A_k = \frac{1}{2}(L_k + L_k^T), \end{array} \tag{4}$$

and L_k is a null matrix with the zeros in its k-subdiagonal replaced by 1's. Note that if we put $A_0 = I_n$, $B_0 = V A_0 V = V$, V being idempotent.

In this Section, we will be concerned by the exact distribution function of $r_k(\bar{x})$ under the null hypotheses that the series $x_1, x_2, ..., x_n$ is generated by a spherically symmetric white noise process with mean vector μ and covariance matrix $\sigma^2 I$. Because $r_k(\bar{x})$ is location and scale invariant, we will assume in the sequel, without loss of generality, that $\mu = 0$ and $\sigma = 1$. Note that we are placed in a more general context than the counterpart method of Provost and Rudiuk (1995), which is restricted to the Gaussian case, with the unnecessary condition $\mu = 0$.

The distribution of $r_k(\bar{x})$ can be written as

$$\Pr(r_k(\bar{x}) \leq r) = \Pr(\frac{x^T V A_k V x}{x^T V x} \leq r), \tag{5}$$

with x, A_k, V as defined in (4). The matrix $V A_k V$ is symmetric, so it can be diagonalized as

$$P^T V A_k V P = \Lambda,$$

where Λ =diag$(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix with eigenvalues on its main diagonal and P an orthogonal matrix of corresponding eigenvectors. Next, we make the orthogonal transformation $z = P^T x$. The quadratic form $x^T V A_k V x$ reduces then to $z^T \Lambda z$ where z has mean 0 and covariance matrix I_n. On the other hand, because VA_kV and V commute, then they can be simultaneously diagonalized with the same orthonormal matrix P (see Horn and Johnson, 1990, p. 235). Hence, noting that the rank of V is $n-1$ and that all its nonzero eigenvalues are equal to one (since V is idempotent), we see that $x^T V x = z^T I_n^0 z$, where I_n^0 is the identity matrix with the n-th diagonal element replaced with 0:

$$\Pr\{r_k(\bar{x}) \leq r\} = \Pr\left\{\frac{\sum_{i=1}^{n-1} \lambda_i z_i^2}{\sum_{i=1}^{n-1} z_i^2} \leq r\right\} = \Pr\left\{\sum_{i=1}^{n-1} (\lambda_i - r) z_i^2 \leq 0\right\} \quad (6)$$

Putting $\alpha_i := \lambda_i - r$, we get

$$\Pr\{r_k(\bar{x}) \leq r\} = \Pr\left\{\sum_{i=1}^{n-1} \alpha_i z_i^2 \leq 0\right\}. \quad (7)$$

Let us remark that VA_kV has generally k eigenvalues equal to 0 for $1 < k < n/2$. We have to substract r from all of the eigenvalues except from exactly one of the null eigenvalue, whatever it is.

Recall that if X and μ are $n \times n$ vectors with X random and μ fixed, then X is said to have a spherically symmetric distribution about μ if and only if $G(X - \mu)$ has the same distribution as $X - \mu$ for all orthogonal $n \times n$ matrices G. The class of spherically symmetric distributions includes such distributions as the multivariate normal and the multivariate Student-t with covariance matrix $\sigma^2 I_n$, the multivariate Cauchy, the multivariate exponential, etc. (see Johnson and Kotz, 1972).

Dufour and Roy (1985) have shown that the distribution of the $r_k(\bar{x})$ for any spherically symmetric distribution about 0 is the same as the distribution of $r_k(\bar{x})$ for a Gaussian white noise. We will therefore assume in the sequel that our series $x_1, x_2, ..., x_n$ (and consequently $z_1, z_2, ..., z_n$) is generated by a Gaussian white noise process, bearing in mind that the derived results will hold even if $x_1, x_2, ..., x_n$ was generated by any spherically symmetric white noise process.

The characteristic function of $\sum_{i=1}^{n-1} \alpha_i z_i^2$ can be obtained and inverted numerically, using Imhof (1961), to evaluate (7). This is detailed because we have to discuss the numerical properties of our solution.

2.2 Inversion of the characteristic function of $\sum_{j=1}^{n-1} \alpha_j z_j^2$

Suppose the x_j's ($j = 1, ..., n$) have a Gaussian distribution, then the z_j's also have a Gaussian distribution and, consequently, the z_j^2's ($j = 1, ..., n-1$) are independent χ^2-distributed variables with one degree of freedom. The characteristic function of z_j^2 is thus given by (see e.g. Manoukian, 1986)

$$h_j(t) = (1 - 2it)^{-1/2}. \tag{8}$$

Since the characteristic function of $\alpha_j z_j^2$ equals $h_j(\alpha_j t)$, the characteristic function of $\sum_{j=1}^{n-1} \alpha_j z_j^2$ is given by

$$\phi(t) = \prod_{j=1}^{n-1} (1 - 2i\alpha_j t)^{-1/2}. \tag{9}$$

We will now obtain (7) by means of the inversion formula (whose general form is due to Gil-Pilaez, 1951)

$$\Pr\left\{\sum_{j=1}^{n-1} \alpha_j z_j^2 \leq 0\right\} = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{t} Im\left[\phi(t)\right] dt. \tag{10}$$

The approach of Imhof (1961), given with more details by Koerts and Abrahamse (1969), allows to express $Im\left[\phi(t)\right]$ in known quantities, as follows. Let $1 - 2i\alpha_j t = r_j e^{i\theta_j}$, where $r_j = (1 + 4\alpha_j^2 t^2)^{1/2}$ and $\theta_j = -\arctan(2\alpha_j t)$. We have

$$\phi(t) = (\prod_{j=1}^{n-1} (1 + 4\alpha_j^2 t^2)^{-1/4}) e^{i\frac{1}{2}\sum_{j=1}^{n-1} \arctan(2\alpha_j t)}.$$

The imaginary part of $\phi(t)$ is given by

$$Im\left[\phi(t)\right] = \sin\left[\arg\phi(t)\right] |\phi(t)| \sin\left[\frac{1}{2}\sum_{j=1}^{n-1} \arctan(2\alpha_j t)\right] \prod_{j=1}^{n-1} (1 + 4\alpha_j^2 t^2)^{-1/4}.$$

After the substitution $u = 2t$ is made, we obtain

$$\Pr\left\{r_k(\bar{x}) \leq r\right\} = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\sin\varepsilon(u)}{u\gamma(u)} du, \tag{13}$$

where

$$\epsilon(u) = \frac{1}{2}\sum_{j=1}^{n-1} \arctan(\alpha_j u), \qquad \gamma(u) = \prod_{j=1}^{n-1} (1 + \alpha_j^2 u^2)^{1/4}. \tag{14}$$

It is worth noting that the representation of the distribution function of $r_k(\bar{x})$ given by (13) is greatly simpler than the following one, given by Provost and Rudiuk (1995):

$$\Pr\{r_k(\bar{x}) \leq r\} = \sum_{v=0}^{\infty}\sum_{w=0}^{\infty}\sum_{j=0}^{\infty} K_v K'_w \frac{\Gamma(\frac{\rho_1+\rho_2}{2}+v+w+j)}{(\frac{\rho_1}{2}+v)_{j+1}2^{\frac{\rho_1+\rho_2}{2}+v+w+j}}, \qquad (15)$$

where ρ_1 and ρ_2 are the numbers of eigenvalues λ_i of VA_kV such that, respectively, $\lambda_i - r > 0$ and $\lambda_i - r < 0$, $(\alpha)_\beta = \Gamma(\alpha+\beta)/\Gamma(\alpha)$,

$$K_v = \sum_{v_1+\ldots+v_{\rho_1}=v} \Big(\prod_{j:\lambda_j - r>0} \mu_j^{-\frac{1}{2}}\Big)\frac{(\frac{1}{2})_{v_1}\ldots(\frac{1}{2})_{v_{\rho_1}} c_1^{v_1}\ldots c_{\rho_1}^{v_{\rho_1}}}{v_1!\ldots v_{\rho_1}!\Gamma(v+\frac{\rho_1}{2})},$$

with $\mu_j = 2(\lambda_i - r)/\theta$, $c_j = (\mu_j - 1)/\mu_j$, and

$$K'_w = \sum_{w_1+\ldots+w_{\rho_2}=w} \Big(\prod_{j:\lambda_j - r<0} \mu_j^{-\frac{1}{2}}\Big)\frac{(\frac{1}{2})_{w_1}\ldots(\frac{1}{2})_{w_{\rho_2}} d_1^{w_1}\ldots d_{\rho_2}^{w_{\rho 2}}}{w_1!\ldots w_{\rho_2}!\Gamma(w+\frac{\rho_2}{2})},$$

with $\mu_j = -2(\lambda_i - r)/\theta, \qquad d_j = (\mu_j - 1)/\mu_j$, .

2.3 Numerical implementation

2.3.1 Upper bound of integration

In the present section we explain how to use numerical methods of integration to compute the integral in (13). The function $u\gamma(u)$ increases monotonically towards infinity. Therefore, the integration in (13) will be carried out over a range $0 \leq u \leq U$ only. We will get the approximation

$$\Pr\{r_k(\bar{x}) \leq r\} \simeq \frac{1}{2} - \frac{1}{\pi}\int_0^U \frac{\sin\varepsilon(u)}{u\gamma(u)}du\ . \qquad (16)$$

The degree of approximation obtained will depend (apart from rounding-off errors) on the *error of integration*, resulting from the use of an approximate rule for computing the integral, and the *error of truncation.* Denote

$$t_U = \frac{1}{\pi}\int_U^\infty \frac{\sin\varepsilon(u)}{u\gamma(u)}du, \qquad (17)$$

which can be bounded above as follows

$$|t_U| = \frac{1}{\pi}\left|\int_U^\infty \frac{\sin\varepsilon(u)}{u\gamma(u)}du\right| \leq \frac{1}{\pi}\int_U^\infty \frac{1}{u\prod_{j=1}^{n-1}(1+\alpha_j^2u^2)^{1/4}}du. \qquad (18)$$

Further, the inequality $(1+\alpha_j^2 u^2)^{\frac{1}{4}} > (\alpha_j^2 u^2)^{\frac{1}{4}} = |\alpha_j|^{\frac{1}{2}}\, u^{\frac{1}{2}}$ implies

$$\prod_{j=1}^{m}(1+\alpha_j^2 u^2)^{1/4} > \prod_{j=1}^{m}|\alpha_j|^{\frac{1}{2}}\, u^{\frac{m}{2}}, \tag{19}$$

where $\alpha_1, ..., \alpha_m$ are the non-zero elements of $\alpha_1, ..., \alpha_{n-1}$. Thus

$$|t_U| \leq \frac{1}{\pi}\int_U^{\infty}\left[\prod_{j=1}^{m}|\alpha_j|^{\frac{1}{2}}\right]^{-1} u^{-\frac{m}{2}-1}du = \frac{2}{m\pi}\left[\prod_{j=1}^{m}|\alpha_j|^{\frac{1}{2}}\right]^{-1} U^{-\frac{m}{2}} := T_U.$$

The numerical integration over the range $0 \leq u \leq U$ can thus be carried out using the following procedure:

1. Specify the truncation error ε you are willing to accept.
2. Determine the corresponding value of U by taking $T_U = \varepsilon$, then deduce

$$U = \left[\frac{2}{\varepsilon m\pi \prod_{j=1}^{m}|\alpha_j|^{\frac{1}{2}}}\right]^{\frac{2}{m}}. \tag{21}$$

Note that the upper bound of the integral is computed here explicitly unlike in Imhof (1961) where the computation of U for a general quadratic form does not seem feasible, and according to Imhof (1961,p. 423): *"One can hopefully expect that T_U will often be satisfactorily small, even for moderate values of U"*. Results for U are provided in Table 17. For the exact distribution of the sample autocorrelations of an ARMA process, the computation of U is feasible but less straightforward (Ali, 1984).

2.3.2 Numerical integration

The integral in (16) can be computed using a numerical rule of integration. In this paper we have used Simpson's rule. If $u = 0$, we take the limiting form (Imhof, 1961)

$$\lim_{u\to 0}\frac{\sin\varepsilon(u)}{u\gamma(u)} = \frac{1}{2}\sum_{i=1}^{n-1}\alpha_i. \tag{22}$$

In computing the exact distribution of $r_k(\bar{x})$, we performed the numerical integration following the procedure outlined in Subsection 2.3, allowing for a truncation error of less than 10^{-10}, and using Simpson's rule as implemented in subroutine QSIMP of Press *et al.* (1988, p. 133) where the integration error was set to 10^{-6}.

2.3.3 Computation of eigenvalues

The eigenvalues of the $n \times n$ symmetric matrix VA_kV involved in the exact distribution are computed using two different methods: the QL method and the Maehly version of Newton's method.

- The QL method consists of two stages: First, VA_kV is reduced to a tridiagonal form T by $n-2$ orthogonal transformations using the Householder algorithm as is classically done for a general real symmetric matrix. Second, the QL algorithm is used to compute the eigenvalues of T. A detailed description of the two stages is found in Wilkinson and Reinsch (1971) and Press *et al.* (1988). The numerical implementation uses subroutines TDIAG and LRVT of Algorithm AS 60.2 given in Sparks and Todd (1973).

- The Maehly version of Newton's method is described in Stoer and Bulirsch (1980) and an algorithm is given. A brief outline of the method is given in Appendix B. Proposition 7 therein provides the method with an initial value.

Preliminary comparisons have revealed that the QL method is faster. Details are omitted for reasons of lack of space.

3 Approximate distributions

3.1 The asymptotic distributions of the centered sample autocorrelations

Tests for white noise that use the sample autocorrelation $r_k(\bar{x})$ are usually based on the asymptotic normal distribution with mean 0 and approximate variance $1/n$ (Box and Pierce, 1970):

$$r_k(\bar{x}) \sim N(0, \frac{1}{n}), \tag{23}$$

or on the asymptotic normal distribution with mean 0 and approximate variance $(n-k)/n(n+2)$ (which are the exact mean and variance of $r_k(0)$) (Ljung and Box, 1978):

$$r_k(\bar{x}) \sim N(0, \frac{n-k}{n(n+2)}). \tag{24}$$

Replacing the approximate mean and variance by the exact mean and the exact variance we get the asymptotic normal distribution (Dufour and Roy, 1985)

$$r_k(\bar{x}) \sim N(-\frac{n-k}{n(n-1)}, \frac{(n-k)(n-2)(n^2+n-2k) - 2n(n-1)(n-2k)^+}{n^2(n-1)^2(n+1)}), \tag{25}$$

where $(a)^+ = \max\{a, 0\}$.

The mean of (25) was given by Moran (1948), Dufour and Roy (1985), and Anderson (1990) and was shown by Dufour and Roy (1985) to hold whenever $x_1, x_2, ..., x_n$ are exchangeable. The variance of (25) was given by Moran (1948) for $k = 1$ under normality; by Dufour and Roy (1985) for any k under spherical symmetry, and by Anderson (1990) for any k under normality).

In order to investigate the accuracy of the above-mentioned normal approximations, Dufour and Roy (1985) conducted Monte Carlo experiments and examined the empirical frequencies of rejection of the null hypothesis of white noise by tests-based on the standardized statistics- with three different nominal levels $(0.05, 0.1$ and $0.2)$. For each value of n and k, they considered three types of tests: one-sided tests against positive serial dependence, one-sided tests against negative serial dependence, and two-sided tests.

As an alternative to Dufour and Roy (1985) approach, we will rather investigate the accuracy of the approximations (23), (24), (25), and two others to be introduced in Sections 3.2 and 4 by using the exact distribution function of the sample autocorrelations $r_k(\bar{x})$. The advantage of the new approach is that for any combination (n, k), all the distributions can be compared at every point of the interval $[-1, 1]$, not only at a few selected critical values, and can be compared at every point of the interval $[0, 1]$, not only at a few significance levels. Further details will be given in Section 5.

3.2 A new approximate distribution of $r_k(\bar{x})$

The evaluation of the exact distribution of $r_k(\bar{x})$ requires the computation of the exact eigenvalues of VA_kV, which become computationally costly when n is large. To overcome this difficulty we propose a new approximation whose derivation requires the eigenvalues of A_k. These are given explicitly in the following:

Proposition 1 (Provost and Rudiuk (1995)) *The eigenvalues of the ma-*

trix A_k, defined in (4), are

$$\mu_{i,n} = \mu_{kj-l,km+l'} = \cos\frac{j\pi}{m+1} \quad \text{if } l' < k-l \tag{26}$$

$$\mu_{i,n} = \mu_{kj-l,km+l'} = \cos\frac{j\pi}{m+2} \quad \text{if } l' \geq k-l, \tag{27}$$

where

$$i = kj - l, \quad j \leq \frac{n+l}{k}, \quad l = 0, 1, ..., k-1, \tag{28}$$

and

$$n = km + l', \quad l' = 0, 1, ..., k-1, \quad m \leq \frac{n-l'}{k}. \tag{29}$$

The eigenvector corresponding to the eigenvalue $\mu_{i,n} = \mu_{kj-l,km+l'}$ is denoted by $x_{i,n}$ and its components are

$$x_i = \begin{cases} \sin(u\frac{j\pi}{m+1}), & i = uk - l, \quad u = 1, ..., [\frac{n+l}{k}] \\ 0 & \textit{otherwise.} \end{cases}$$

The row vectors $x'_{i,n}$, $i = 1, ..., n$, form an orthogonal matrix.

We deduce the following proposition in Appendix 1.

Proposition 2 *The eigenvalues $\mu_{i,n}$ $(i = 1, ..., n)$ of A_k occur in pairs of opposite signs. That is, if $\mu_{i,n}$ is an eigenvalue of A_k, then so is $-\mu_{i,n}$. More precisely,*

$$\mu_{n-i+1,n} = -\mu_{i,n}.$$

Provost and Rudiuk (1995) seem not to have noticed this useful result. In Section 4 we use it to prove the symmetry of the exact distribution of $r_k(0)$ and give simplified forms for its moments.

Now, the eigenvalues $\lambda_{i,n}$ of VA_kV could be approximated using the asymptotic result of the following conjecture.

Conjecture 3 *Consider the $n \times n$ matrices A_k and V defined in (4). Let $\mu_{1,n} \geq .... \geq \mu_{n,n}$ be the eigenvalues of A_k, let $\mu^*_{1,n} \geq .... \geq \mu^*_{n,n}$ be the eigenvalues of A_k where $\mu_{1,n}$ is replaced with 0, and let $\lambda_{1,n} \geq .... \geq \lambda_{n,n}$ be the eigenvalues of VA_kV. Then,*

$$\lim_{n\to\infty} \lambda_{i,n} = \mu^*_{i,n}, \ i = 1, ..., n.$$

This conjecture means that the eigenvalues of VA_kV are asymptotically the same as those of A_k, except that the largest eigenvalue of the latter should be replaced with 0.

Large scale numerical experiments have given evidence that the $\lambda_{i,n}(VA_kV) = \lambda_{i,n}$ are well approximated by $\mu^*_{i,n}(A_k) = \mu^*_{i,n}$ for all $i = 1, ...n$, even for small values of n, as is shown in the following table for some combinations of n and k. The table entries contain the largest differences $\left|\lambda_{i,n}(VA_kV) - \mu^*_{i,n}(A_k)\right|$ over all $i = 1, ..., n$.

			k			
n	1	2	3	7	15	20
10	2.8×10^{-2}	5.8×10^{-2}	6.2×10^{-2}			
100	3.9×10^{-4}	1.5×10^{-3}	2.3×10^{-3}	8.8×10^{-3}		
400	2.5×10^{-5}	1.0×10^{-6}	2.0×10^{-4}	9.1×10^{-4}	4.3×10^{-3}	
500	1.6×10^{-5}	6.4×10^{-5}	1.4×10^{-4}	6.5×10^{-4}	2.3×10^{-3}	5.8×10^{-3}
700	8.2×10^{-6}	3.3×10^{-5}	6.9×10^{-5}	3.9×10^{-4}	1.6×10^{-3}	3.1×10^{-3}

Let r be a real number and $\alpha^*_j = \mu^*_{j,n}(A_k) - r$, the exact distribution function of $r_k(\bar{x})$ can thus be approximated by

$$\Pr\{r_k(\bar{x}) \leq r\} = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\sin \varepsilon^*(u)}{u\gamma(u)}du, \tag{30}$$

where

$$\epsilon^*(u) = \frac{1}{2}\sum_{j=1}^{n-1}\arctan(\alpha^*_j u), \qquad \gamma^*(u) = \prod_{j=1}^{n-1}(1 + \alpha^{*2}_j u^2)^{1/4}.$$

The numerical results of Section 5 show that the above approximation to (13) is quite accurate even for very small sample sizes. Before that, let us also consider the centered autocorrelation which can also be used as an approximation of the noncentered autocorrelation.

4 The exact distribution of the noncentered autocorrelations

In this section we establish new results on the exact distribution of $r_k(0)$ for a spherically symmetric white noise process. More specifically, we

i) render the exact distribution of $r_k(0)$ computationally less demanding,

ii) prove that the exact distribution of $r_k(0)$ is symmetric about the origin,

iii) establish new results on the moments,

iv) assess the risk incurred if $r_k(0)$ is used instead of $r_k(\bar{x})$, when the true mean is unknown.

Using the same arguments as for $r_k(\bar{x})$, it is easy to see that

$$\begin{aligned}\Pr\{r_k(0) \leq r\} &= \Pr\left\{x^T(A_k - rI_n)x) \leq 0\right\} \\ &= \Pr\left\{\sum_{i=1}^{n}(\mu_i - r)z_i^2 \leq 0\right\} \\ &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\sin\left[\frac{1}{2}\sum_{j=1}^{n}\arctan((\mu_j - r)u)\right]}{u\prod_{j=1}^{n}(1+(\mu_j - r)^2u^2)^{1/4}}du, \quad (31)\end{aligned}$$

where the μ_j's are the eigenvalues of A_k. Their explicit expression and some of their properties were given in Propositions 1 and 2. The following is proved in Appendix 1.

Proposition 4 *The distribution of $r_k(0)$, at any lag k, of a spherically symmetric white noise process, is symmetric around 0. That is, for any real number r,*

$$\Pr(r_k(0) \leq -r) = \Pr(r_k(0) \geq r).$$

Note that Watson and Durbin (1951) have given a closed-form expression for the exact distribution of $r_1(0)$ and pointed out its symmetry, in the Gaussian case. To our knowledge, apart from this paper, no such statement has been made regarding $r_k(0)$ for higher lags. Note also that the exact distribution function of $r_k(\bar{x})$ is not symmetric about the origin since $E(r_k(\bar{x})) = -(n-k)/n(n+2)$ (which is different from 0), and $r_k(\bar{x})$ takes on values on $[-1, 1]$.

Corollary 5 *Odd order moments about the origin and odd order cumulants of the noncentered sample autocorrelations $r_k(0)$ of a spherically symmetric white noise process are zero.*

We also have the following corollary of Proposition 2.

Corollary 6 *Consider the $n \times n$ matrix A_k defined in (4). Then*

$$tr(A_k^{2i+1}) = 0, \text{ for all } i = 0, 1, 2, ...$$

It is known that the moments and cumulants of the quadratic forms in n independent standardized variables are identified with the traces of the matrix of the form (A_k in our case). Refer to Mathai and Provost (1992) for more details. The above corollary could therefore be used to simplify those moments and cumulants. For example, it is quite useful for simplifying the following expressions of the first four moments about the origin of $r_k(0)$ given by Provost and Rudiuk (1995, Equations (34)):

$$\begin{aligned}
\mu'_1 &= \frac{trA_k}{n} \\
\mu'_2 &= \frac{2trA_k^2 + (trA_k)^2}{n(n+2)} \\
\mu'_3 &= \frac{8trA_k^3 + 6trA_k^2 trA_k + (trA_k)^3}{n(n+2)(n+4)} \\
\mu'_4 &= \frac{48trA_k^4 + 32trA_k^3 trA_k + 12trA_k^2(trA_k)^2 + 12(trA_k^2)^2 + trA_k}{n(n+2)(n+4)(n+6)}.
\end{aligned} \tag{32}$$

Using the above corollary, Equations (32) simplify to

$$\mu'_1 = \mu'_3 = 0,\ \mu'_2 = \frac{2trA_k^2}{n(n+2)}, \mu'_4 = \frac{48trA_k^2 + 12(trA_k^2)^2}{n(n+2)(n+4)(n+6)}. \tag{33}$$

5 Numerical Results

In the present section we compare the distributions considered in Sections 2, 3 and 4. The comparison was carried out using two methods:

i) Graphically, at every point of the interval $[-1, 1]$, but for a limited number of sample sizes and time lags. This is particularly useful when we need to pay more attention to specific sample sizes and time lags, and

ii) Numerically, by comparing the critical values, associated with the standard significance levels, of the various distributions by considering a wide range of sample sizes and time lags.

General conclusions about the tests based on the above distributions are drawn and a few recommendations on their use are given.

5.1 Graphical comparisons

In Figure 1 is plotted the graphs of all the distribution functions considered in the paper. These have been evaluated at 101 equally-spaced points on the interval $[-1, 1]$. We considered several values of (n, k) but only the case $(10, 1)$,is shown here. The graphical comparisons are particularly useful when we need to get additional insight in specific cases of sample sizes and time lags without any restriction on the significance level.

5.2 Critical values, exact case

For testing purposes, it is useful to obtain critical values. Here we use the *bisection* method for computing lower critical values r of order α on the basis of the distribution function $F(r) = Pr(r_k(\bar{x}) \leq r) = \alpha$. Let r_1 and r_2 be such that $F(r_1) < \alpha$ and $F(r_2) > \alpha$ and compute $F(r)$ for $r = \bar{r} = (r_1 + r_2)/2$. If $F(\bar{r}) < \alpha$, then we replace r_1 by $\bar{r}$; otherwise we replace r_2 by $\bar{r}$. This procedure is iterated until $|r_1 - r_2|$ becomes smaller than a preassigned level ϵ (which was set to 10^{-6} in the numerical implementation).

Tables 1 through 4 report the lower critical values $\rho^{\alpha}_{k,n}$ (associated with the significance levels $\alpha = 0.025$, 0.05, 0.95 and 0.975) of the exact distribution of $r_k(\bar{x})$ for (n, k) such that $n = 4, ..., 50$ and $k = 1, ..., [n/3]$. The Fortran 77 program allowing to obtain other tables is available from the authors (http://homepages.ulb.ac.be/~gmelard).

5.3 Critical values, a tabular comparison

Table 5 reports the lower critical values $\rho^{\alpha}_{k,n}$ of the exact distribution of $r_k(\bar{x})$ together with the deviations from it of the lower critical values $\tilde{\rho}^{l,\alpha}_{k,n}$ of the various approximate distributions for the following combinations of k and n :

		k				
	10	1	3			
n	20	1	3	7		
	50	1	3	7	15	
	100	1	3	7	15	20

(34)

We make the following observations:

First of all, the following remark is of order: any tentative to draw a general conclusion from the above table would be very awkward; one should be specific at least about the levels and the sample size. The following

features are, however, worth noting although their validity is limited to the cases in hand.

For levels 0.025 and 0.05, the AM approximation generally gives the best approximation to the exact critical values. On the other hand, it systematically leads to non conservative tests of autocorrelation for all considered lags and sample sizes.

For levels 0.95 and 0.975, the Dufour-Roy approximation is generally the most precise and gives conservative tests for practically all considered sample sizes and lags. The AM approximation comes second regarding precision but always leads to conservative tests of autocorrelation.

Based on the two previous paragraphs, we conjecture that the use of the AM approximation for lower tails and the Dufour-Roy approximation for upper tails would give us tests with the most precise size.

For $n = 50, 100$ and for all considered critical values and autocorrelation orders, the error on the critical values is never larger than 0.007 for the AM approximation and never larger than 0.009 for the Dufour-Roy approximation. These two approximations are clearly more convenient for small samples than the other considered approximations.

The classically used approximation of Box-Pierce could be considered as satisfactory for levels 0.025 and 0.05 but as rather poor for levels 0.95 and 0.975. It is even the poorest approximation for these levels and gives the most conservative tests of autocorrelation.

Now let us address the following question which was already asked in the introduction: *"what is the risk incurred if $r_k(0)$ is used instead of $r_k(\bar{x})$ when the true mean is unknown ?":*

Table 5 shows that dropping $\bar{x}$, by using the exact distribution of $r_k(0)$, leads to quite large errors on critical values for $n = 10, 20$ and 50. Indeed, the error can be as large as 0.105 for $n = 10, k = 1$ and $\alpha = 0.95$, and about 0.01 for $n = 100$ and for all lags and levels.

5.4 Critical values, a graphical overview

Now, Figures 2 through Figure 5 give us a general view of the accuracy of all approximations.

On Figures 2 and 3 are plotted, for each autocorrelation lag k, the mean absolute deviation (MAD) obtained by averaging over sample sizes $n = K_k, ..., 50$, where $K_k = \max(3k, 4)$, the absolute values of the errors on the critical values of the various approximations, subscripted by l, to the exact

distribution:

$$MAD_k^{l,\alpha} = \frac{1}{51 - K_k} \sum_{n=K_k}^{50} \left| \rho_{k,n}^{\alpha} - \tilde{\rho}_{k,n}^{l,\alpha} \right|.$$

It emerges from those figures that, for levels 0.025 and 0.05, the AM approximation outperforms all the other approximations for almost all lags, and is even far superior for high lags. For levels 0.95 and 0.975, the Dufour-Roy approximation is globally the best approximation but it only slightly outperforms the AM approximation and the difference gets less and less significant as the lag increases.

On Figures 4 and 5 are plotted, for each sample size n, the MAD computed by averaging over lags $k = 1, ..., \lfloor n/3 \rfloor$, the absolute values of the errors on the critical values of the various approximations to the exact distribution:

$$MAD_n^{l,\alpha} = \frac{1}{\lfloor n/3 \rfloor} \sum_{k=1}^{\lfloor n/3 \rfloor} \left| \rho_{k,n}^{\alpha} - \tilde{\rho}_{k,n}^{l,\alpha} \right|.$$

For levels 0.025 and 0.05, and for all sample sizes larger than $n = 15$, the AM approximation is in average the most accurate, but for $n = 4, ..., 15$ the picture is less clear-cut. For levels 0.95 and 0.975 and for sample sizes $n \geq 5$, the approximations may be ordered in the following descending order of accuracy: Dufour-Roy, AM approximation, Box-Pierce, Exact (of $r_k(0)$) and Ljung-Box. Note that the AM and Dufour-Roy approximations are almost coincident.

6 Examples

The autocorrelation tests, for lags $k = 1, ..., [n/3]$, based on the exact and approximate distributions of $r_k(\bar{x})$ were applied to the series described in the following two examples. The sample autocorrelations, jointly with the p-values (p) (significance levels) associated with the two-sided tests based on the various distributions, are reported in Table 2 and Table 3. Let F denote either the exact or an approximate distribution function of $r_k(\bar{x})$. For a double-sided test, the p-value for an observed sample autocorrelation r is by definition, $p = 2\min(1 - F(r), F(r))$. For one-sided tests against positive or negative autocorrelation, these p-values are to be divided by 2.

6.1 An example of Bartels (1982)

Bartels (1982) considered the following series of first differences of annual deflated aggregate undistributed income in corporate trading enterprises ($n = 18$):

$$-82.29, -31.14, 136.58, 85.42, 42.96, -122.72, 0.59, 55.77, 117.62,$$
$$18.88, -48.21, -63.70, -10.95, -211.38, -304.02, 30.72, 238.19, 140.98.$$

These data were also used by Hallin and Mélard (1988) to illustrate rank autocorrelations but the some of the data were badly reported (the first three values of the second line were moved at the end). We see from Table 6 that the exact test is significant for autocorrelations at lag 1 at the level of 0.05. Only the tests based on the AM and Dufour-Roy approximate distributions could detect that significant autocorrelation. This stresses once more the need in small sample size cases to use tests based on either the exact distribution or one of its approximations: AM or Dufour-Roy rather than the more traditional Box-Pierce limits.

6.2 Log forward-log spot prediction errors

As explained in Dufour (1981), a standard problem in studies of foreign exchange market consists in testing whether the forward exchange rate F_t is an 'optimal' predictor of the corresponding future spot rate S_{t+1}, both usually in log form. This is usually interpreted as implying that the errors of prediction $S_{t+1} - F_t$ have mean zero and are uncorrelated. Like Dufour (1981), we will consider here the case where S_t is the logarithm of the exchange rate between the German Mark and the US Dollar (DM/$US) and F_t is the logarithm of the one-month forward exchange rate, during the interesting episode of the German hyperinflation. The series studied is monthly and covers the period January 1921-August 1923.

We see from Table 7 that no test is significant (at a level of less than 0.09) for any of the lags considered. The exact test is significant at a level of 0.103 for lag 1. It has greater significance levels for greater lags. If we test against positive dependence only, the significance level for lag 1 is 0.0515.

Dufour (1981) applied a signed-rank test to the same data. The two-sided test was found to be significant for lag 1, with significance level 0.047. The one-sided test against positive autocorrelation was found to be significant with significance level 0.024 for lag 1 (0.050 and 0.057 for lags 2 and 3).

This discrepancy between Dufour's test and ours deserves some remarks. Dufour (1981) method is nonparametric and assumes only marginal symme-

try. Moreover, the series in hand exhibits signs of non-normality and heteroskedasticity, which pose no problem to Dufour's method, but they probably do for our method. We have seen before that for the latter method, deviation from normality poses no problem as far as the distribution remains spherically symmetric. Extension to processes with symmetric marginal distribution deserves to be investigated. On the other hand, heteroskedasticity may favor Dufour's method over ours. A study of robustness to heteroskedasticity may be needed.

6.3 An example of Provost and Rudiuk (1995)

The numerical computation of the exact distribution is corroborated by a numerical example. For comparison purposes, we consider the example of Provost and Rudiuk (1995) of a series of length 7. The exact distribution of the autocorrelation at lag 2 is computed using our exact method and the exact method of Provost and Rudiuk (1995) Exact(PR). The distribution values taken at points $r = -0.5 : 0.1 : 0.4$ by the two approaches are compared with the empirical distribution values obtained by generating 5000000 series of observations from independent standard normal variables. This was carried out by computing the inverse of the normal distribution function (using Function PPND of Algorithm AS11) of pseudo-random numbers between 0 and 1, generated using Function RANDOM of Algorithm AS 183, where the seeds IX, IY and IZ were set to 14, 4 and 1966 respectively.

From Table 8 we observe that the distribution values given by the AM method are closer to the empirical values (EMPIRICAL) than the distribution values given by the PR method. Indeed, the difference $|\text{EMPIRICAL} - \text{PR}|$ is as large as 0.0039 for $r = -0.3$ while $|\text{EMPIRICAL} - \text{AM}|$ is never larger than 0.0003. This shows that our method is numerically more accurate than Provost and Rudiuk (1995) method. This can be explained by the fact that in their method three infinite sums and one infinite integral are involved, without the ability to control the truncation errors. At the opposite, our method uses a simpler representation that allows us to have control over truncation error.

7 Simulation results

It is clear from Section 2 that our exact distribution is based on the assumption of normality. It seems natural to use that exact distribution for Gaussian white noise processes as an approximation for the other distributions which may be more accurate than the asymptotic distribution We

want to examine the quality of that approximation in some cases. We have generated 10000 series of length 40 from white noise processes using the following distributions: (a) normal, (b) uniform, (c) double exponential, (d) Cauchy, (e) Student with 2 degrees of freedom, (f) logistic. This was done using RANDOM (Wichmann and Hill, 1982).and the inverse method except for (d) and (e) and PPND (Beasley and Springer, 1977). We have considered the first 10, 20, 30 and all observations from each series and computed the autocorrelations for each lag from 1 to 4 (if $n = 10$), 9 (if $n = 20$), 14 (if $n = 30$ or 40). The difference d between the empirical quantiles from the simulations and the normal exact quantiles are displayed in Table 10. According to these results, using the normal exact critical values provides a conservative test for most of the distributions tried except the uniform distribution.and sometimes the logistic distribution.

We have considered the empirical distribution, over all simulations, of the p-values obtained from the Gaussian exact distribution. The average and standard deviations should be respectively 0.5 and 0.29.

Table 11. Kolmogorov-Smirnov test statistics for testing uniformity of the p-values of the test of white noise based on autocorrelation at lag k using the normal exact distribution for 10000 white noise time series of length n generated using the specified distributions. Values in italics are significant at the 5% level and values in bold are significant at the 1% level.

		(a)	(b)	(c)	(d)	(e)	(f)
n	k	normal	uniform	dble. expon.	Cauchy	t_2	logistic
10	1	0.009	*0.014*	**0.037**	**0.189**	**0.074**	**0.018**
	4	0.013	**0.020**	**0.050**	**0.205**	**0.089**	**0.026**
20	1	0.010	**0.017**	**0.025**	**0.250**	**0.088**	0.012
	4	0.007	0.014	**0.033**	**0.252**	**0.088**	*0.016*
30	1	0.010	0.014	**0.018**	**0.277**	**0.091**	0.008
	4	0.007	0.010	**0.026**	**0.284**	**0.090**	0.013
	8	0.008	*0.016*	**0.023**	**0.288**	**0.097**	0.008
	12	0.010	*0.016*	**0.021**	**0.314**	**0.098**	0.005
40	1	0.007	0.013	**0.018**	**0.298**	**0.092**	0.007
	4	0.010	0.007	**0.024**	**0.306**	**0.088**	*0.015*
	8	0.009	0.007	**0.028**	**0.303**	**0.095**	*0.016*
	12	0.006	0.010	**0.026**	**0.319**	**0.100**	0.013

We have performed a Kolmogorov-Smirnov test for uniformity of these distributions of p-values. Some results are given in Table 11 where the Kolmogorov-Smirnov test statistic is given. Asymptotic critical values for

$n = 10000$ observations are 0.0135 at the 5 % level and 0.01628 at the 1 % level. For samples of $n = 20$ or longer, the adequacy is not rejected systematically for uniform and logistic distributions. However, it is rejected for double exponential, Cauchy and t_2 distributions for all n.

8 Conclusions

The purpose of this paper is to stress once more, twenty years after Ali (1984), that it is easy to use the exact distribution of the autocorrelations at least in the case of a test for randomness. It is true that the procedure for an ARMA process is more elaborate. Aside from rank autocorrelations (e.g. Hallin and Mélard, 1988), that would be a better way to go than using the ubiquious asymptotic critical values, or worse the still more standard Bartlett limits (which are valid under stronger assumptions and which are badly interpreted most of the time). Unfortunately, our first investigations don't reveal a way to compute exact limits for the portmanteau test, where a closer approximation than the traditional Ljung-Box approach is studied by Dufour and Roy (1985). Doing a similar job for partial autocorrelations seems also beyond our capability.

APPENDICES

A Proofs of propositions in sections 3 and 4

Proposition 1. **Proof.** Write n and i as in Proposition 1:

$$n = km + l', \text{ where } m \leq \frac{n - l'}{k}, \quad l' = 0, 1, ..., k - 1,$$

and

$$i = kj - l, \text{ where } j \leq \frac{n + l}{k}, \quad l = 0, 1, ..., k - 1.$$

- If $l' < k - l$, write $\mu_{n-i+1,n}$ as

$$\begin{aligned} \mu_{n-i+1,n} &= \mu_{km+l'-(kj-l)+1,n} \\ &= \mu_{k(m-j+1)-(k-l'-l-1),n} \\ &\qquad \text{(by (26), since } l' < k - l' - l - 1) \\ &= \cos\tfrac{(m-j+1)\pi}{m+1} \\ &= \cos(\pi - \tfrac{j\pi}{m+1}) \\ &= -\cos\tfrac{j\pi}{m+1} \\ &= -\mu_{i,n} \qquad \text{(by (26), since } l' < k - l). \end{aligned}$$

- If $l' \geq k - l$, write $\mu_{n-i+1,n}$ as

$$\begin{aligned}
\mu_{n-i+1,n} &= \mu_{km+l'-(kj-l)+1,n} \\
&= \mu_{k(m-j+2)-(2k-l'-l-1),n} \\
&\quad \text{(by (27), since } l' \geq k-(2k-l'-l-1)), \\
&= \cos\tfrac{(m-j+2)\pi}{m+2} \\
&= \cos(\pi - \tfrac{j\pi}{m+2}) \\
&= -\cos\tfrac{j\pi}{m+2} \\
&= -\mu_{i,n} \qquad \text{(by (27), since } l' \geq k-l).
\end{aligned}$$

■

Proposition 4. **Proof.** The value of the exact distribution function of $r_k(0)$ at $-r$ is given by

$$F_k(-r) = \Pr(r_k(0) \leq -r) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\sin\left[\frac{1}{2}\sum_{j=1}^n \arctan((\mu_j + r)u)\right]}{u\prod_{j=1}^n (1+(\mu_j+r)^2u^2)^{1/4}}du,$$

where the μ_j's are the eigenvalues of A_k. Using Proposition 2, we observe that the arguments of the arctan function occur, in the sum corresponding to $F_k(-r)$, in pairs $(\mu_j + r)u$ and $(-\mu_j + r)u$. In $F_k(r)$, they occur in pairs $(\mu_j - r)u$ and $(-\mu_j - r)u$.
Noting that $(\mu_j + r)u = -(-\mu_j - r)u$ and $(-\mu_j + r)u = -(\mu_j - r)u$, and that the arctan function is odd, we obtain

$$\frac{1}{2}\sum_{j=1}^n \arctan((\mu_j - r)u) = -\frac{1}{2}\sum_{j=1}^n \arctan((\mu_j + r)u).$$

Similar observations yield

$$\prod_{j=1}^n (1+(\mu_j - r)^2u^2)^{1/4} = \prod_{j=1}^n (1+(\mu_j + r)^2u^2)^{1/4}.$$

Finally, since the *sine* function is odd

$$\begin{aligned}
\Pr(r_k(0) &\leq -r) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\sin[\epsilon(u)]}{u\gamma(u)}du \\
&= 1 - (\frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\sin[\epsilon(u)]}{u\gamma(u)}du) \\
&= 1 - \Pr(r_k(0) \leq r) \\
&= \Pr(r_k(0) \geq r),
\end{aligned}$$

which completes the proof. ■

B Newton's algorithm for computing the eigenvalues involved in the exact distribution of $r_k(\bar{x})$

In this section we describe how the Newton method applies to finding the eigenvalues of the symmetric tridiagonal matrix T obtained by applying the Householder algorithm to the symmetric matrix VA_kV. Those eigenvalues are the roots of the characteristic polynomial $p(x)$ associated with the tridiagonal matrix. In order to evaluate the iteration function of Newton's method,

$$x_{k+1} = x_k - \frac{p(x_k)}{p'(x_k)},$$

we have to calculate the value of the polynomial p, as well as the value of its first derivative, at the point $x = x_k$.

Write T as follows

$$T = \begin{pmatrix} \alpha_1 & \beta_2 & & 0 \\ \beta_2 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_n \\ 0 & & \beta_n & \alpha_n \end{pmatrix}, \quad \alpha_i,\ \beta_i \text{ real.}$$

Denoting by $p_i(x)$ the characteristic polynomial

$$p_i(x) = \det \begin{pmatrix} \alpha_1 - x & \beta_2 & & 0 \\ \beta_2 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_i \\ 0 & & \beta_i & \alpha_i - x \end{pmatrix}$$

of the principal minor formed by the first i rows and columns of the matrix T, we have the recursions (see Stoer and Bulirsch (1980))

$$\begin{aligned} p_0(x) &= 1, \\ p_1(x) &= (\alpha_1 - x).1, \\ p_i(x) &= (\alpha_i - x)p_{i-1}(x) - \beta_i^2 p_{i-2}(x), \quad i = 2, 3, ..., n, \\ p(x) &= \det(T - xI_n) = p_n(x). \end{aligned} \tag{35}$$

A similar recursion for calculating $p'(x)$ by differentiating equations (35):

$$
\begin{aligned}
p'_0(x) &= 0, \\
p'_1(x) &= -1, \\
p'_i(x) &= -p_{i-1}(x) + (\alpha_i - x)p'_{i-1}(x) - \beta_i^2 p'_{i-2}(x), \quad i = 2, 3, ..., n, \\
p'(x) &= p'_n(x).
\end{aligned} \tag{36}
$$

The two recursions (35) and (36) can be evaluated concurrently.

Since T is symmetric, all its eigenvalues are real, that is all roots λ_i, $\lambda_1 \geq .... \geq \lambda_n$ of $p(x)$ are real and we have, by Theorem 5.5.5 of Stoer and Bulirsch (1980), that Newton's method yields a convergent strictly decreasing sequence x_k for any initial value $x_0 > \lambda_1$.

A suitable value for x_0 in our case (where $p(x)$ is the characteristic polynomial of VA_kV) is 1 as implied by the following proposition:

Proposition 7 *Consider the $n \times n$ matrices A_k and V defined in (4), and let $\lambda_1, ..., \lambda_n$ be the eigenvalues of VA_kV. Then,*

$$1 \leq i \leq n \max |\lambda_i| \leq 1.$$

Proof. Let $\|.\|_2$ be the spectral norm such that for any square matrix A, $\|A\|_2 = \rho(A^T A)$, where $\rho(A) = \max_{1 \leq i \leq n} |\alpha_i|$ is the spectral radius of A and $\alpha_1, ..., \alpha_n$ are its eigenvalues.
Let $A = VA_kV$ with eigenvalues $\lambda_1, ..., \lambda_n$, and let $\gamma_1, ..., \gamma_n$ be the eigenvalues of A_k. The proof consists in showing that $\max_{1 \leq i \leq n} |\lambda_i|^2 \leq 1$.
Using the fact that

i) the eigenvalues of $(VA_kV)^2$ are λ_1^2,...,λ_n^2,

ii) the matrix VA_kV is symmetric, and

iii) $\|V\|_2 = 1$,

we get

$$
\begin{aligned}
\max_{1 \leq i \leq n} |\lambda_i|^2 &= \rho((VA_kV)^2) \\
&= \|VA_kV\|_2 \\
&\leq \|V\|_2 \|A_k\|_2 \|V\|_2 = \|A_k\|_2 \\
&\leq \rho({A_k}^2) = \max_{1 \leq i \leq n} |\gamma_i|^2 .
\end{aligned} \tag{37}
$$

Moreover, it is easy to show that (see, e.g., Horn and Johnson (1990)),

$$\max_{1 \leq i \leq n} |\gamma_i| \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |(A_k)_{ij}| .$$

For $k < (n-1)/2$ (which case is of most interest to us), it is easy to see that $\max\limits_{1\leq i\leq n} \sum_{j=1}^{n} |(A_k)_{ij}| = 1$. We thus have that $\max\limits_{1\leq i\leq n} |\gamma_i| \leq 1$ and $\max\limits_{1\leq i\leq n} |\gamma_i|^2 \leq 1$. In view of (37), it follows that

$$\max_{1\leq i\leq n} |\lambda_i|^2 \leq 1.$$

Thus,

$$\max_{1\leq i\leq n} |\lambda_i| \leq 1.$$

■

If the initial value x_0 is far from a root, then the sequence x_k obtained by Newton's method may converge very slowly in the beginning. This observation has led to considering the following *double-step* method :

$$x_{k+1} = x_k - 2\frac{p(x_k)}{p'(x_k)}, \quad k = 0, 1, 2, ...$$

instead of the straightforward Newton method.

Of course, there is now the danger of "overshooting"; for an initial point $x_0 > \lambda_1$, some x_{k+1} may overshoot λ_1, negating the benefit of the above-mentioned theorem (Theorem 5.5.5 of Stoer and Bulirsch (1980)). However, this overshooting can be detected, and, due to some remarkable properties of polynomials, a good initial value y ($\lambda_1 \geq y > \lambda_2$) with which to start a subsequent Newton procedure for the calculation of λ_2 can be recovered. The latter is a consequence of Theorem 5.5.9 of Stoer and Bulirsch (1980).

Having found the largest root λ_1 of the polynomial p, there are still the other roots $\lambda_2, \lambda_3, ..., \lambda_n$ to be found. The following idea suggests itself immediately: "divide off" the known root λ_1, that is, form the polynomial

$$p_1(x) = \frac{p(x)}{x - \lambda_1}$$

of degree $n-1$. This process is called deflation. The largest root of $p_1(x)$ is λ_2, and may be determined by the previously described procedures. Here λ_1 or, even better, the value $y = x_{k_0}$ found by overshooting may serve as a starting point. In this fashion, all roots will be found eventually.

Deflation, in general is not without hazard, because roundoff will preclude an exact determination of $p_1(x)$. The polynomial actually found in place of $p_1$will have roots different from $\lambda_2, \lambda_3, ..., \lambda_n$. These are then found by means of further approximations, with the result that the last roots may

be quit inaccurate. However, deflation has been found to be numerically stable if done with care (see Stoer and Bulirsch (1980, p. 278)).

Deflation can be avoided altogether by using the Maehly version of the (straightforward) Newton method for finding the root λ_{j+1} (see Maehly (1954)) :

$$x_{k+1} = \Phi_j(x_k) \quad \text{with} \quad \Phi_j(x) = x - \frac{p(x)}{p'(x) - \sum_{i=1}^{j-1} \frac{p(x)}{x-\lambda_i}}.$$

The advantage of this formula lies in the fact that the iteration given by Φ_j converges quadratically to λ_{j+1} even if the numbers $\lambda_1, ..., \lambda_j$ in Φ_j are not roots of p. Note that $\Phi_j(x)$ is not defined if $x = \lambda_k, k = 1, ..., j$, is a previous root of $p(x)$. Such roots cannot be selected as starting values. Instead one may use the values found by overshooting if the double-step method is employed. Stoer and Bulirsch (1980, p.279) give the following pseudo-ALGOL program for finding all roots of a polynomial p having only real roots. It incorporates all those features. The initial value x_0 may be taken equal to 1 in virtue of Proposition 7.

```
z0 :=starting point x0;
for j := 1 step 1 until n do
    begin m := 2;    zs := z0;
        Iteration: z := zs;    s := 0;
        for i := 1 step 1 until j − 1 do
            s := s + 1/(z − xi[i]);
            zs := p(z);    zs := z − m ∗ zs/(p'(z) − zs ∗ s);
            if zs < z then goto iteration;
            if m = 2 then
                begin zs := z;    m := 1; goto Iteration end;
    λj := z
end;
```

References

[1] ALI, M. M. (1984). Distribution of the sample autocorrelations when observations are from a stationary autoregressive moving average process. *J. Bus. Econ. Statist., **2**, 271-278.*

[2] ANDERSON, O. D. (1990). Moments of the sampled autocovariances and autocorrelations for a Gaussian white noise process. *The Canadian Journal of Statistics., **18**, No. 3, 271-284.*

[3] ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series.* New York: Wiley.

[4] BARTELS, R. (1982). The rank Version of Von Neumann's Ratio Test for Randomness. *Journal of The American Statistical Association*, **77**, 40-46.

[5] BARTLETT, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. *J. Roy. Statist. Soc. Suppl., **8**, 27-41 (Corrigenda 10 (1948) No. 1).*

[6] BEASLEY, J. D. and SPRINGER, S. G. (1977). Algorithm AS 111. The percentage points of the normal distribution. *J. Roy. Statist. Soc. Ser. C Appl. Statist., **26**, 118-120.*

[7] BUTLER, R. W. and PAOLELLA, M. S. (1998). Approximate distributions for the various serial correlograms, *Bernoulli*, **4**, 497-518.

[8] BOX, G. E. P., JENKINS, G. M. and REINSEL, G. C. (1994). *Time Series Analysis: Forecasting and Control.* Prentice Hall Press, Englewood Cliffs, N.J. (3rd edition).

[9] BOX, G. E. P., and PIERCE, D. A. (1970). Distribution of residual autocorrelations in Autoregressive-Integrated Moving Average time series models. *J. Amer. Statist. Assoc.,* **65**, 1509-1526.

[10] CHOI, Byoung Seon (1992). *ARMA model identification.* New York : Springer-Verlag.

[11] DUFOUR, J. M. (1981). Rank tests for serial dependence. *Journal of Time Series Analysis*, **2**, 117-128.

[12] DUFOUR, J. M., and R. ROY. (1985). Some robust exact results on sample autocorrelations and tests of randomness. *J. Econometrics, **29**, 257-273.*

[13] FORCHINI, G. (2002) The exact cumulative distribution function of a ratio of quadratic forms in normal variables, with applications to the AR(1) model, *Econometric Theory*, **18**, 833-852.

[14] GIL-PILAEZ, J. (1951). Note on the inversion theorem. *Biometrika, **38**, 481-2.*

[15] HALLIN, M. and MELARD, G. (1988). Rank-based tests for randomness against first-order serial dependence. *Journal of the American Statistical Association,* **83**, 404, 1117-1128.

[16] HILL, I. D. (1973). Algorithm AS 66. The Normal Integral. *J. Roy. Statist. Soc. Ser. C Appl. Statist. **22**, 3, 424-427.*

[17] HORN, R. A. and C. R. JOHNSON (1990). *Matrix Analysis*. University Press: Cambridge.

[18] IMHOF, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika, **48**, 3 and 4, p. 419.*

[19] JOHNSON, N. L., S. KOTZ (1972). Distributions in Statistics: Continuous Multivariate Distributions. New York: Wiley&Sons.

[20] KOERTS, J., and A. P. J. ABRAHAMSE (1969). *On the Theory and Application of the General Linear Model*. Rotterdam University Press, Rotterdam.

[21] LJUNG, G. M. and BOX, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, **65**, 2, 297-303.*

[22] MANOUKIAN, E. B. (1986). *Modern Concepts and Theorems of Mathematical Statistics*. New York : Springer-Verlag.

[23] MATHAI, A. M., and PROVOST, S. B. (1992). *Quadratic Forms in Random Variables*. Marcel Dekker, Inc., New York.

[24] MORAN, P. A. P. (1948). Some theorems on time series II: The significance of the serial correlation coefficient. *Biometrika, **35**, 255-260.*

[25] PAN, JIE JIAN (1968). Distributions of the noncircular serial correlation coefficients. *American Mathematical Society and Institute of Mathematical Statistics - Selected translations in Probability and Statistics*, **7**, 281-291.

[26] PAOLELLA, M. S. (2003). Computing moments of ratios of quadratic forms in normal variables, *Computational Statistics and Data Analysis*, **42**, 313-331.

[27] PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A., and VETTERLING, W. T. (1988). *Numerical Recipes in Fortran - The Art of Scientific Computing.* Cambridge: University Press.

[28] PROVOST, B. and RUDIUK, E. M. (1995). The sampling distribution of the serial correlation coefficient. *American Journal of Mathematical and Management Sciences., **15**, No. 1&2, 57-81.*

[29] PROVOST, B. and RUDIUK, E. M. (1996). The exact distribution of indefinite quadratic forms in noncentral normal vectors. *Ann. Inst. Statist. Math., **48**, No. 2, 381-394.*

[30] SPARKS, D. N., and TODD, A. D. (1973). Algorithm AS 60.1, Latent roots and vectors of a symmetric matrix, *J. Roy. Statist. Soc. Ser. C Appl. Statist.*, **22**, 260-264.

[31] STOER, J., and BULIRSCH, R. (1980). *Introduction to Numerical Analysis.* New York: Springer-Verlag.

[32] WATSON, G. S. and DURBIN, J. (1951). Exact tests of serial correlation using noncircular statistics. Ann. Math. Statist., **22**, 446-451.

[33] WICHMANN, B. A. and HILL, I. D. (1982). An efficient and portable pseudo-random number generator, *J. Roy. Statist. Soc. Ser. C Appl. Statist.*, **31**, 88-190 (correction: **33**, 1984, 123).

[34] WILKINSON, J. H., and C. REINSCH, editors (1971). *Linear Algebra*, Volume II of Handbook for Automatic Computation. Springer-Verlag: Berlin.

Table 1. Values of U, equation (21), granting a truncation error in $P(r_k(\overline{x}) \le r)$ of less than Precision.

			Precision=10^{-6}				Precision=10^{-10}			
k	***r***	***n***	10	20	50	100	10	20	50	100
1	-0.8		19.43	5.78	2.94	2.39	122.62	14.52	4.24	2.88
	-0.2		22.79	6.38	3.05	2.49	143.82	16.04	4.41	3.00
	0.0		25.31	33.88	5.88	3.36	196.00	85.09	8.49	4.04
	0.2		22.74	6.19	3.02	2.48	143.47	15.56	4.37	2.98
	0.8		15.67	5.19	2.81	2.34	98.90	13.04	4.07	2.82
3	-0.8		23.87	5.46	2.90	2.49	150.63	13.72	4.18	2.99
	-0.2		24.82	6.72	3.11	2.46	156.62	16.87	4.50	2.96
	0.0		981078	1263.88	12.24	4.80	6190183	3174.72	17.70	5.77
	0.2		22.98	6.37	3.09	2.45	144.98	16.01	4.46	2.95
	0.8		16.43	4.82	2.76	2.50	103.65	12.12	4.00	3.01
7	-0.8			5.78	2.83	2.51		14.53	4.09	3.02
	-0.2			6.74	3.09	2.54		16.94	4.47	3.05
	0.0			1885184	414.75	6.50		4735369	599.49	7.81
	0.2			6.53	3.06	2.56		16.39	4.43	3.08
	0.8			5.14	2.70	2.43		12.92	3.90	2.93
20	-0.8					2.36				2.83
	-0.2					2.44				2.93
	0.0					1806				2176
	0.2					2.44				2.94
	0.8					2.28				2.74

Table 2. Lower critical values of the exact distribution of the centered sample autocorrelations of a spherically symmetric white noise processα = 0.025.

Sample Sizes	Lags															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
4	-0.586															
5	-0.747															
6	-0.713	-0.613														
7	-0.679	-0.615														
8	-0.651	-0.605														
9	-0.624	-0.584	-0.531													
10	-0.600	-0.564	-0.524													
11	-0.578	-0.547	-0.515													
12	-0.558	-0.531	-0.504	-0.470												
13	-0.540	-0.516	-0.491	-0.463												
14	-0.523	-0.502	-0.479	-0.455												
15	-0.508	-0.489	-0.468	-0.448	-0.424											
16	-0.494	-0.476	-0.458	-0.439	-0.419											
17	-0.481	-0.465	-0.448	-0.431	-0.413											
18	-0.469	-0.454	-0.439	-0.423	-0.407	-0.389										
19	-0.457	-0.444	-0.430	-0.415	-0.400	-0.385										
20	-0.447	-0.434	-0.421	-0.408	-0.394	-0.380										
21	-0.437	-0.425	-0.413	-0.401	-0.388	-0.375	-0.361									
22	-0.427	-0.416	-0.405	-0.394	-0.382	-0.370	-0.357									
23	-0.419	-0.408	-0.398	-0.387	-0.376	-0.365	-0.353									
24	-0.410	-0.401	-0.391	-0.381	-0.371	-0.360	-0.350	-0.338								
25	-0.402	-0.393	-0.384	-0.375	-0.365	-0.356	-0.346	-0.335								
26	-0.395	-0.387	-0.378	-0.369	-0.360	-0.351	-0.342	-0.332								
27	-0.388	-0.380	-0.372	-0.364	-0.355	-0.347	-0.338	-0.329	-0.319							
28	-0.381	-0.374	-0.366	-0.358	-0.350	-0.342	-0.334	-0.325	-0.317							
29	-0.375	-0.368	-0.360	-0.353	-0.346	-0.338	-0.330	-0.322	-0.314							
30	-0.369	-0.362	-0.355	-0.348	-0.341	-0.334	-0.326	-0.319	-0.311	-0.303						
31	-0.363	-0.357	-0.350	-0.343	-0.337	-0.330	-0.323	-0.316	-0.308	-0.301						
32	-0.357	-0.351	-0.345	-0.339	-0.332	-0.326	-0.319	-0.312	-0.305	-0.298						
33	-0.352	-0.346	-0.340	-0.334	-0.328	-0.322	-0.316	-0.309	-0.303	-0.296	-0.289					
34	-0.347	-0.341	-0.336	-0.330	-0.324	-0.318	-0.312	-0.306	-0.300	-0.293	-0.287					
35	-0.342	-0.337	-0.331	-0.326	-0.320	-0.315	-0.309	-0.303	-0.297	-0.291	-0.285					
36	-0.338	-0.332	-0.327	-0.322	-0.317	-0.311	-0.306	-0.300	-0.294	-0.289	-0.283	-0.277				
37	-0.333	-0.328	-0.323	-0.318	-0.313	-0.308	-0.303	-0.297	-0.292	-0.286	-0.281	-0.275				
38	-0.329	-0.324	-0.319	-0.314	-0.309	-0.304	-0.299	-0.294	-0.289	-0.284	-0.278	-0.273				
39	-0.325	-0.320	-0.315	-0.311	-0.306	-0.301	-0.296	-0.292	-0.287	-0.282	-0.276	-0.271	-0.266			
40	-0.321	-0.316	-0.312	-0.307	-0.303	-0.298	-0.294	-0.289	-0.284	-0.279	-0.274	-0.269	-0.264			
41	-0.317	-0.312	-0.308	-0.304	-0.300	-0.295	-0.291	-0.286	-0.282	-0.277	-0.272	-0.267	-0.262			
42	-0.313	-0.309	-0.305	-0.301	-0.296	-0.292	-0.288	-0.284	-0.279	-0.275	-0.270	-0.265	-0.261	-0.256		
43	-0.309	-0.305	-0.301	-0.297	-0.293	-0.289	-0.285	-0.281	-0.277	-0.272	-0.268	-0.264	-0.259	-0.255		
44	-0.306	-0.302	-0.298	-0.294	-0.290	-0.287	-0.283	-0.279	-0.274	-0.270	-0.266	-0.262	-0.257	-0.253		
45	-0.302	-0.299	-0.295	-0.291	-0.288	-0.284	-0.280	-0.276	-0.272	-0.268	-0.264	-0.260	-0.256	-0.252	-0.247	
46	-0.299	-0.296	-0.292	-0.288	-0.285	-0.281	-0.277	-0.274	-0.270	-0.266	-0.262	-0.258	-0.254	-0.250	-0.246	
47	-0.296	-0.293	-0.289	-0.286	-0.282	-0.279	-0.275	-0.271	-0.268	-0.264	-0.260	-0.256	-0.253	-0.249	-0.245	
48	-0.293	-0.290	-0.286	-0.283	-0.279	-0.276	-0.273	-0.269	-0.266	-0.262	-0.258	-0.255	-0.251	-0.247	-0.243	-0.239
49	-0.290	-0.287	-0.283	-0.280	-0.277	-0.274	-0.270	-0.267	-0.263	-0.260	-0.256	-0.253	-0.249	-0.246	-0.242	-0.238
50	-0.287	-0.284	-0.281	-0.278	-0.274	-0.271	-0.268	-0.265	-0.261	-0.258	-0.255	-0.251	-0.248	-0.244	-0.241	-0.237

Table 3. Lower critical values of the exact distribution of the centered sample autocorrelations of a spherically symmetric white noise process. Significance level = 0.05.

Sample Size	Lags															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
4	-0.567															
5	-0.686															
6	-0.643	-0.569														
7	-0.613	-0.557														
8	-0.583	-0.540														
9	-0.556	-0.518	-0.477													
10	-0.533	-0.499	-0.465													
11	-0.511	-0.483	-0.454													
12	-0.492	-0.467	-0.442	-0.415												
13	-0.475	-0.453	-0.430	-0.407												
14	-0.459	-0.439	-0.419	-0.398												
15	-0.445	-0.427	-0.409	-0.390	-0.371											
16	-0.432	-0.416	-0.399	-0.382	-0.365											
17	-0.419	-0.405	-0.390	-0.375	-0.359											
18	-0.408	-0.395	-0.381	-0.367	-0.353	-0.339										
19	-0.398	-0.386	-0.373	-0.360	-0.347	-0.334										
20	-0.388	-0.377	-0.365	-0.353	-0.341	-0.329										
21	-0.379	-0.369	-0.358	-0.347	-0.336	-0.324	-0.313									
22	-0.371	-0.361	-0.351	-0.341	-0.330	-0.320	-0.309									
23	-0.363	-0.353	-0.344	-0.335	-0.325	-0.315	-0.305									
24	-0.355	-0.347	-0.338	-0.329	-0.320	-0.311	-0.302	-0.292								
25	-0.348	-0.340	-0.332	-0.324	-0.315	-0.307	-0.298	-0.289								
26	-0.341	-0.334	-0.326	-0.318	-0.311	-0.302	-0.294	-0.286								
27	-0.335	-0.328	-0.321	-0.313	-0.306	-0.298	-0.291	-0.283	-0.275							
28	-0.329	-0.322	-0.315	-0.309	-0.302	-0.294	-0.287	-0.280	-0.272							
29	-0.323	-0.317	-0.310	-0.304	-0.297	-0.291	-0.284	-0.277	-0.270							
30	-0.318	-0.312	-0.306	-0.300	-0.293	-0.287	-0.280	-0.274	-0.267	-0.261						
31	-0.313	-0.307	-0.301	-0.295	-0.289	-0.283	-0.277	-0.271	-0.265	-0.258						
32	-0.308	-0.302	-0.297	-0.291	-0.286	-0.280	-0.274	-0.268	-0.262	-0.256						
33	-0.303	-0.298	-0.293	-0.287	-0.282	-0.276	-0.271	-0.265	-0.260	-0.254	-0.248					
34	-0.298	-0.294	-0.289	-0.283	-0.278	-0.273	-0.268	-0.263	-0.257	-0.252	-0.246					
35	-0.294	-0.289	-0.285	-0.280	-0.275	-0.270	-0.265	-0.260	-0.255	-0.249	-0.244					
36	-0.290	-0.285	-0.281	-0.276	-0.272	-0.267	-0.262	-0.257	-0.252	-0.247	-0.242	-0.237				
37	-0.286	-0.282	-0.277	-0.273	-0.268	-0.264	-0.259	-0.255	-0.250	-0.245	-0.240	-0.235				
38	-0.282	-0.278	-0.274	-0.270	-0.265	-0.261	-0.257	-0.252	-0.248	-0.243	-0.238	-0.234				
39	-0.278	-0.274	-0.270	-0.266	-0.262	-0.258	-0.254	-0.250	-0.245	-0.241	-0.237	-0.232	-0.228			
40	-0.275	-0.271	-0.267	-0.263	-0.259	-0.255	-0.251	-0.247	-0.243	-0.239	-0.235	-0.230	-0.226			
41	-0.272	-0.268	-0.264	-0.260	-0.257	-0.253	-0.249	-0.245	-0.241	-0.237	-0.233	-0.229	-0.225			
42	-0.268	-0.265	-0.261	-0.257	-0.254	-0.250	-0.246	-0.243	-0.239	-0.235	-0.231	-0.227	-0.223	-0.219		
43	-0.265	-0.262	-0.258	-0.255	-0.251	-0.248	-0.244	-0.240	-0.237	-0.233	-0.229	-0.225	-0.222	-0.218		
44	-0.262	-0.259	-0.255	-0.252	-0.249	-0.245	-0.242	-0.238	-0.235	-0.231	-0.227	-0.224	-0.220	-0.216		
45	-0.259	-0.256	-0.253	-0.249	-0.246	-0.243	-0.239	-0.236	-0.233	-0.229	-0.226	-0.222	-0.219	-0.215	-0.211	
46	-0.256	-0.253	-0.250	-0.247	-0.244	-0.240	-0.237	-0.234	-0.231	-0.227	-0.224	-0.221	-0.217	-0.214	-0.210	
47	-0.253	-0.250	-0.247	-0.244	-0.241	-0.238	-0.235	-0.232	-0.229	-0.225	-0.222	-0.219	-0.216	-0.212	-0.209	
48	-0.251	-0.248	-0.245	-0.242	-0.239	-0.236	-0.233	-0.230	-0.227	-0.224	-0.221	-0.217	-0.214	-0.211	-0.208	-0.204
49	-0.248	-0.245	-0.242	-0.240	-0.237	-0.234	-0.231	-0.228	-0.225	-0.222	-0.219	-0.216	-0.213	-0.210	-0.206	-0.203
50	-0.245	-0.243	-0.240	-0.237	-0.234	-0.232	-0.229	-0.226	-0.223	-0.220	-0.217	-0.214	-0.211	-0.208	-0.205	-0.202

Table 4. Lower critical values of the exact distribution of the centered sample autocorrelations of a spherically symmetric white noise process. Significance level = 0.95.

Sample sizes	Lags															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
4	0.067															
5	0.309															
6	0.334	0.370														
7	0.349	0.349														
8	0.353	0.339														
9	0.351	0.337	0.340													
10	0.347	0.334	0.330													
11	0.342	0.330	0.322													
12	0.337	0.325	0.315	0.311												
13	0.331	0.320	0.310	0.304												
14	0.325	0.315	0.306	0.298												
15	0.319	0.310	0.301	0.293	0.287											
16	0.313	0.305	0.297	0.289	0.282											
17	0.308	0.300	0.292	0.285	0.278											
18	0.303	0.295	0.288	0.281	0.274	0.268										
19	0.298	0.291	0.284	0.277	0.270	0.264										
20	0.293	0.286	0.280	0.273	0.267	0.261										
21	0.288	0.282	0.276	0.270	0.263	0.258	0.252									
22	0.284	0.278	0.272	0.266	0.260	0.255	0.249									
23	0.279	0.274	0.268	0.263	0.257	0.252	0.246									
24	0.275	0.270	0.265	0.259	0.254	0.249	0.244	0.239								
25	0.271	0.266	0.261	0.256	0.251	0.246	0.241	0.236								
26	0.267	0.263	0.258	0.253	0.248	0.243	0.239	0.234								
27	0.264	0.259	0.255	0.250	0.246	0.241	0.236	0.232	0.227							
28	0.260	0.256	0.252	0.247	0.243	0.238	0.234	0.230	0.225							
29	0.257	0.253	0.249	0.244	0.240	0.236	0.232	0.228	0.223							
30	0.253	0.250	0.246	0.242	0.238	0.234	0.230	0.225	0.221	0.217						
31	0.250	0.247	0.243	0.239	0.235	0.231	0.227	0.223	0.220	0.216						
32	0.247	0.244	0.240	0.236	0.233	0.229	0.225	0.222	0.218	0.214						
33	0.244	0.241	0.237	0.234	0.230	0.227	0.223	0.220	0.216	0.212	0.209					
34	0.241	0.238	0.235	0.232	0.228	0.225	0.221	0.218	0.214	0.211	0.207					
35	0.239	0.236	0.232	0.229	0.226	0.223	0.219	0.216	0.213	0.209	0.206					
36	0.236	0.233	0.230	0.227	0.224	0.221	0.217	0.214	0.211	0.208	0.204	0.201				
37	0.233	0.231	0.228	0.225	0.222	0.219	0.215	0.212	0.209	0.206	0.203	0.200				
38	0.231	0.228	0.225	0.222	0.220	0.217	0.214	0.211	0.208	0.205	0.202	0.198				
39	0.229	0.226	0.223	0.220	0.218	0.215	0.212	0.209	0.206	0.203	0.200	0.197	0.194			
40	0.226	0.224	0.221	0.218	0.216	0.213	0.210	0.207	0.204	0.202	0.199	0.196	0.193			
41	0.224	0.221	0.219	0.216	0.214	0.211	0.208	0.206	0.203	0.200	0.197	0.195	0.192			
42	0.222	0.219	0.217	0.214	0.212	0.209	0.207	0.204	0.201	0.199	0.196	0.193	0.191	0.188		
43	0.220	0.217	0.215	0.212	0.210	0.208	0.205	0.203	0.200	0.197	0.195	0.192	0.190	0.187		
44	0.218	0.215	0.213	0.211	0.208	0.206	0.203	0.201	0.199	0.196	0.194	0.191	0.188	0.186		
45	0.216	0.213	0.211	0.209	0.207	0.204	0.202	0.200	0.197	0.195	0.192	0.190	0.187	0.185	0.182	
46	0.214	0.211	0.209	0.207	0.205	0.203	0.200	0.198	0.196	0.193	0.191	0.189	0.186	0.184	0.181	
47	0.212	0.210	0.207	0.205	0.203	0.201	0.199	0.197	0.194	0.192	0.190	0.187	0.185	0.183	0.180	
48	0.210	0.208	0.206	0.204	0.202	0.199	0.197	0.195	0.193	0.191	0.189	0.186	0.184	0.182	0.179	0.177
49	0.208	0.206	0.204	0.202	0.200	0.198	0.196	0.194	0.192	0.190	0.187	0.185	0.183	0.181	0.179	0.176
50	0.206	0.204	0.202	0.200	0.198	0.196	0.194	0.192	0.190	0.188	0.186	0.184	0.182	0.180	0.178	0.175

Table 5. Lower critical values of the exact distribution function of the centered sample autocorrelations of a spherically symmetric white noise process. Significance level = 0.975.

Sample sizes	Lags															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
4	0.086															
5	0.377															
6	0.413	0.463														
7	0.425	0.438														
8	0.431	0.426														
9	0.429	0.415	0.416													
10	0.423	0.407	0.403													
11	0.416	0.402	0.394													
12	0.409	0.395	0.386	0.377												
13	0.402	0.389	0.378	0.369												
14	0.394	0.382	0.371	0.362												
15	0.387	0.376	0.365	0.356	0.347											
16	0.380	0.370	0.359	0.350	0.342											
17	0.373	0.364	0.354	0.345	0.336											
18	0.367	0.358	0.348	0.340	0.332	0.323										
19	0.360	0.352	0.343	0.335	0.327	0.319										
20	0.354	0.346	0.338	0.330	0.323	0.315										
21	0.348	0.341	0.333	0.326	0.318	0.311	0.304									
22	0.343	0.336	0.329	0.322	0.314	0.307	0.300									
23	0.337	0.331	0.324	0.317	0.311	0.304	0.297									
24	0.332	0.326	0.320	0.313	0.307	0.300	0.294	0.288								
25	0.328	0.322	0.315	0.309	0.303	0.297	0.291	0.285								
26	0.323	0.317	0.311	0.306	0.300	0.294	0.288	0.282								
27	0.318	0.313	0.307	0.302	0.296	0.291	0.285	0.279	0.274							
28	0.314	0.309	0.304	0.298	0.293	0.288	0.282	0.277	0.271							
29	0.310	0.305	0.300	0.295	0.290	0.285	0.279	0.274	0.269							
30	0.306	0.301	0.296	0.292	0.287	0.282	0.277	0.272	0.267	0.262						
31	0.302	0.297	0.293	0.288	0.284	0.279	0.274	0.269	0.265	0.260						
32	0.298	0.294	0.290	0.285	0.281	0.276	0.272	0.267	0.262	0.258						
33	0.295	0.290	0.286	0.282	0.278	0.273	0.269	0.265	0.260	0.256	0.251					
34	0.291	0.287	0.283	0.279	0.275	0.271	0.267	0.262	0.258	0.254	0.249					
35	0.288	0.284	0.280	0.276	0.272	0.268	0.264	0.260	0.256	0.252	0.248					
36	0.285	0.281	0.277	0.273	0.270	0.266	0.262	0.258	0.254	0.250	0.246	0.242				
37	0.281	0.278	0.274	0.271	0.267	0.263	0.260	0.256	0.252	0.248	0.244	0.240				
38	0.278	0.275	0.272	0.268	0.265	0.261	0.257	0.254	0.250	0.246	0.242	0.239				
39	0.275	0.272	0.269	0.266	0.262	0.259	0.255	0.252	0.248	0.244	0.241	0.237	0.233			
40	0.273	0.269	0.266	0.263	0.260	0.256	0.253	0.250	0.246	0.243	0.239	0.236	0.232			
41	0.270	0.267	0.264	0.261	0.257	0.254	0.251	0.248	0.244	0.241	0.238	0.234	0.231			
42	0.267	0.264	0.261	0.258	0.255	0.252	0.249	0.246	0.242	0.239	0.236	0.233	0.229	0.226		
43	0.265	0.262	0.259	0.256	0.253	0.250	0.247	0.244	0.241	0.238	0.234	0.231	0.228	0.225		
44	0.262	0.259	0.256	0.254	0.251	0.248	0.245	0.242	0.239	0.236	0.233	0.230	0.227	0.223		
45	0.260	0.257	0.254	0.251	0.249	0.246	0.243	0.240	0.237	0.234	0.231	0.228	0.225	0.222	0.219	
46	0.257	0.255	0.252	0.249	0.247	0.244	0.241	0.238	0.236	0.233	0.230	0.227	0.224	0.221	0.218	
47	0.255	0.252	0.250	0.247	0.245	0.242	0.239	0.237	0.234	0.231	0.228	0.225	0.223	0.220	0.217	
48	0.253	0.250	0.248	0.245	0.243	0.240	0.237	0.235	0.232	0.230	0.227	0.224	0.221	0.218	0.216	0.213
49	0.250	0.248	0.246	0.243	0.241	0.238	0.236	0.233	0.231	0.228	0.225	0.223	0.220	0.217	0.215	0.212
50	0.248	0.246	0.244	0.241	0.239	0.236	0.234	0.232	0.229	0.226	0.224	0.221	0.219	0.216	0.213	0.211

Table 6. Lower critical values of the exact distribution of the sample autocorrelation $r_k(\ )$ of a white noise and the differences d between the various approximate distributions and the exact one ([approximate - exact]*1000 rounded to the nearest integer) for n = 10, 20, 50, 100, and k = 1, 3, 7, 15, 20.

			Levels for lower critical values															
			2.5%				**5%**				**95%**				**97.5%**			
k	**Distribution**	***n***	10	20	50	100	10	20	50	100	10	20	50	100	10	20	50	100
1	Exact		-0.600	-0.447	-0.287	-0.203	-0.533	-0.388	-0.245	-0.172	0.347	0.293	0.206	0.152	0.423	0.354	0.248	0.183
	AM		**41**	**18**	**5**	**2**	**41**	**17**	**5**	**2**	**-41**	**-17**	**-5**	**-2**	**-45**	**-19**	**-5**	**-2**
	Dufour-Roy		-90	-32	-8	-3	-64	-22	-5	-2	28	12	3	1	45	19	5	2
	Ljung-Box		**34**	**29**	**15**	**8**	**58**	**37**	**17**	**9**	128	58	22	11	143	64	23	11
	Box-Pierce		-20	**8**	**10**	**6**	**13**	**20**	**13**	**8**	173	75	26	12	197	84	29	13
	Exact (of $r_k(0)$)		**74**	**44**	**19**	**10**	**81**	**45**	**19**	**10**	105	50	20	10	103	49	20	10
3	Exact		-0.524	-0.421	-0.281	-0.200	-0.465	-0.365	-0.240	-0.170	0.330	0.280	0.202	0.151	0.403	0.338	0.244	0.181
	AM		**31**	**16**	**5**	**2**	**30**	**14**	**4**	**1**	**-48**	**-18**	**-5**	**-2**	**-54**	**-20**	**-6**	**-2**
	Dufour-Roy		-88	-32	-8	-3	-66	-24	-6	-2	**-20**	4	2	1	**-13**	10	4	2
	Ljung-Box		-42	**3**	**9**	**6**	-9	**14**	**12**	**7**	145	71	26	12	163	80	28	13
	Box-Pierce		-96	-17	**4**	**4**	-55	-3	**7**	**6**	191	88	30	14	217	100	34	15
	Exact (of $r_k(0)$)		**59**	**40**	**18**	**10**	**65**	**41**	**19**	**10**	71	44	19	10	62	43	19	10
7	Exact			-0.365	-0.268	-0.196		-0.317	-0.229	-0.166		0.254	0.195	0.148		0.306	0.234	0.178
	AM			**12**	**4**	**1**		**11**	**3**	**1**		**-15**	**-6**	**-2**		**-16**	**-6**	**-2**
	Dufour-Roy			-35	-9	-3		-27	-7	-2		**-16**	0	1		**-12**	2	1
	Ljung-Box			-53	-4	**2**		-34	**1**	**4**		96	34	15		112	38	17
	Box-Pierce			-74	-9	0		-51	-4	**2**		114	38	17		132	43	18
	Exact (of $r_k(0)$)			**31**	**17**	**9**		**33**	**17**	**9**		30	17	9		27	17	9
15	Exact				-0.241	-0.187			-0.205	-0.159			0.178	0.142			0.213	0.170
	AM				**3**	**1**			**2**	**1**			**-6**	**-2**			**-7**	**-2**
	Dufour-Roy				-10	-3			-9	-3			**-5**	0			**-3**	0
	Ljung-Box				-31	-7			-23	-4			50	21			58	24
	Box-Pierce				-37	-9			-27	-6			55	23			64	26
	Exact (of $r_k(0)$)				**14**	**8**			**14**	**8**			14	8			13	8
20	Exact					-0.181				-0.154				0.138				0.165
	AM					**1**				**1**				**-2**				**-2**
	Dufour-Roy					-4				-3				**-1**				**-1**
	Ljung-Box					-13				-9				25				29
	Box-Pierce					-15				-11				27				31
	Exact (of $r_k(0)$)					**8**				**8**				8				8

NOTE. Boldface identifies entries for which the test of autocorrelation is non conservative.

Table 7. Significance levels for a two-sided test based on the exact and approximate distributions for the first difference of annual deflated aggregate for undistributed income in corporate trading enterprises (n = 18)

Lag	Autocorrelation	Exact	AM appr.	Dufour-Roy	Ljung-Box	Exact (of rk(0))	Box-Pierce
1	0.409	0.030	0.021	0.040	0.068	0.058	0.083
2	-0.306	0.233	0.205	0.262	0.171	0.151	0.194
3	-0.386	0.095	0.078	0.124	0.084	0.056	0.101
4	-0.072	0.903	0.905	0.949	0.748	0.724	0.760
5	-0.052	0.969	0.976	0.972	0.817	0.792	0.826
6	-0.158	0.531	0.506	0.598	0.479	0.398	0.502

Table 8. Significance levels for a two-sided test based on the exact and approximate distributions for LOG FORWARD-LOG SPOT prediction errors (February 1921-August 1923) (n = 31)

Lag	Autocorrelation	Exact	AM appr.	Dufour-Roy	Ljung-Box	Exact (of rk(0))	Box-Pierce
1	0.252	0.098	0.087	0.106	0.148	0.144	0.161
2	0.127	0.353	0.335	0.355	0.466	0.459	0.479
3	0.168	0.236	0.217	0.236	0.334	0.317	0.349
4	-0.174	0.379	0.367	0.400	0.318	0.291	0.333
5	-0.105	0.637	0.633	0.662	0.547	0.518	0.560
6	-0.029	0.990	0.999	0.980	0.866	0.854	0.870
7	0.100	0.418	0.393	0.395	0.566	0.521	0.578
8	-0.071	0.765	0.764	0.806	0.683	0.642	0.693
9	0.163	0.209	0.188	0.190	0.349	0.273	0.364
10	0.192	0.140	0.119	0.124	0.271	0.185	0.286

Table 9. Provost&Rudiuk (exact(PR)), exact (this paper), and empirical distribution values taken at points $r = -0.5(0.1)0.4$. The empirical distribution is based on 5000000 series of observations from independent standard normal variables ($n = 7$ and $k = 2$).

r	-0.5	-0.4	-0.3	-0.2	-0.1	0.0	0.1	0.2	0.3	0.4
Exact (PR)	0.0817	0.1619	0.2642	0.3984	0.5372	0.6698	0.7835	0.8703	0.9297	0.9655
Exact (this paper)	0.0831	0.1623	0.2684	0.3986	0.5374	0.6715	0.7845	0.8708	0.9303	0.9657
Empirical	0.0832	0.1623	0.2681	0.3984	0.5374	0.6712	0.7844	0.8706	0.9302	0.9657

Table 10. Lower critical values of the exact distribution of the sample autocorrelation $r_k(\ )$ of a white noise and the differences d between simulation results over 10000 experiments and the normal exact one ([simulated - normal]*1000 rounded to the nearest integer) for n = 10, 20, 30, 40, and k = 1, 3, 7, 15, 20.

			Levels for lower critical values															
			2.5%				**5%**				**95%**				**97.5%**			
k	**Distribution**	***n***	10	20	30	40	10	20	30	40	10	20	30	40	10	20	30	40
1	Exact distribution		-0.600	-0.447	-0.369	-0.321	-0.533	-0.388	-0.318	-0.275	0.347	0.293	0.253	0.226	0.423	0.354	0.306	0.273
	Normal series		1	-1	2	3	4	-3	-1	0	-9	-5	-1	1	-14	-4	-5	-4
	Uniform series		**-6**	**-7**	**-11**	**-6**	**-11**	**-11**	**-6**	**-2**	**12**	**5**	**6**	**4**	**12**	**7**	**3**	-1
	Double exponential series		22	13	10	6	12	14	9	7	-32	-17	-9	-11	-36	-10	-13	-10
	Cauchy series		93	73	63	47	88	82	71	66	-89	-80	-79	-80	-84	-75	-70	-66
	Student with 2 d.f. series		39	39	24	19	38	36	23	21	-41	-40	-26	-21	-46	-30	-22	-18
	Logistic series		7	6	6	5	7	5	4	4	-21	-9	-5	-4	-21	-7	-9	-5
3	Exact		-0.524	-0.421	-0.355	-0.312	-0.465	-0.365	-0.306	-0.267	0.330	0.280	0.246	0.221	0.403	0.338	0.296	0.266
	Normal series		6	4	2	-1	8	1	1	-2	-1	-2	5	-10	-7	1	8	3
	Uniform series		0	**-5**	**-8**	**-15**	**-3**	**-6**	1	**-3**	**14**	**8**	**12**	**9**	**23**	**13**	**14**	**6**
	Double exponential series		19	14	10	6	20	14	11	7	-24	-12	-3	-2	-33	-12	**3**	-3
	Cauchy series		53	65	56	50	57	76	67	71	-92	-82	-70	-70	-89	-60	-49	-46
	Student with 2 d.f. series		27	27	21	24	31	29	27	25	-53	-31	-27	-26	-60	-29	-18	-18
	Logistic series		9	8	4	0	14	5	5	1	-10	-6	**3**	**2**	-16	-7	**4**	**1**
7	Exact				-0.326	-0.294			-0.280	-0.243			0.230	0.210			0.277	0.253
	Normal series				-2	0			-3	-11			1	1			3	6
	Uniform series				**-6**	**-6**			**-6**	**-14**			**6**	**6**			**12**	**12**
	Double exponential series				8	6			2	**-3**			-6	-3			-6	-1
	Cauchy series				39	45			62	52			-69	-64			-55	-47
	Student with 2 d.f. series				15	12			24	10			-30	-26			-26	-22
	Logistic series				2	3			**-1**	**-8**			-3	0			-2	**1**

NOTE. Boldface identifies entries for which the test of autocorrelation using the normal exact distributionn is non conservative according to the simulations.

Figure 1. Exact and approximate distributions for *n* = *10* and *k* = *1*

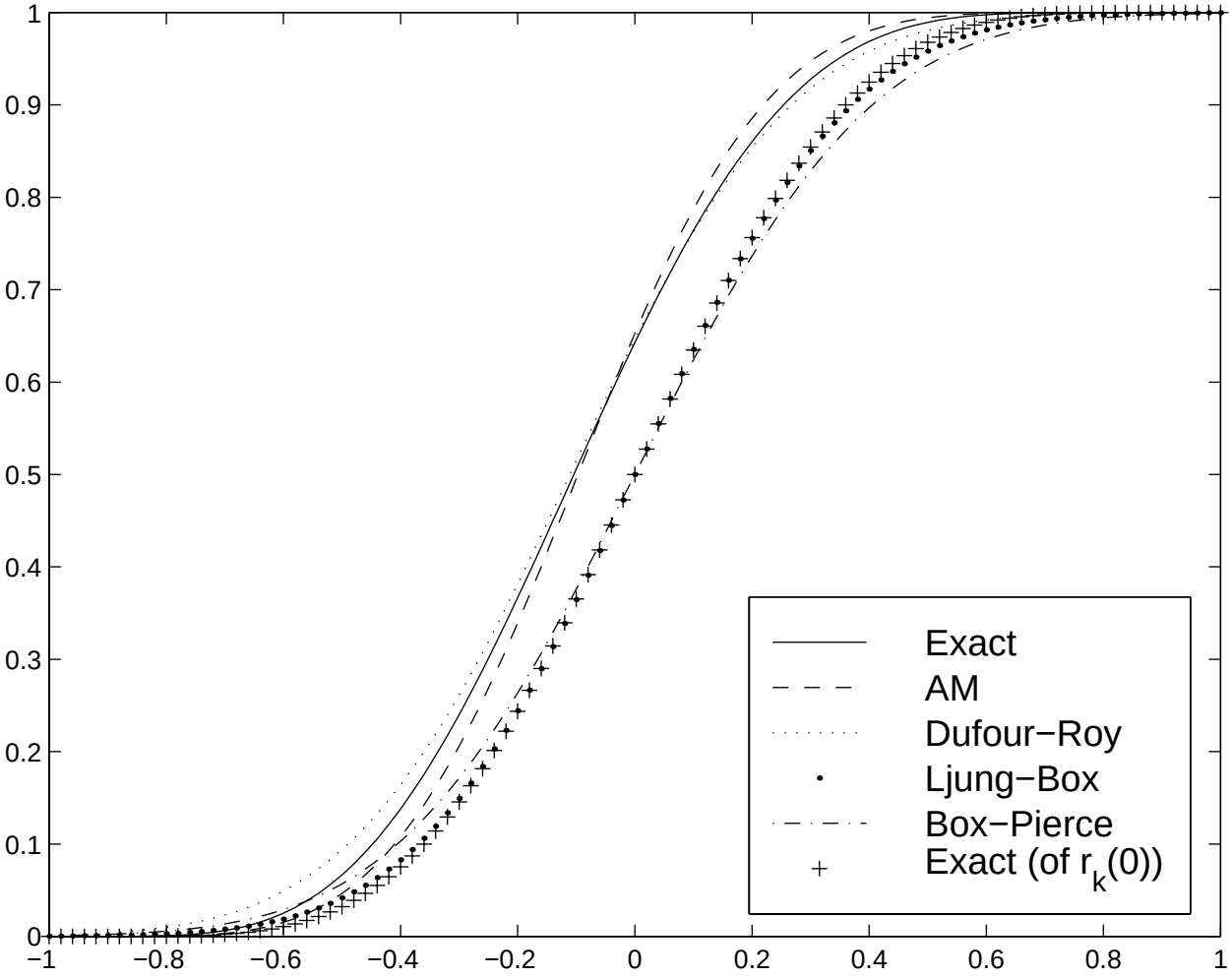


Figure 2: MAD (multiplied by 10^3) across sample sizes n = max(3k,4),...,50. Signif. level =0.025

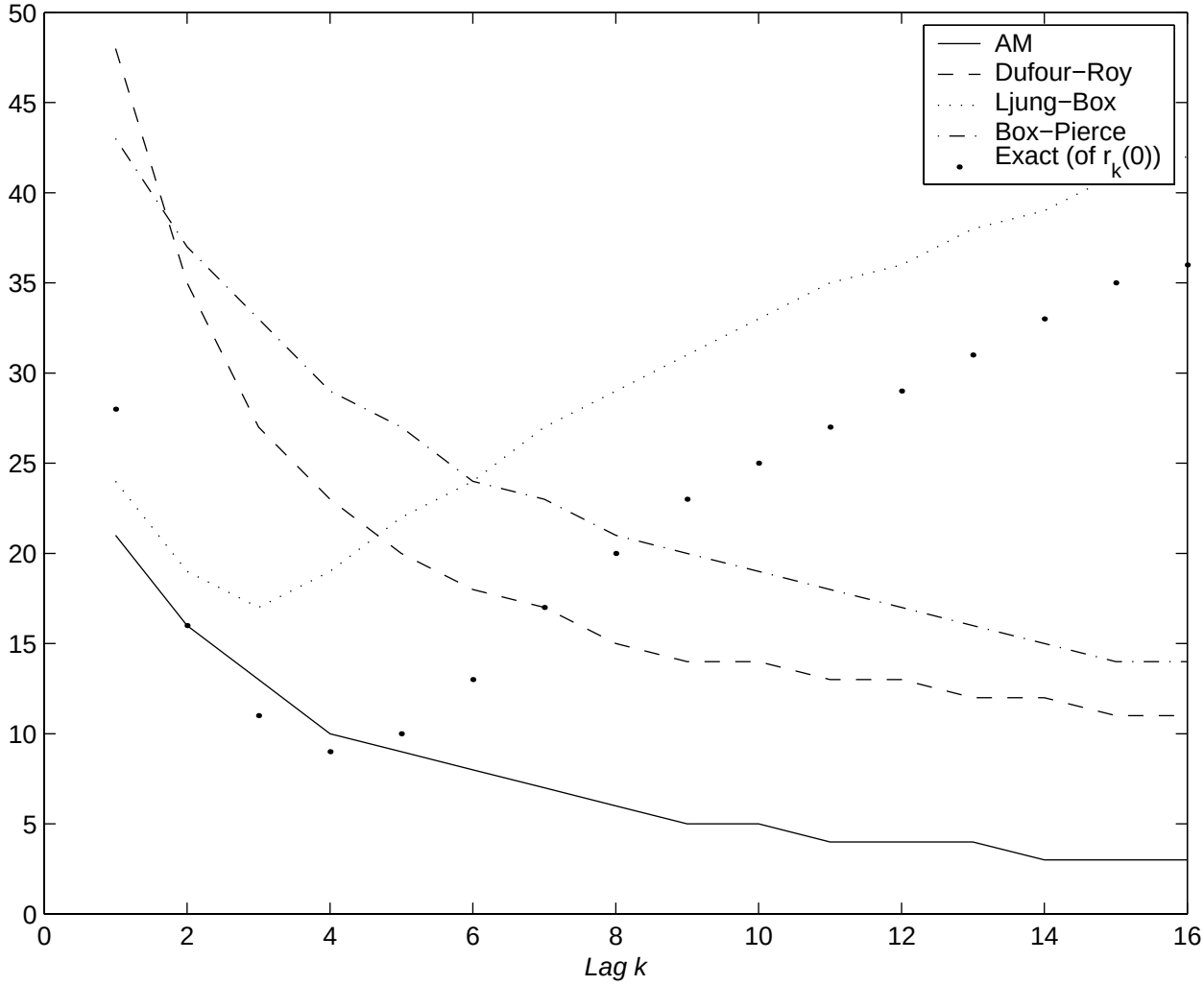


Figure 3: MAD (multiplied by 10^3) across sample sizes n = max(3k,4),...,50. Signif. level =0.975

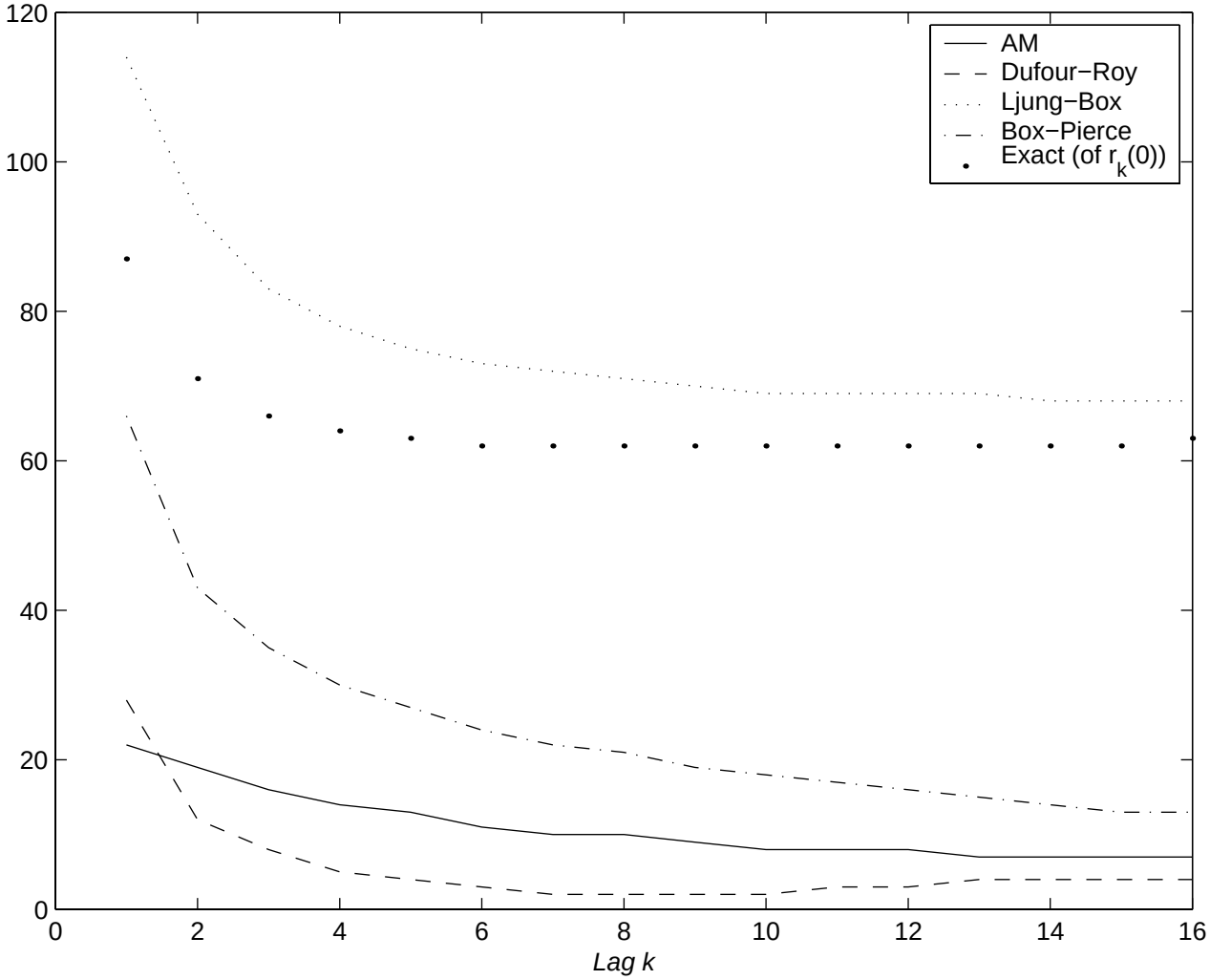


Figure 4 : MAD (multiplied by 10^3) across lags k=1,...,[n/3]. Signif. level =0.025

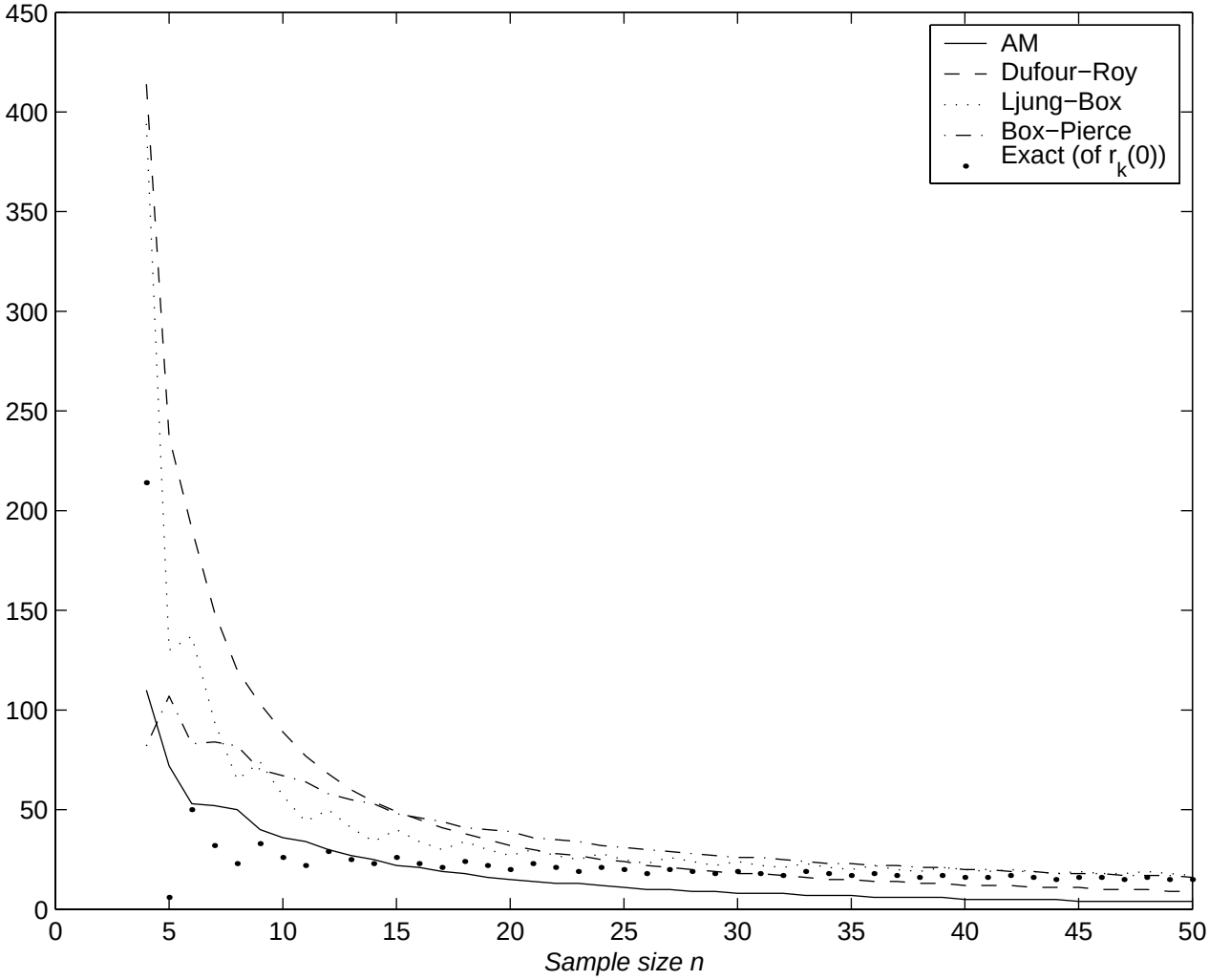


Figure 5 : MAD (multiplied by 10^3) across lags k=1,...,[n/3]. Signif. level =0.975