

# On confidence intervals and tests for autocorrelations

Guy MÉLARD

*Institut de Statistique, Université Libre de Bruxelles, Belgium*

Roch ROY

*Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, Canada*

Received 11 October 1985

Revised 30 September 1986

**Abstract:** We propose a statistical procedure for estimating the asymptotic variances and covariances of sample autocorrelations from a stationary time series so that confidence regions and tests on a finite subset of autocorrelations can be implemented. The corresponding algorithm is described. The accuracy of the asymptotic confidence intervals for finite samples is studied by Monte Carlo simulations. Further, our method is illustrated with examples from the literature.

**Keywords:** Stationary time series, Autocorrelations, Confidence intervals, Small sample properties, Simulation, Algorithm.

## 1. Introduction

A question that is sometimes addressed in time series analysis is the following: given a subset of autocorrelations, are they significantly different from zero, or from given values? Routinely the asymptotic standard error of the autocorrelations is used, under the assumption of either a white noise or a moving average process. More general expressions for variances and covariances of autocorrelations go back to Bartlett [3] but rely on the knowledge of the whole autocorrelation function. The purposes of this paper are: first, to provide a sound asymptotic procedure for estimating these variances and covariances so that confidence regions and tests can be implemented; second, to check the validity of the method for small samples; third, to illustrate its use.

Tests and confidence intervals on autocorrelation coefficients are useful both at the identification and checking stages of the iterative ARIMA model building procedure. For example, at the identification stage, the knowledge of the autocorrelation coefficients which are significant can help to specify a more parsimonious model. At the checking stage, an alternative way of assessing the adequacy of

a fitted model consists in testing whether the sample autocorrelations are compatible with the theoretical autocorrelation function of that model.

Let  $D = \{j_1, \dots, j_J\}$  be the finite set of lags in which we are interested, and  $\rho = (\rho_{j_1}, \dots, \rho_{j_J})'$  the corresponding vector of autocorrelation coefficients. We want either to test the hypothesis  $H_0: \rho = \rho_0$ , where  $\rho_0$  is a given  $J \times 1$  vector, or to determine a confidence region for  $\rho$ . The method described in Section 2 is based on a quadratic form in  $r$ , the vector of sample autocorrelations associated with  $\rho$ . The main difficulty is to obtain a consistent nonnegative definite estimator of the covariance matrix of  $r$ . Such an estimator for the autocorrelations is obtained by adapting the method developed by M  lard and Roy [12,13] for the autocovariances. The special case  $J = 1$  was handled by Robinson [15] in the time domain and by Brillinger [5] in the frequency domain.

The theoretical aspects of our method are described in Section 2. The algorithm is given in Section 3. In Section 4, results from Monte Carlo simulations are discussed. The purpose is to study the accuracy of the asymptotic nonparametric (in the sense model-free) confidence intervals for finite samples. These intervals are also compared with the exact asymptotic confidence intervals and with the ones obtained under the assumption of a moving average process. Two time series taken from the literature are used to illustrate the usefulness of our method.

## 2. Inference on a finite subset of autocorrelations

Consider a time series  $X_1, \dots, X_n$ , of length  $n$ , generated from a stationary stochastic process  $\{X_t: t \in \mathbb{Z}\}$  satisfying the condition

$$\sum_{j=-\infty}^{\infty} \gamma_j^2 < \infty$$

where  $\gamma_j = \text{Cov}(X_t, X_{t+j})$  is the autocovariance at lag  $j$ . We also suppose that the cumulants of order 4 of the process are zero. Hence the class of processes considered includes the stationary Gaussian processes. Let  $\rho_j = \gamma_j/\gamma_0$  denote the autocorrelation at lag  $j$  which is estimated by  $r_j^{(n)} = c_j^{(n)}/c_0^{(n)}$ , where

$$c_j^{(n)} = (1/n) \sum_{t=1}^{n-j} (X_t - \bar{X})(X_{t+j} - \bar{X}), \quad 0 \leq j \leq n-1, \quad (2.1)$$

and  $\bar{X} = (1/n) \sum_{t=1}^n X_t$  is the sample mean. We have  $\rho_{-j} = \rho_j$  and we take  $c_{-j}^{(n)} = c_j^{(n)}$  and  $r_{-j}^{(n)} = r_j^{(n)}$ ,  $j > 0$ .

We know from Anderson [1], p. 489, that the asymptotic covariance between  $r_i$  and  $r_j$  is given by

$$\lim_{n \rightarrow \infty} n \text{Cov}(r_i, r_j) = \sigma_{ij}, \quad i, j \in \mathbb{Z},$$

where

$$\sigma_{ij} = (\lambda_{i+j} + \lambda_{i-j} - 2\gamma_i \lambda_j / \gamma_0 - 2\gamma_j \lambda_i / \gamma_0 + 2\gamma_i \gamma_j \lambda_0 / \gamma_0^2) / \gamma_0^2 \quad (2.2)$$

and

$$\lambda_i = \sum_{k=-\infty}^{\infty} \gamma_k \gamma_{k+i}, \quad i \in \mathbb{Z}. \quad (2.3)$$

Note that  $\lambda_{-i} = \lambda_i$ ,  $i > 0$ . Furthermore, the asymptotic covariance (2.2) remains valid without any assumption on the cumulants of order 4. However, that assumption is used to derive consistent estimators of the  $\gamma_j$ 's and of the  $\lambda_i$ 's along the line followed by Mélard and Roy [13].

Denote by  $\mathbf{r}^{(n)}$  the column vector whose components are  $r_{j_i}^{(n)}$ ,  $i = 1, \dots, J$ ,  $J$  being independent of  $n$ , and let  $D = \{j_1, \dots, j_J\} \subset \{1, \dots, n-1\}$ . Under fairly general conditions (see Anderson [1], p. 489, Hannan and Heyde [7]), we know that  $\sqrt{n}(\mathbf{r}^{(n)} - \boldsymbol{\rho})$  is asymptotically multinormal  $N(\mathbf{0}, \boldsymbol{\Sigma})$  where  $\boldsymbol{\rho}$  is the column vector  $(\rho_i : i \in D)$  and  $\boldsymbol{\Sigma}$  is the  $J \times J$  matrix whose elements are  $\sigma_{ij}$ ,  $i, j \in D$ . Since  $\boldsymbol{\Sigma}$  is unknown in general, we will replace it by  $\hat{\boldsymbol{\Sigma}}^{(n)}$  whose elements  $\hat{\sigma}_{ij}$  converge in probability to  $\sigma_{ij}$ .

Under the assumption that the underlying process is  $MA(q)$ , the estimation of  $\boldsymbol{\Sigma}$  is much simpler. Indeed, the covariance matrix  $\boldsymbol{\Sigma}$  is a function of  $\rho_1, \dots, \rho_q$  only and it is natural to replace them by their consistent estimators  $r_1^{(n)}, \dots, r_q^{(n)}$ . This method provides a consistent estimator  $\hat{\boldsymbol{\Sigma}}$  of  $\boldsymbol{\Sigma}$  but it does not insure that  $\hat{\boldsymbol{\Sigma}}$  is nonnegative definite. We know from Priestley [14], p. 437, that the sequence  $(\delta_j : j \in \mathbb{Z})$  where  $\delta_j = 1$ ,  $|j| \leq q$  and  $\delta_j = 0$ ,  $|j| > q$  is not nonnegative definite. Therefore we cannot assert that the sequence  $(\delta_j r_j : j \in \mathbb{Z})$  is nonnegative definite (see Loève [9], p. 134). In general, the construction of a nonnegative definite estimator of  $\boldsymbol{\Sigma}$  is not immediate.

Mélard and Roy [13] studied a similar problem concerning the sample autocovariances. In this latter case, the elements of the matrix  $\boldsymbol{\Sigma}$  have the simpler form  $\lambda_{i+j} + \lambda_{i-j}$  rather than (2.2). Robinson [15] had proposed an estimator of  $\lambda_i$  which converges in the  $L_1$  norm. The present authors modified that estimator in order to preserve the nonnegative definiteness of the matrix  $\hat{\boldsymbol{\Sigma}}$ . The latter estimator is obtained by substituting  $w_i^{(n)} c_i^{(n)}$  to  $\gamma_i$  into (2.3) where  $(w_i^{(n)} = w(i/b_n) : i \in \mathbb{Z})$  is a positive definite sequence; the function  $w : \mathbb{R} \rightarrow \mathbb{R}$  is continuous at the origin,  $w(0) = 1$ , has at most a finite number of discontinuities and is square integrable and  $(b_n : n \geq 1)$  is a sequence of real numbers such that  $b_n \rightarrow \infty$  in such a way that  $b_n/n \rightarrow 0$  when  $n \rightarrow \infty$ .

Consider now the matrix  $\boldsymbol{\Sigma}$  whose elements  $\sigma_{ij}$  are given by (2.2). This matrix is at least nonnegative definite for any autocovariance function  $(\gamma_i : i \in \mathbb{Z})$ . Since the sequence  $(w_i^{(n)} c_i^{(n)} : i \in \mathbb{Z})$  is also an autocovariance function (Loève [9], p. 134), the substitution of  $w_i^{(n)} c_i^{(n)}$  to  $\gamma_i$  preserves the nonnegative definiteness of the resulting matrix  $\hat{\boldsymbol{\Sigma}}^{(n)}$ . It also leads to a consistent estimator. Indeed, from Theorem 1 of Mélard and Roy [13], the estimator  $\hat{\lambda}_i^{(n)}$  obtained by the same substitution in (2.3) converges in probability to  $\lambda_i$ . Since  $w_i^{(n)} \rightarrow 1$  as  $n \rightarrow \infty$  and that  $c_i^{(n)} \rightarrow \gamma_i$  in probability, it follows that  $w_i^{(n)} c_i^{(n)} \rightarrow \gamma_i$  in probability. The element  $\sigma_{ij}$  being a continuous function in all its arguments, the estimator  $\hat{\sigma}_{ij}$  converges also in probability to  $\sigma_{ij}$ .

**Remark.** We have not proved that  $\hat{\Sigma}^{(n)}$  is positive definite but only that it is nonnegative definite. Therefore, in the next section, the inverse of  $\hat{\Sigma}^{(n)}$  will be taken in the generalized sense.

### 3. Algorithm for computing confidence intervals and test statistics

The following algorithm relies upon the results of the previous section. It consists in the following steps. The superscript  $n$  will be omitted unless it is essential for clarity.

- A. Compute  $c_j = c_{-j}$  by formula (2.1),  $j = 0, 1, \dots, n-1$ .
- B. Choose a window  $w_i = w(i/b_{n,H})$ ,  $|i| \leq n-1$ , with  $b_{n,H} = H\sqrt{n}$ .
  - B.1. Modified Bartlett window:

$$w(x) = \begin{cases} 1 - |x|, & |x| \leq 1, \\ 0, & |x| > 1. \end{cases}$$

- B.2. Parzen window:

$$w(x) = \begin{cases} 1 - 6x^2 + 6|x|^3, & |x| \leq \frac{1}{2}, \\ 2(1 - |x|)^3, & \frac{1}{2} \leq |x| \leq 1, \\ 0, & |x| > 1. \end{cases}$$

- C. Compute for  $i = 0, 1, \dots, n-1$ :

$$\hat{\lambda}_i = \hat{\lambda}_{-i} = \sum_{k=-n+1}^{n-i-1} w_k w_{k+i} c_k c_{k+i}.$$

Note that  $\hat{\lambda}_i = 0$  for  $|i| \geq n$ .

- D. Compute for  $i, j \in D$  (a given subset of  $\{1, \dots, n-1\}$ ):

$$\hat{\sigma}_{ij} = (\hat{\lambda}_{i+j} + \hat{\lambda}_{i-j} - 2w_i c_i \hat{\lambda}_i / c_0 - 2w_j c_j \hat{\lambda}_j / c_0 + 2w_i w_j c_i c_j \hat{\lambda}_0 / c_0^2) / c_0^2. \quad (3.1)$$

- E. To test the null hypothesis  $H_0: \rho = \rho_0$  against the alternative  $H_1: \rho \neq \rho_0$  at the significance level  $\alpha$ , compute

$$Q = n^{-1}(\mathbf{r} - \rho_0)' \hat{\Sigma}^{-1}(\mathbf{r} - \rho_0) \quad (3.2)$$

and reject  $H_0$  if  $Q > \chi_{j, 1-\alpha}^2$ , the quantile of order  $1 - \alpha$  of the  $\chi_j^2$  distribution.

- F. To build a confidence interval for  $\rho_j$  at the confidence level  $1 - \alpha$ , compute

$$r_j \pm u_{1-\alpha/2} (\hat{\sigma}_{jj}/n)^{1/2}$$

where  $u_{1-\alpha/2}$  is the quantile of order  $1 - \frac{1}{2}\alpha$  of the  $N(0, 1)$  distribution.

**Remarks.** 1. The choice of  $b_{n,H}$  is justified by the condition imposed on the sequence  $\{b_n\}$  in Section 2.

2. The use of the chi-square distribution at step E follows from an argument similar to the one evoked in Méléard and Roy [11]. As indicated in the remark at

the end of Section 2,  $\hat{\Sigma}^{-1}$  is the generalized inverse of  $\hat{\Sigma}$ . Therefore, the number of degrees of freedom of the chi-square distribution should be decreased by the number of zero eigenvalues of  $\hat{\Sigma}$ .

3. The use of the  $N(0, 1)$  distribution at step F relies on results of Section 2.

4. The test statistic computed at step E can serve either to test that a finite number of autocorrelations of the time series under investigation are zero or to test that the first few autocorrelations of the generating process are those of a completely specified ARMA  $(p, q)$  process. In the latter case, the autocorrelations of the ARMA process can be obtained with the algorithms of McLeod [10] or of Tunncliffe Wilson [16].

5. More generally than F, a confidence region for  $\rho$  is given by

$$n^{-1}(\mathbf{r} - \rho)' \hat{\Sigma}^{-1}(\mathbf{r} - \rho) \leq \chi_{J; 1-\alpha}^2.$$

The usefulness of that procedure is restricted to small values of  $J$ .

#### 4. Simulation results

Since the methodology described in the previous sections relies on asymptotic results, it is important to study the accuracy of the corresponding approximations for finite samples. To investigate this issue, we conducted the following Monte Carlo experiment. For three different processes: the moving average process of order 4 (MA(4)) with  $\theta_1 = \theta_2 = \theta_3 = 0$ ,  $\theta_4 = \theta = 0.5$ , the autoregressive process of order 1 (AR(1)) with  $\phi = 0.5$  and the autoregressive process of order 4 (AR(4)) with  $\phi_1 = \phi_2 = \phi_3 = 0$ ,  $\phi_4 = \phi = 0.5$ , the following eight intervals were compared for the series lengths  $n = 50$  and  $n = 100$ .

For the lag  $j$  autocorrelation  $\rho_j$ , these intervals are the following:

$I_1$ . Interval based on the exact asymptotic variance:

$$r_j \pm u_{1-\alpha/2} (\sigma_{jj}/n)^{1/2}$$

where  $\sigma_{jj}$  is given by (2.2).

For the MA(4) process, we find

$$\gamma_j = \begin{cases} (1 + \theta^2) \sigma_a^2, & j = 0, \\ -\theta \sigma_a^2, & j = -4, 4, \\ 0, & \text{otherwise,} \end{cases}$$

where  $\sigma_a^2$  denotes the variance of the innovation process and

$$\lambda_j = \begin{cases} \sigma_a^4 (1 + 4\theta^2 + \theta^4), & j = 0, \\ -2\sigma_a^4 \theta (1 + \theta^2), & j = -4, 4, \\ \sigma_a^4 \theta^2, & j = -8, 8, \\ 0, & \text{otherwise.} \end{cases}$$

For the AR(1) process, we have

$$\gamma_j = \frac{\sigma_a^2}{1 - \phi^2} \phi^{|j|}, \quad j \in \mathbb{Z},$$

and

$$\lambda_i = \frac{\sigma_a^4}{(1 - \phi^2)^2} \phi^i \left( \frac{1 + \phi^2}{1 - \phi^2} + i \right), \quad i \geq 0.$$

For the AR(4) process, we get

$$\gamma_j = \begin{cases} \frac{\sigma_a^2}{1 - \phi^2} \phi^{|k|}, & j = 4k, k \in \mathbb{Z}, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\lambda_i = \begin{cases} \frac{\sigma_a^4}{(1 - \phi^2)^2} \phi^k \left( \frac{1 + \phi^2}{1 - \phi^2} + k \right), & i = 4k, k \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

$I_2$ . Confidence interval for  $\rho_j$  under the assumption that the underlying process is MA( $j - 1$ ):

$$r_j \pm u_{1-\alpha/2} \hat{\sigma}_{Mj}$$

where

$$\sigma_{Mk}^2 = \frac{1}{n} \left( 1 + 2 \sum_{i=1}^{k-1} \rho_i^2 \right) \quad (4.1)$$

is the asymptotic variance of  $r_k$  for  $k \geq j$  and  $\hat{\sigma}_{Mk}^2$  is obtained by substituting  $r_i$  to  $\rho_i$  in (4.1).

In time series analysis, the intervals defined by the values  $\pm u_{1-\alpha/2} \hat{\sigma}_{Mj}$  are routinely computed for testing sequentially that the underlying process is MA( $j - 1$ ),  $1 \leq j \leq K \leq n - 1$ .

$I_i$ ,  $i = 3, 4, 5$ . These intervals are obtained at step F of the algorithm using the modified Bartlett window with  $b_{n,H} = H\sqrt{n}$ ,  $H = 1, 3$  and 5 respectively.

$I_i$ ,  $i = 6, 7, 8$ . These intervals are calculated in a similar way using the Parzen window.

For each model and for each series length, 500 independent realizations were generated with  $N(0, 1)$  innovations using the subroutine GGUBS of IMSL (1984). For each realization and for each lag  $k$  ( $1 \leq k \leq K$ ), the sample autocorrelation  $r_k$  and the confidence intervals  $I_i$ ,  $i = 1, \dots, 8$ , were calculated for three different nominal levels (80, 90 and 95 percent). For each confidence interval, we checked if it contained the exact value of the theoretical autocorrelation  $\rho_j$ . The accuracy of  $I_i$ ,  $i = 1, \dots, 8$ , was appreciated by comparing its empirical confidence levels to the corresponding nominal ones.

The results of the experiment are reported in Table 1 for the moving average process and in Tables 2 and 3 for the autoregressive processes. The entries in these tables are empirical confidence levels in percentage, i.e. the percentage among the computed confidence intervals which contain the true value of the

Table 1  
Empirical confidence levels (in percentage) of different confidence intervals for autocorrelations of a moving average process of order 4 with  $\theta_1 = \theta_2 = \theta_3 = 0$  and  $\theta_4 = 0.5$

Confidence level	Interval	$n = 50, k =$										$n = 100, k =$									
		1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
80	$I_1$	80	80	81	82	83	86	86	84	84	87	82	82	82	78	82	83	82	84	82	81
	$I_2$	75	94	77	92	84	87	89	87	89	90	77	95	76	92	83	84	84	87	85	85
	$I_3$	74	88	76	89	80	83	84	82	83	83	78	88	77	85	78	80	79	81	80	79
	$I_4$	78	85	79	86	82	85	86	83	83	85	82	85	80	82	81	82	83	84	82	83
	$I_5$	80	86	80	87	83	86	87	85	85	87	84	87	82	84	84	84	84	87	84	85
90	$I_1$	91	91	91	91	92	94	94	93	93	94	92	92	91	91	90	92	92	93	92	92
	$I_2$	85	99	88	97	93	95	96	95	96	96	87	98	87	97	92	93	93	95	94	94
	$I_3$	83	97	87	96	88	92	94	93	93	94	88	95	87	96	88	87	91	92	90	91
	$I_4$	87	94	90	93	89	91	93	93	93	93	89	93	90	93	89	90	92	94	91	92
	$I_5$	88	94	90	94	90	93	95	93	92	94	91	94	91	94	91	92	93	95	93	93
95	$I_1$	97	97	96	96	97	97	98	97	97	98	96	96	95	96	95	97	97	97	97	97
	$I_2$	92	100	94	99	98	99	99	98	99	99	93	99	92	100	96	98	98	97	98	98
	$I_3$	90	99	92	99	95	97	98	97	97	98	93	97	92	98	93	95	95	96	96	96
	$I_4$	93	98	94	97	94	97	97	96	96	97	95	96	94	97	95	97	96	96	97	97
	$I_5$	94	99	95	97	95	97	98	96	96	97	97	97	95	98	95	97	96	97	97	97

The standard error of the confidence levels is 1.8% for the nominal level 80%, 1.3 for 90 and 1.0 for 95.

Table 2  
Empirical confidence levels (in percentage) of different confidence intervals for autocorrelations of an autoregressive process of order 1 with  $\phi = 0.5$

Confidence level	Interval	$n = 50, k =$										$n = 100, k =$									
		$n = 50, k =$					$n = 50, k =$					$n = 100, k =$					$n = 100, k =$				
		1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
80	$I_1$	74	78	81	80	85	86	88	89	86	88	78	79	78	80	79	83	83	84	84	82
	$I_2$	81	79	79	79	85	86	89	90	88	91	83	81	79	79	79	83	83	85	84	83
	$I_3$	74	72	75	76	80	81	86	86	84	88	75	76	75	75	74	80	81	83	81	80
	$I_4$	77	74	75	76	79	78	85	85	83	88	78	78	77	77	76	80	82	83	82	80
	$I_5$	78	77	77	77	80	81	87	86	83	89	80	80	78	79	77	82	83	84	83	80
90	$I_1$	84	86	90	92	94	95	95	95	94	95	86	89	90	88	91	93	92	92	93	91
	$I_2$	89	87	89	91	95	96	96	96	96	97	90	91	90	88	91	93	92	92	95	92
	$I_3$	84	84	86	87	91	93	95	96	94	95	85	85	85	86	87	90	89	90	93	90
	$I_4$	86	85	87	85	90	91	93	94	92	94	87	87	87	86	88	91	90	90	93	89
	$I_5$	88	86	88	86	91	92	93	94	92	94	88	88	88	88	89	92	91	90	93	90
95	$I_1$	91	92	95	96	97	98	98	99	98	98	91	94	96	95	97	96	97	96	97	95
	$I_2$	94	91	94	95	98	98	99	99	99	99	95	94	95	95	96	97	98	97	98	97
	$I_3$	90	88	91	93	96	97	98	99	99	99	91	91	92	91	94	95	96	95	97	95
	$I_4$	91	89	91	92	95	96	96	97	96	96	93	92	93	92	94	94	95	94	97	94
	$I_5$	92	90	92	93	95	97	97	97	96	97	94	93	94	94	95	95	95	96	97	94

The standard error of the confidence levels is 1.8% for the nominal level 80%, 1.3 for 90 and 1.0 for 95.



Table 3  
Empirical confidence levels (in percentage) of different confidence intervals for autocorrelations of an autoregressive process of order 4 with  $\phi_1 = \phi_2 = \phi_3 = 0$  and  $\phi_4 = 0.5$

Confidence level	Interval	$n = 50, k =$										$n = 100, k =$									
		1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
80	$I_1$	85	87	84	72	84	91	84	77	89	89	84	79	84	76	84	82	83	81	84	84
	$I_2$	73	61	77	83	84	81	85	81	90	88	71	55	75	84	81	73	82	86	86	83
	$I_3$	71	68	75	78	77	75	79	74	83	82	71	67	75	77	77	69	76	80	79	76
	$I_4$	77	76	77	75	79	79	81	77	85	84	77	74	79	78	80	75	79	81	83	80
	$I_5$	79	80	80	78	81	82	81	80	86	85	80	77	81	80	83	78	82	83	85	82
90	$I_1$	93	96	93	84	94	97	94	89	95	97	92	90	93	86	92	92	91	92	93	92
	$I_2$	85	72	87	92	94	93	95	90	97	97	83	67	87	92	91	86	92	93	95	94
	$I_3$	84	78	85	89	88	87	89	86	92	92	84	77	86	87	86	79	86	91	90	86
	$I_4$	89	87	88	89	90	90	88	88	92	92	89	84	90	89	90	86	89	92	92	89
	$I_5$	91	89	90	91	91	91	90	89	95	94	92	87	91	90	91	88	91	93	93	91
95	$I_1$	97	98	98	91	98	98	98	94	98	99	95	97	97	92	96	98	96	96	97	99
	$I_2$	90	83	93	95	98	98	98	94	99	99	90	74	91	95	95	93	97	96	98	98
	$I_3$	91	86	91	95	93	94	97	92	98	98	91	84	92	94	92	86	92	95	95	94
	$I_4$	96	92	93	94	94	94	95	93	96	97	94	91	95	95	94	91	95	96	96	95
	$I_5$	97	94	95	95	95	95	96	94	98	97	96	94	97	95	96	94	95	96	97	97

The standard error of the confidence levels is 1.8% for the nominal level 80%, 1.3 for 90 and 1.0 for 95.

parameter. We do not present the results for the Parzen window since they are very similar to those obtained with the modified Bartlett window. We make the following observations. For the three models, we notice that for each nominal level and for each interval, the corresponding empirical levels increase in general with the lag  $k$ . However, the increasing trend is less accentuated for  $n = 100$  than for  $n = 50$ . Also, the trend is more irregular with the AR(4) model. The interval  $I_1$  cannot be used in a practical situation since the unknown theoretical autocorrelations are involved. However, in the simulation experiment, it will give us an idea of the accuracy of the asymptotic variances for finite samples and will serve as a reference point.

From now on, we will discuss the three models separately. For the moving average model, the empirical levels of  $I_1$  are always greater than or equal to the nominal levels. The interval  $I_2$  is meaningful only for the lag  $k = 5$  and even for this particular lag, its empirical levels are significantly different from the nominal ones when  $n = 50$ . Among the nonparametric intervals  $I_3$ ,  $I_4$  and  $I_5$  considered, the interval  $I_5$  seems the best globally. However, for larger lags,  $I_3$  or  $I_4$  is sometimes better. For all three intervals, the corresponding empirical levels behave favorably well with respect to those of  $I_1$ . Further, their empirical levels are closer to the nominal ones than those of  $I_2$ . This latter observation emphasizes the importance of using nonparametric intervals. Also, the empirical levels of  $I_2$  are generally greater than those of  $I_3$ ,  $I_4$  and  $I_5$ . Therefore, a test based on one of these last three intervals will be in general more powerful than a test based on  $I_2$ . Finally, we see that the empirical levels of the nonparametric intervals are very often greater than the corresponding nominal levels. Hence, tests of hypotheses based on these confidence intervals will tend to be conservative.

With the AR(1) process, the empirical levels of  $I_1$  for the lags  $k = 1, 2$  are systematically smaller than the corresponding nominal levels. The interval  $I_2$  has no statistical meaning for any lag  $k$  since it is based on the assumption that the process is MA( $k - 1$ ). However, its empirical levels are astonishingly close to the nominal levels. Among the nonparametric intervals,  $I_5$  seems also the best globally although for  $k \geq 5$ ,  $I_3$  and  $I_4$  give very similar results. The empirical levels of  $I_5$  are as close as or sometimes closer to the nominal levels than those of  $I_1$ .

With the AR(4) process, the empirical levels of  $I_1$  are quite different from the nominal ones. At lags 4 and 8, they are much smaller and at the other lags, they are systematically larger. As expected, the interval  $I_2$  behaves rather poorly, in particular at lag 2 where the empirical levels are much smaller than the nominal ones. Among the nonparametric intervals,  $I_5$  seems the best globally although at the 90% and 95% confidence levels,  $I_4$  gives comparable results. It is worthwhile noticing that the empirical levels of  $I_5$  are almost systematically smaller than those of  $I_1$ . Also, the empirical levels of  $I_5$  are in general closer to the nominal levels than those of  $I_3$ . With our method, not only the estimation of the standard error does not seem to reduce the accuracy of the confidence intervals but it also leads to shorter intervals in some cases.

## 5. Examples

In the following examples, unless specified, the modified Bartlett window with  $H = 5$  was used. Some preprocessing was needed since the series analyzed in our examples are nonstationary whereas our method applies only to stationary ones.

Let us first consider the airline passengers data (series G of Box and Jenkins, [4]). Using the transformation  $\nabla \nabla_{12} \log X_t$ , where  $\nabla$  and  $\nabla_{12}$  are respectively the ordinary and seasonal difference operators, the sample autocorrelation function is given in Table 4.

Box and Jenkins [4] proposed the multiplicative seasonal moving average model

$$\nabla \nabla_{12} \log X_t = A_t - \theta A_{t-1} - \Theta A_{t-12} + \theta \Theta A_{t-13}, \quad (5.1)$$

where the  $A_t$ 's are independent normal random variables with mean 0 and variance  $\sigma^2$ . The absence of the terms  $A_{t-2}, A_{t-3}, \dots, A_{t-11}$  in (5.1) can be tested by the method of Section 3. Indeed, it is equivalent to test the hypothesis that  $\rho_2 = \rho_3 = \dots = \rho_{10} = 0$ . The  $Q$  statistic (3.2) with  $D = \{2, 3, \dots, 10\}$  is equal to 13.6, which corresponds to a probability of significance of 0.19 for the chi-square distribution with 9 degrees of freedom. Hence, there is no reason to reject the hypothesis at the usual probability levels. With  $H = 1$  and  $H = 3$  for the modified Bartlett window, we have respectively  $Q = 15.4$  and  $Q = 14.0$ . With the Parzen window, the conclusions are the same.

It sometimes happens that an autocorrelation coefficient is very high but we are not sure that it reveals something special about the data. For example, let us consider the Chatfield and Prothero [6] series whose length is 77. For the series  $\nabla \nabla_{12} \log X_t$ ,  $r_{11} = 0.44$  is found. The standard error of  $r_{11}$  computed by (3.1) with  $i = j = 11$  is equal to 0.155 for  $H = 5$  and 0.161 for  $H = 3$ . Hence  $r_{11}$  is significantly different from 0. Note that the standard error of  $r_{11}$  under the hypothesis of a MA(10) model is equal to 0.189. There is less evidence of an autocorrelation at lag 11 by that more frequently used method. Table 5 contains the confidence intervals obtained by (3.1) for the lags  $k$  from 1 to 12, at the confidence level 0.95.

If we test  $H_0: \rho_2 = \rho_3 = \dots = \rho_{10} = 0$  on the same transformed series, we find  $Q = 34.5$  with a probability of significance of about 0.0001. Thus, a model of the

Table 4  
Autocorrelations  $r_j$  of  $\nabla \nabla_{12} \log X_t$  for the airline data

$j = 1$	-0.34	$j = 9$	0.18
2	0.11	10	-0.08
3	-0.20	11	0.06
4	0.02	12	-0.39
5	0.06	13	0.15
6	0.03	14	-0.06
7	-0.06	15	0.15
8	0.00		

Table 5

Confidence intervals for  $\rho_k$ , at the 0.95 confidence level, for the Chatfield–Prothero data

$k = 1$	$-0.58 \pm 0.23$	$k = 7$	$-0.17 \pm 0.32$
2	$0.36 \pm 0.30$	8	$-0.03 \pm 0.36$
3	$-0.22 \pm 0.37$	9	$0.10 \pm 0.35$
4	$0.05 \pm 0.37$	10	$-0.26 \pm 0.34$
5	$-0.05 \pm 0.29$	11	$0.44 \pm 0.30$
6	$0.10 \pm 0.26$	12	$-0.36 \pm 0.32$

Table 6

Test on four models for the Chatfield and Prothero data, with  $W_t = \nabla \nabla_{12} \log X_t$ ;  $P$  is the probability of significance of the  $Q$  statistic

Model	$Q$	$P$
(A) $W_t = -0.47W_{t-1} + A_t - 0.81A_{t-12}$	755	0
(B) $W_t = -0.51W_{t-1} - 0.47W_{t-12} + 0.24W_{t-13} + A_t$	28.3	0.013
(C) $W_t = -0.56W_{t-1} + A_t - 0.49A_{t-1}$	74.7	0
(D) $W_t = A_t - 0.44A_{t-1} - 0.85A_{t-12} + 0.37A_{t-13}$	67.3	0

form (5.1) is rejected. Chatfield and Prothero [6] fitted however that model as well as three other ones. These models are given in Table 6. We shall illustrate the use of the method of Section 3 at the checking stage of the model building procedure. In that purpose, we compare the sample autocorrelations of the transformed data with the autocorrelation function of each of these models by means of the  $Q$  test (3.2) with  $D = \{1, 2, \dots, 13\}$ . The results in Table 6 show that all the models are rejected except possibly model B, if a significance level of 1% or less is used. This is not surprising since Chatfield and Prothero finally retained model B, essentially because of its predictive capability. Obviously none of these models are compatible with the large value of  $r_{11}$ . Mélard [11] provides a deeper analysis of that series. Note that the tests should have been performed with two independent realizations, the first one giving the parameter estimates. This is not possible here because the series is too short.

## 6. Conclusion

The use of formal inference for autocorrelations has often been restricted to some specific cases such as moving average processes. There are at least two reasons for that. First, estimating the variance of autocorrelations of a general stationary process is cumbersome. Second, that variance depends on the autocorrelations themselves, so it has to be estimated in some consistent way. The use of more than one autocorrelation has been restricted except in the case of port-manteau statistics. This is probably due to the complex correlation patterns between autocorrelation estimators.

In this paper, we propose sound asymptotic methods for confidence intervals and tests for autocorrelation coefficients of general stationary processes. Examples illustrate how to apply the proposed methods. Simulation results indicate that the probability levels of the confidence intervals introduced are quite close to their nominal values, and that the corresponding significance tests are generally conservative.

In the course of the simulation experiment, we observed that deviations from the theoretical values are mainly due to the fact that sample autocorrelations are biased estimators. However, their standard deviations are generally well estimated. This explains why the empirical confidence levels of the intervals  $I_1$  and  $I_5$  are often closer to each other than they are to the corresponding nominal confidence level. A summary of the main results on the bias of sample autocorrelations is given in Kendall, Stuart and Ord [8], Chapter 48.

Another improvement of our method would result from using a more accurate formula than (2.2) for small and moderate sample sizes. Anderson and De Gooijer [2] derived an exact formula for the covariance between two sample autocovariances. Of course, an approximation occurs when passing from autocovariances to autocorrelations. Further work is thus needed on the distribution of sample autocorrelations.

## Acknowledgment

This work was supported by the Cooperation between the Province of Quebec and the French Community of Belgium, the Natural Sciences and Engineering Research Council of Canada and the Foundation FCAR (Government of Quebec). The authors thank Normand Ranger for his programming assistance in the simulation study.

## References

- [1] T.W. Anderson, *The Statistical Analysis of Time Series* (Wiley, New York, 1971).
- [2] O.D. Anderson and J.G. De Gooijer, Formulae for the covariance structure of the sampled autocovariances from series generated by general autoregressive integrated moving average process of order  $(p, d, q)$ ,  $d = 0$  or  $1$ , *Sankhyā Ser. B* **45** (1983) 249–256.
- [3] M.S. Bartlett, On the theoretical specification and sampling properties of autocorrelated time-series, *J. Roy. Statist. Soc. Suppl.* **8** (1946) 27–41, 85–97. Corrigenda **10** (1948) 200.
- [4] G.E.P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control* (Holden-Day, San Francisco, CA, 1976) revised edition.
- [5] D.R. Brillinger, Confidence intervals for the cross-covariance function, *Selecta Statist. Canad.* **5** (1978) 1–16.
- [6] C. Chatfield and D.L. Prothero, Box–Jenkins seasonal forecasting: problem in a case study, *J. Roy. Statist. Soc. Ser. A* **136** (1983) 295–336 (with discussion).
- [7] E.J. Hannan and C.C. Heyde, On limit theorems for quadratic functions of discrete time series, *Ann. Math. Statist.* **43** (1972) 2058–2066.

- [8] M.G. Kendall, A. Stuart and J.K. Ord, *The Advanced Theory of Statistics*, Volume 3 (Griffin, London, 1983) 4th edition.
- [9] M. Loève, *Probability Theory II* (Springer, New York, 1978) 4th edition.
- [10] A.I. McLeod, Derivation of the theoretical autocovariance function of autoregressive-moving average time series, *J. Roy. Statist. Soc. Ser. C Appl. Statist.* **24** (1975) 255–256.
- [11] G. Mélard, Illustration of the use of a general time series model, in: O.D. Anderson (Ed.), *Time Series Analysis: Theory and Practice 7*, Proceedings of the (general interest) international conference held at Toronto, Canada, August 1983 (North-Holland, Amsterdam, 1985) 53–75.
- [12] G. Mélard and R. Roy, Testing for homogeneity and stability of time series, *1983 Proceedings of the Business and Economic Statistics Section* (American Statistical Association, Washington, DC, 1983) 385–389.
- [13] G. Mélard and R. Roy, Sur un test d'égalité des autocovariances de deux séries chronologiques, *Canad. J. Statist.* **12** (1984) 333–342.
- [14] M.B. Priestley, *Spectral Analysis and Time Series*, Volume 1 (Academic Press, New York, 1981).
- [15] P.M. Robinson, Estimating variances and covariances of sample autocorrelations and autocovariances, *Austral. J. Statist.* **19** (1977) 236–240.
- [16] G. Tunnicliffe Wilson, Some efficient computational procedures for high order ARMA models, *J. Statist. Comput. Simul.* **8** (1979) 301–309.