NOTE ON THE AUTOREGRESSIVE SPECTRAL ESTIMATOR

Raúl P. Mentz and Carlos I. Martínez

Instituto de Investigaciones Estadísticas
Facultad de Ciencias Económicas
Universidad Nacional de Tucumán
and
CONICET

Summary

An estimator of the spectral density of a stationary time series is obtained by fitting to the observations an autoregressive model (often including the estimation of its order), and computing with sample values the spectrum of the indicated model. In the present note we consider the calculation of simultaneous confidence bands, according to Newton and Pagano (1984). The procedure is illustrated by means of Monte Carlo simulations, for series generated by autoregressive models or orders up to 5.

Key words and phrases. Confidence bands, asymptotic properties, Monte Carlo, Estimation of autoregressive order.

Correspondence:
R. P. Mentz
Casilla de Correo 209
(4000) Tucumán, Argentina
rmentz@herrera.unt.edu.ar
cimartine@herrera.unt.edu.ar

NOTE ON THE AUTOREGRESIVE SPECTRAL ESTIMATOR

1. Introduction

Given a time series (or stochastic process) y_t , t=0, ±1, ±2, etc. with the property of being stationary, we define its *autocovariance function* as the set of expected values

$$(1) \qquad \sigma_{st} = E(y_t - \mu)(y_s - \mu), \; s, t = 0,\pm1,\pm2, \; etc.$$

where μ is the expected value of the process.

These expectations are assumed to exist, in the sense that the defining series or integrals converge, and be positive semidefinite, in the sense that a matrix of arbitrary size formed with them should possess this property.

We also define its *autocorrelation function*

$$(2) \qquad \rho_s = \frac{\sigma_s}{\sigma_0}, \qquad s = 0,\pm1,\pm2, \; etc.$$

It is often convenient to consider transforming these parametric functions by defining the *spectral density function* of the process, namely

$$(3) \qquad f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \sigma_h \cos(\lambda h) = \frac{\sigma_0}{2\pi} \sum_{h=-\infty}^{\infty} \rho_h \cos(\lambda h), \; -\pi \le \lambda \le \pi.$$

This will exist provided the series converges. If the series of absolute values of the covariances converges, then (3) converges uniformly; see, for example, Anderson (1971).

An important example of time series is that corresponding to an *autoregressive model of order p*, that we identify by AR(*p*), with coefficients $\alpha_1, \alpha_2, ...,\alpha_p$ and variance σ^2 ; it satisfies the (stochastic) finite difference equation

$$(4) \qquad \sum_{j=0}^{p} \alpha_j (y_{t-j} - \mu) = \varepsilon_t, \qquad t = 0,\pm1,\pm2, \textit{etc},$$

where we define $\alpha_0=1$, and the ε_t constitute a stochastic process of uncorrelated (or independent) random variables with 0 expected value and constant variance σ^2. The α_j satisfy the condition that the associated polynomial equation, $\sum \alpha_j z^j = 0$ has all of its roots, real or complex, larger than 1 in absolute value. Under the stated condition y_t can be "inverted", i.e., expressed as an infinite linear combination of the ε_t's, current and past, and further, y_t is independent of ε_t for s>t, i.e., independent of future values of the ε_t 's.

Under these conditions, it can shown that the spectral density function of the AR(*p*) process is

$$(5) \qquad f_{AR}(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{\left| \sum_{j=0}^{p} \alpha_j e^{i\lambda j} \right|^2}, \qquad -\pi \le \lambda \le \pi,$$

so that if, for example, p=1, defining $\alpha=\alpha_1$ we find that the spectral density function of the AR(1) model is

$$(6) \qquad f_{AR(1)}(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{1 + \alpha^2 + 2\alpha \cos\lambda}, \qquad -\pi \le \lambda \le \pi.$$

The condition on the roots now implies that α is less than 1 in absolute value.

The *autoregressive estimator* of the spectral density function of a given (stationary) time series is defined as follows: (1) The series is approximated by an AR($\hat{p}$) model, whose

order $\hat{p}$ is suitably chosen; (2) The $\hat{p}$ + 1 parameters of this model are estimated; (3) With the estimates we form a function similar to (5), and we take this function (of frequency λ) as the sample estimator of the spectral density.

This procedure is related to that of Maximum Entropy. Priestley (1981, page 604) states: The method of maximum entropy is based on choosing a spectral estimate $\hat{h}(\omega)$ which is such that the entropy

$E = \int_{-\pi}^{\pi} \log\left\{\hat{h}(\omega)\right\} d\omega$ is maximized subject to the constraints (in our notation)

$\int_{-\pi}^{\pi} \hat{h}(\omega) e^{i\omega r} d\omega = c_r$, $r = 0, \pm 1, ..., \pm k$. Akaike showed that the resulting form of $\hat{h}(\omega)$ us exactly the same as for the AR spectral estimate, where the c_r satisfy the Yule-Walker equations.

Newton and Pagano (1984) considered the problem of computing simultaneous confidence bands for the spectral density function. The problem arises because the estimator consists of a set of values (in general it is computed at frequencies $\lambda_j = 2\pi j/T$, j=0,1,2,…,T/2, where T is sample size), each one having sampling variability.

In this note we analyze through simulated examples, the behavior of the proposed confidence band. We want to compare results with those in Newton and Pagano (1984) and with those in Mentz *et al* (2003). In the former no logarithmic transformation is introduced, so that we follow this option. Logarithms are used in the second reference.

Newton and Pagano (1984) approached the problem by considering that the reciprocal of the spectral density function of an AR(p) model, has the form of the spectral density function of a moving average model of the same order, with the same coefficients and with the variance of the error term equal to $4\pi^2/\sigma^2$; we denote this model by MA(p).

2. The AR(1) Model

With sample values $y_1,...,y_T$, parameters μ, α and σ^2 are estimated by one of the known methods of parametric estimation: Yule-Walker, least squares, what is known as *Burg's*

algorithm, or maximum likelihood under the assumption that the errors in the models have a normal distribution.. We can also estimate the autocovariances of the process, which means that the output of the estimation stage is, in the case of an AR(1) model, the following:

$$(7) \qquad \hat{\mu}, \hat{\alpha}, \hat{\sigma}^2, c_j = \frac{1}{T} \sum_{j=1}^{T-j} (y_t - \bar{y})(y_{t+j} - \bar{y}),\ j = 0,1,2.$$

With these sample values we define a set of autocovariances corresponding to an MA(*p*) model namely,

$$(8) \qquad \hat{\gamma}_0 = \frac{4\pi^2}{\hat{\sigma}^2}(1 + \hat{\alpha}^2), \qquad \hat{\gamma}_1 = \frac{4\pi^2}{\hat{\sigma}^2}\hat{\alpha},$$

and the empirical spectral density function of the MA(1) model is

$$(9) \qquad \hat{h}(\omega) = \frac{1}{\hat{f}(\omega)} = \frac{1}{2\pi}\{\hat{\gamma}_0 + 2\hat{\gamma}_1 \cos(\omega)\}.$$

We next define two matrices as follows:

$$(10) \qquad \Omega(\theta) = diag(\sigma^2 \Gamma^{-1}, 2\sigma^{-4})$$

where Γ is the matrix of variances and covariances of the process; in the special case of an AR(1) model this is σ_0, and is estimated by the sample value c_0. The other matrix is

$$\mathbf{B} = \begin{pmatrix} \frac{4\pi^2}{\sigma^2} 2\alpha & \sigma^2 \gamma_0 \\ \frac{4\pi^2}{\sigma^2} & \sigma^2 \gamma_1 \end{pmatrix}. \tag{11}$$

These two matrices lead to estimates of $\mathbf{C}$ and $\mathbf{B}$ by substituting sample values for α and σ^2 and further estimate γ using definitions (8). Defining

$$\hat{\mathbf{D}} = \hat{\mathbf{B}}\hat{\mathbf{C}}\hat{\mathbf{B}}' \tag{12}$$

the confidence band for the spectral density function is

$$\frac{1}{\hat{h}(\omega) + \hat{s}(\omega)} \leq f(\omega) \leq \frac{1}{\hat{h}(\omega) - \hat{s}(\omega)}, \tag{13}$$

where

$$\hat{s}^2(\omega) = \frac{1}{T} \chi^2_{\delta, p+1} \mathbf{x}'(\omega) \hat{\mathbf{D}}(\omega) \mathbf{x}(\omega), \tag{14}$$

$\chi^2_{\delta, p+1}$ is the percentile of the chi-square distribution with $p+1=2$ degrees of freedom, and we introduce the two-dimensional vector $\mathbf{x} = (1/\pi)(1/2 \;\; \cos(\omega))'$.

If it happens that $\hat{h}(\omega) - \hat{s}(\omega) \leq 0$, the upper bound of the confidence band is taken to be infinite.

Under these conditions, the (simultaneous) confidence bands have confidence coefficient approximately equal to δ. The statistical argument is based on projections proposed by Scheffé (1959, page 407), as cited by Newton and Pagano (1984).

3. The AR(5) Model

As a more general example we consider the AR(5) model, one that is studied by Newton and Pagano (1984). Observations y_t, t=1,2,...,T come from the model

$$y_t + \sum_{j=1}^{5} \alpha_j y_{t-j} = \varepsilon_t, \tag{15}$$

where the ε_t are independent with expected value 0 and constant variance $\sigma^2>0$. The output of the computer program includes the estimates $\hat{\alpha}_j, \hat{\sigma}^2$ and $c_j = (1/T)\sum_{t=1}^{T-j}(y_t - \bar{y})(y_{t+j} - \bar{y})$, where $\hat{\mu} = \bar{y} = \sum_{t=1}^{T} y_t / T$

The MA(5) used to approximate the autoregression has autocovariances estimated by

$$\hat{\gamma}_j = \frac{4\pi^2}{\hat{\sigma}^2} \sum_{s=0}^{5-j} \hat{\alpha}_s \hat{\alpha}_{s+j}, \qquad |j| = 0,1,...,5, \tag{16}$$

and by 0 for all other values of j. The spectral density of the MA(5)model is estimated by

$$\hat{h}(\omega) = \frac{1}{2\pi}\left\{\hat{\gamma}_0 + 2\sum_{j=1}^{5}\hat{\gamma}_j \cos(j\omega)\right\}, \qquad -\pi \le \omega \le \pi. \tag{17}$$

Defining matrices

(18) $$\hat{\boldsymbol{\Gamma}} = \left\| c_{|i-j|} \right\|,$$

(19) $$\hat{\mathbf{C}} = \mathrm{diag}(\hat{\sigma}^2 \hat{\boldsymbol{\Gamma}}^{-1}, 2\hat{\sigma}^{-4}),$$

and defining c as

(20) $$c = \frac{4\pi^2}{\hat{\sigma}^2}$$

we form the following matrix of order $(p+1)\mathrm{x}(p+1) = 6\mathrm{x}6$:

(21)

$$\hat{\mathbf{B}} = \frac{1}{c} \begin{pmatrix} 2\hat{\alpha}_1 & 2\hat{\alpha}_2 & 2\hat{\alpha}_3 & 2\hat{\alpha}_4 & 2\hat{\alpha}_5 & c\hat{\gamma}_0\hat{\sigma}^2 \\ 1+\hat{\alpha}_2 & \hat{\alpha}_1+\hat{\alpha}_3 & \hat{\alpha}_2+\hat{\alpha}_4 & \hat{\alpha}_3+\hat{\alpha}_5 & \hat{\alpha}_4 & c\hat{\gamma}_1\hat{\sigma}^2 \\ \hat{\alpha}_3 & 1+\hat{\alpha}_4 & \hat{\alpha}_1+\hat{\alpha}_5 & \hat{\alpha}_2 & \hat{\alpha}_3 & c\hat{\gamma}_2\hat{\sigma}^2 \\ \hat{\alpha}_4 & \hat{\alpha}_5 & 1 & \hat{\alpha}_1 & \hat{\alpha}_2 & c\hat{\gamma}_3\hat{\sigma}^2 \\ \hat{\alpha}_5 & 0 & 0 & 1 & \hat{\alpha}_1 & c\hat{\gamma}_4\hat{\sigma}^2 \\ 0 & 0 & 0 & 0 & 1 & c\hat{\gamma}_5\hat{\sigma}^2 \end{pmatrix}$$

This matrix contains those of smaller order. For example, in the case of the AR(4) model, we omit the last row and the fifth column, and set $\hat{\alpha}_5 = 0$, since $\alpha_5=0$.

With these matrices we form

(22) $$\hat{\mathbf{D}} = \hat{\mathbf{B}}\hat{\mathbf{C}}\hat{\mathbf{B}}'$$

and

(23) $$\hat{s}^2(\omega) = \frac{1}{T}\chi_{\delta,6}\,\mathbf{x}'\hat{\mathbf{D}}\mathbf{x}$$

where the vector $\mathbf{x}$ is defined by

(24) $$\mathbf{x} = \frac{\mathbf{1}}{\pi}\left(\frac{1}{2}\ \cos\omega\ \ \cos 2\omega\ \ \cos 3\omega\ \ \cos 4\omega\ \ \cos 5\omega\right)'.$$

Finally, the confidence bands for the spectral density can be written as

(25) $$\frac{1}{\hat{h}(\omega) + \hat{s}(\omega)} \le f(\omega) \le \frac{1}{\hat{h}(\omega) - \hat{s}(\omega)}.$$

This band has, approximately, level δ. Again, this is the key result in Newton and Pagano's main theorem.

As indicated, Newton and Pagano's approach is to set a band to $h(\omega) = 1/f(\omega)$. The band is $\hat{h}(\omega) - s(\omega) \le h(\omega) \le \hat{h}(\omega) + s(\omega)$, and has no problem even if the left hand side is negative for some frequency. When deriving the band for $f(\omega)$ they write: "It could happen that $\hat{h}(\omega) - s(\omega) < 0$ for some ω in which case infinity for the upper limit which does not diminish the probability contents of the band." One could solve this problem by letting this band to be $\max\left\{\hat{h}(\omega) - s(\omega), d\right\}$ where d is some constant, for example, d = 0.001. This procedure could alter the probability contents of the band or, to

say it in other words, the level α will be only approximately met. In our rather extensive simulations with the given models, we found no cases of negative estimates.

4. Numerical Examples

In this section we present and analyze some numerical examples. Observations were generated by Monte Carlo simulations, taking some values for the parameters and taking the error variance equal to 1. This choice coincides with that of Newton and Pagano, which means that the variance is treated as a scaling factor, the main interest being in estimating the spectral density. In all cases δ=0.05, which means that the confidence bands have (approximately) simultaneous 95% confidence level.

We use the S-PLUS package. The functions we use are: “arima.sim” to simulate the errors; “ar.burg” to fit AR models using Burg’s algorithm; “arima.mle” to fit AR models via maximum likelihood estimation under the Gaussian distribution.

4.1. Models

Observations were generated by four models, namely:

I. AR(1) with $\alpha_1 = -0.60$

II. AR(2) with $\alpha_1 = -0.40$ and $\alpha_2 = -0.45$.

III. AR(4) with $\alpha_1 = -2.760$, $\alpha_2 = 3.8106$, $\alpha_3 = -2.6536$ and $\alpha_4 = 0.9238$.

IV. AR(5) with $\alpha_1 = 1.7$, $\alpha_2 = 2.4$, $\alpha_3 = 1.634$, $\alpha_4 = 0.872$ and $\alpha_5 = 0.168$.

In the use of models II, III and IV we follow Beamish and Priestley (1981), also used by Newton and Pagano (1984) and elsewhere. The argument is that they are “easy, moderately easy and very difficult to fit, respectively”. These models have roots (1.11, 2.00), (1.12, 1.12, 1.19, 1.19, 3.33), and (1.02, 1.02, 1.02, 1.02), respectively, so that the roots of model IV are all close to 1 in absolute value. Model I with parameter 0.60 is expected to behave like model II.

4.2. Dealing with *p* and parameter estimation.

4.2.1. Dealing with *p*.

The autoregressive model has known order *p*, since this is the model generating the simulated observations. In the computer program used to estimate the parameters, this value of *p* can be forced, which implies that exactly as many α_j's as is the value of *p* are estimated. In our case, parameter estimation is done by maximum likelihood using the normal distribution; if observations have only approximately a normal distribution, asymptotic theory shows that use of this procedure provides a valid approximation; see, for example, Brockwell and Davis, 1987, Section 10.8..

With the observations generated by a model with known *p*, they can be used to estimate the order $\hat{p}$ of the AR model that provides the best fit to the data. In the S-PLUS program, this is done by *Burg's algorithm* for parameter estimation, and the AIC procedure to estimate the order. In Newton and Pagano, Parzen's CAT procedure is used together with Burg's algorithm. Our choice of AIC come from the program we use, namely, S-PLUS. One of our objectives is to explore the behavior of the computer programs in this package.

Table 1 shows the results of applying the indicated procedure to the 1000 Monte Carlo replications. Values concentrate for $\hat{p} = p$, but we also observe an important number of cases in which $\hat{p} > p$. In Table 1a) the estimated $\hat{p}$ is restricted to be ≤ 5 while in Table 1b) it is restricted to be ≤ 10.

Shibata (1976) derived the explicit asymptotic distribution of the estimated order $\hat{p}$. If this order is selected by means of the AIC criterion in the range from 0 to, say, P, the probability distribution of $\hat{p}$ has values from the true order to P. The distribution does not depend on the α_j, and these are estimated by maximum likelihood. In a simulation study for the AR(1) model with parameter α=0.80, he showed that the frequency of selecting the order $\hat{p} = 1$ is roughly 0.70, and our results in Table 1 agree with this.

Since the models proposed as examples have $p \leq 5$, the restriction that $\hat{p} \leq 10$ is reasonable: larger values of this estimate will introduce further estimates of α parameters and raise sample variability. Note that in Table 1, $p=1$ and $\alpha=0.60$, $\hat{\alpha}_1$ is close to 0.60 and $\hat{\alpha}_j$, j=2,3, 4 and 5 are small. Similar behavior is observed in the case of $p=2$, $\hat{p} \leq 5$. Further discussion of the effect of estimating the AR order is presented below.

For models I and II ($p=1$ and 2), changing from the restriction $\hat{p} \leq 5$ to $\hat{p} \leq 10$ alters the selection of order in about 5% of the cases: for model I selecting $\hat{p} = p$ reduces from 733 to 699 cases (out of 1000): in the remaining cases some value $\hat{p} > p$ is selected. For models III and IV (p=4 and 5) the effect is larger: for p=4, selecting $\hat{p} = p = 4$ reduces in about 15% of the cases in favor of choosing values of $\hat{p} > 4$; for $p=5$, selecting $\hat{p} = p = 5$ reduces in about a 30% of the cases in favor of choosing values of $\hat{p} \neq 5$ (for some series the selected order is 4).

4.2.2. Parameter estimation.

Table 2 shows the average results in the estimation of parameters. Sample sizes were $T=$ 100, 200 and 400, and 1000 replications (repetitions) were done in each case. Table 2 is done under the restriction $\hat{p} \leq 10$. Another table under the restriction $\hat{p} \leq 5$ is available from the authors.

For model AR(1) with parameters $\alpha_1 = -0.60$ and $\sigma^2 = 1$, the table contains the averages of the estimates, and of the corresponding standard deviation. These values must be compared with $\alpha_1 = -0.60$, $Var(\hat{\alpha}_1) \approx (1 - \alpha^2) / T$, with numerical values 0.08, 0.056 y 0.04, respectively.

Next we find the averages corresponding to estimates of the parameters obtained when the order value $\hat{p}$ is used, and estimation is done by using Burg's algorithm. The

estimation of α_1 has improved to some extent, while the averages for the other parameters are (comparatively) small. We recall that estimates are based upon the number of series indicated in tables 1 and 2, except that α_1 is always estimated with the 1000 series.

Similar considerations correspond to the other three models. In the case of model AR(2), use of the approximation

$$Var(\hat{\alpha}_j) \approx (1-\alpha_j^2)/T.$$

for the variance appears valid. This variance corresponds to the asymptotically normal distribution of the maximum likelihood estimator, see, for example, Brockwell and Davis (1987, page 252).

In general, changing the restriction on the chosen order has very little effect on the estimation of parameters for the chosen parameter values..

4.3. Confidence bands for the spectral density

4.3.1. Graphical Illustration.

In this section we show some examples of use of the techniques developed for the construction of simultaneous confidence bands. In all cases the logarithm of the spectral estimates is presented. The abscissas, which are usually given by *T*/2+ 1 frequency points from 0 to π , are presented in a transformed scale from 0 to 50, and labeled "Index" in the graphs. These graphs were also made by an S-PLUS function.

Figure 1. Observations were generated by simulation with the AR(1) model with parameter α = -0.60. The graph covers only the frequencies from 0 to π (0 to 50), since the remaining values, from -π to 0 can be omitted due to symmetry. The central dotted line is the theoretical spectral density of the AR(1) model used to generate the data, and hence corresponds to (6) with σ^2 = 1. The solid line is the spectral density coming from the MA(1) approximation, namely $\hat{h}^{-1}(\lambda)$ as defined in (9). Broken lines are the

simultaneous, 95%, confidence bands, defined by (13), one above and one below. They were computed with a sample of size T=100.

Figures 2 and 3 . Observations generated by simulation with the AR(2) model with parameters -0.40 and -0.45 This model was considered by Newton and Pagano. We observe the following: (1) There are 4 examples out of the 1000 replicates, (2) The theoretical spectral density and its MA(1) approximation in some case are parallel and in other have one or more crossings, (3) The theoretical spectral density in some cases falls entirely inside the confidence bands, in which case we say that there are 0 points outside the bands, and in other there 1 or more points outside the bands, (4) Since the bands are computed at frequencies $\lambda_j = 2\pi j / T$ for j = 0, 1, 2,…,T/2 if T is even, we compare the number of points falling out of the bands for $j \geq 0$ with T/2+1, which for T=100 equals 51. (5) The histogram in Figure 3 is an example of the distribution of points falling outside the 95% confidence bands, for model AR(2) with parameters –0.40 and –0.45, and with 1000 replications. We find a bar of size approximately equal to 300 at 0, and then a distribution reaching 24. The average is 2.25 (Table 6); its expected number is 0.05x51=2.55. The shape is typical of a mixture of distributions.

Figures 4 and 5 correspond, respectively, to models AR(4) and AR(5) with parameters defined in Section 4.1. They compare with Figure 1 in Newton and Pagano (op. cit.). In Figure 4, we find a good agreement between the two spectral densities, except at the peaks; the true spectral density is seen to lie (approximately) entirely within the confidence bands; the same occurs in Figure 5 for the AR(5) model: the approximation fails at the peak but the true AR spectral density falls entirely within the confidence bands. The bands are wide at the extremes of the range of frequency values, and narrow at the peaks.

4.3.2. Monte Carlo Simulation

Table 3 presents summaries of the simulations performed with the four models mentioned at the beginning of Section 4.1. Table 3 has results corresponding to restricting $\hat{p} \leq 5$ and also to restricting $\hat{p} \leq 10$. The first three columns of results are the averages

over the 1000 replications, of the number of points of the theoretical spectral density function, that occurred outside the 95% (simultaneous) confidence band. The expected numbers are 2.55 for T=100, 5.05 for T=200 and 10.05 for T=400. We find that for models I and II the observed (average) values are in good agreement with these figures. For model IV the averages are larger than expected, in the order of twice the theoretical figure. For model III the averages are even larger. One possible explanation for the left hand part of the tables, is that for these larger values of the order p (4 and 5), the limitation imposed that the order of the fitting AR is the same p used to generate the data, results being too restrictive.

The importance of order selection is considered in the next three columns. In them the averages are computed when the order is chosen according to the AIC criterion, but still maintaining the restriction of its being less than or equal to 5 or to 10, respectively. We find a sizeable improvement for T=400, for models I and II. Here the improvement occurs because about ¼ of the series led to changes in the value of the order, which, as indicated in Table 1, consisted in increasing the value of p. Further, for orders p = 1 or 2, the restriction on the order is not important. When models III and IV are considered there are more points (on average) outside the confidence band than expected. This is particularly true of model III (AR(4)), confirming the finding of previous authors, Newton and Pagano (1984), Beamish and Priestley (1981).

4.3.2. Comparison with Previous Work

In a paper by the authors (Mentz, Viollaz and Martínez, 2003, unpublished) , standard, windowed spectral estimators were studied by Monte Carlo. They were: (1) Confidence bands available in the S-PLUS package, in which the χ^2 approximation is used, with 8.29 degrees of freedom; (2) Similar bands with 9.14 degrees of freedom; (3) Confidence bands using the normal approximation, and (4) Similar bands with a correction for asymptotic bias. A summary of results is Table 4.

It follows that (except for Model III with T=100), the bands based on windowed spectral estimates, leave, on average, less points outside the bands than the AR spectral estimates.

One main point is that while the windowed estimates are based on (95%) point-by-point or marginal analysis, the bands for the AR spectral estimators have 95% joint confidence coefficient, and we expect the former to leave less points outside of them.

5. Summary and Conclusions

In this paper we analyze the behavior of the AR spectral estimator in simulated time series. There exist a set of models frequently studied in the literature, namely, models AR(*p*) with p=2, 4 and 5. There were considered by the present authors in a previous paper dealing with standard windowed spectral estimates, that is those coming from smoothing the periodogram (Mentz, Viollaz and Martínez, 2003). They are confidence intervals computed separately at each frequency with a common confidence coefficient (e.g. 95%). Hence, the level of confidence for the whole set of frequencies is considerably low. In the indicated reference point-by-point inference is compared with joint inference based on extreme value distributions (Woodroofe and Van Ness, 1967).

The bands for the AR spectral estimator developed by Newton and Pagano (op. cit.) are based on projections proposed by Scheffé (1959), and hence have a joint confidence level (e.g. 95%). This is an asymptotic approximation.

The behavior of the AR bands is studied in some detail: (1) The role of selecting the order of the approximating AR model is studied, and found that for the models considered, restricting it to be less than or equal to 10 improves over restricting it to be less than or equal to 5, in particular, for the more complicated models AR(4) and AR(5). However, it affects little the estimation of parameters. (2) Graphical examples are provided to illustrate the results. (3) Tabled results show average numbers of sample points falling outside the bands.

In summary, the AR spectral simultaneous confidence bands are seen to behave reasonably well for the models considered.

As pointed out by a referee, bootstrap estimation of the standard error of the confidence bands could be attempted, but we do not include this work in the present paper.

There has been a good deal of concern about the merits of the AR spectral estimator. See, for example, Priestley (1981, p. 612) who wrote: “It is seen that the AR estimate is highly irregular, with quite marked oscillations over the whole frequency range”. Harvey (1994, p. 205) commented on the importance of selecting an adequate (estimated) AR order. See also Brockwell and Davis (1987, Section 10.6), Koopmans (1974, Section 9.4).

Figure 1

Spectral density function of AR(1) model with parameter -0.60, MA(1) approximation, and 95% simultaneous confidence bands

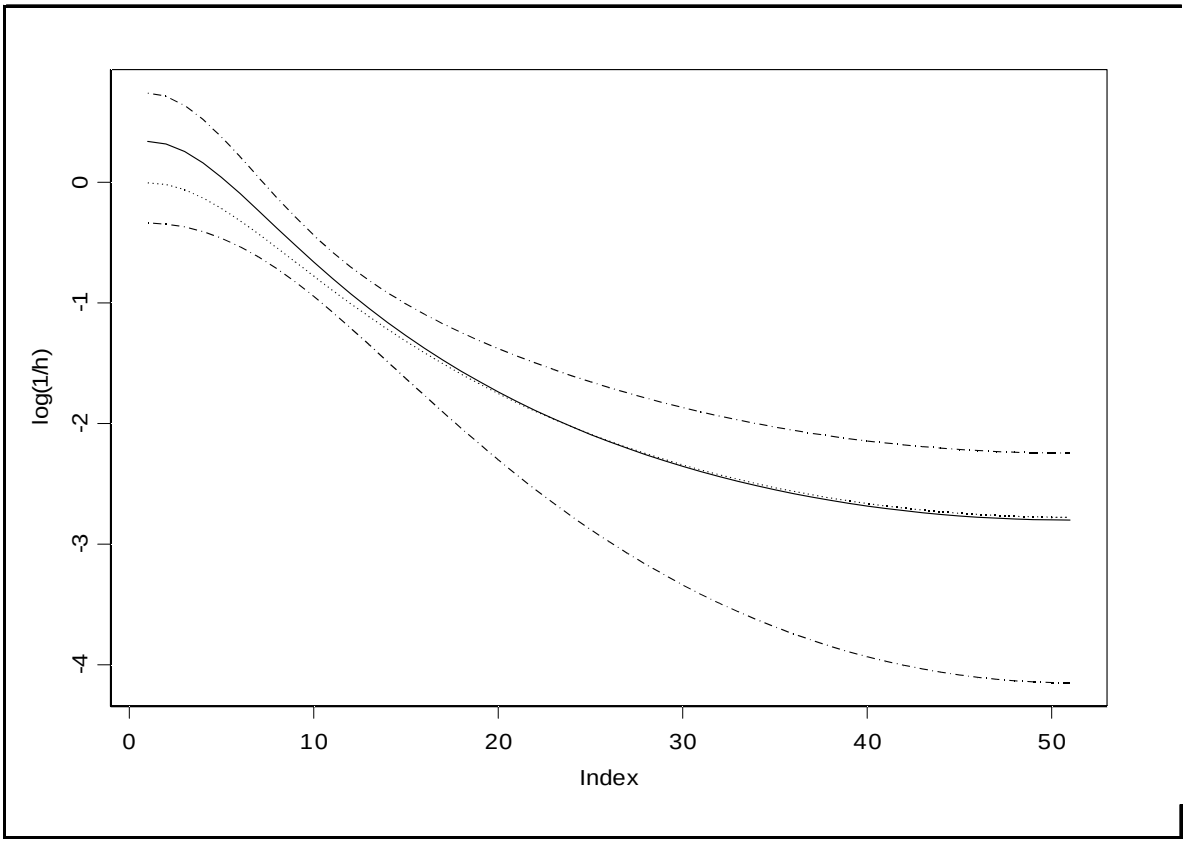

Figure 2

Spectral density function of AR(2) model with parameter -0.40 and -0.45, MA(2) approximation, and 95% simultaneous confidence bands

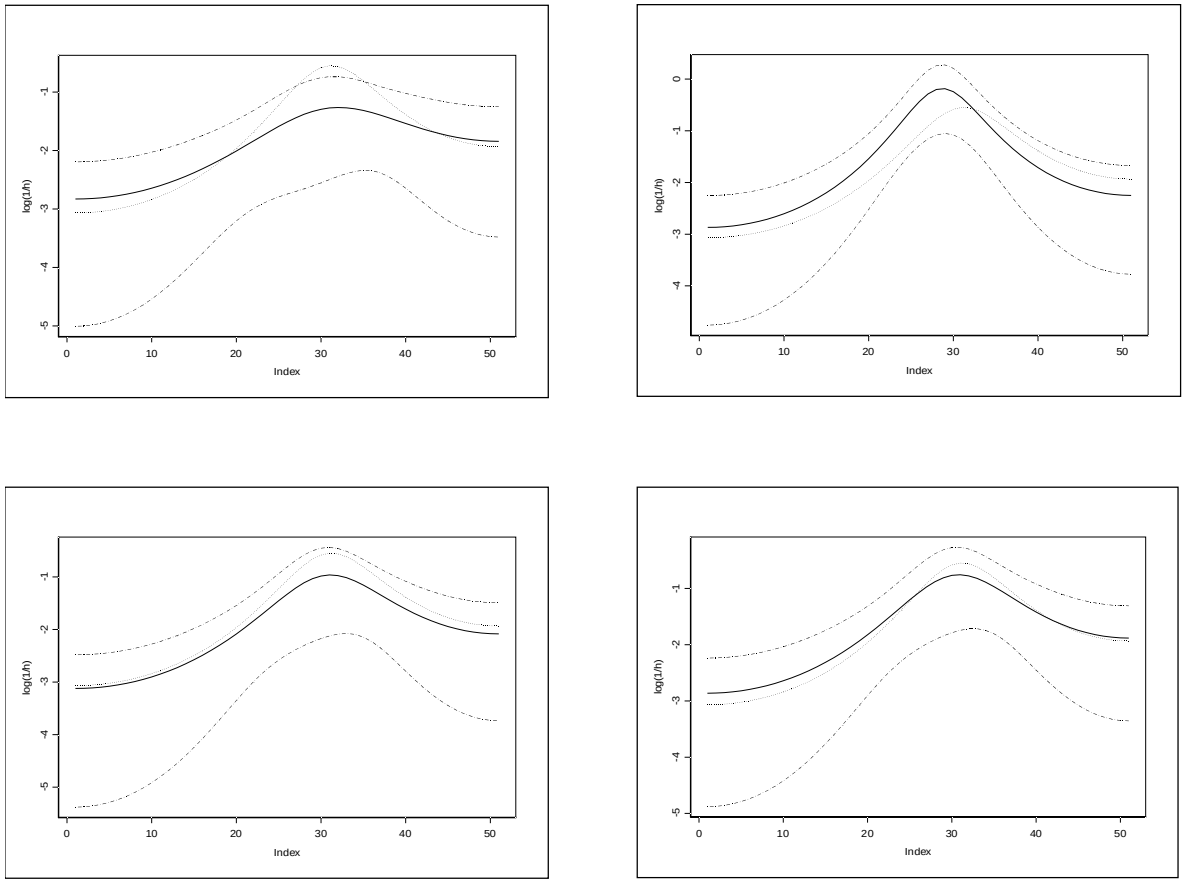

Figure 3

Histogram of points folling outside the 95% simultaneous confidence bands, AR(2) model, AR spectral estimations

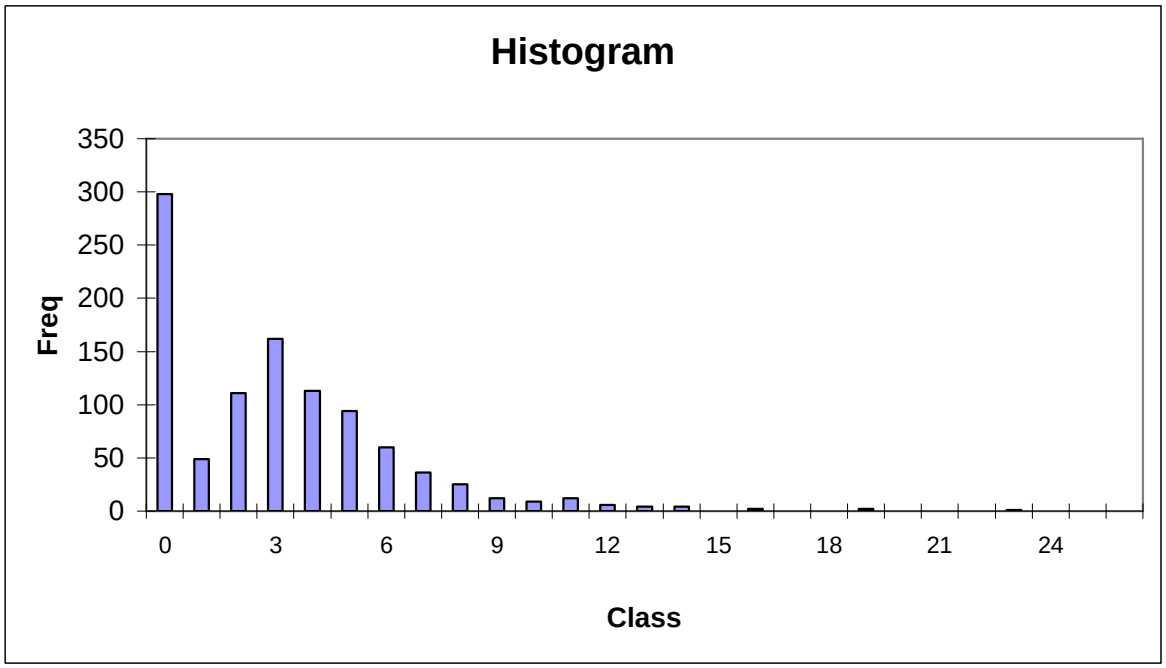

Figure 4

Spectral density function of AR(4) model with parameter (see text), MA(4) approximation, and 95% simultaneous confidence bands

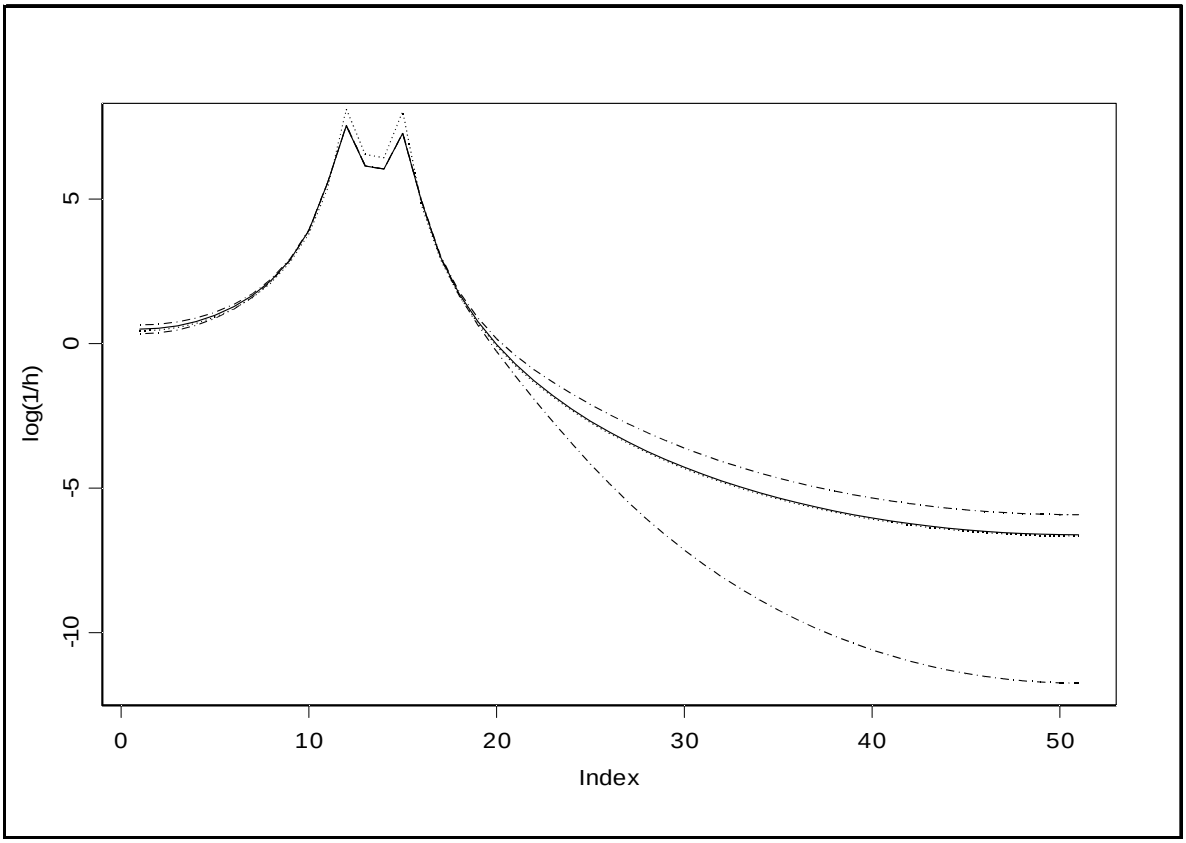

Figure 5

Spectral density function of AR(5) model with parameter (see text), MA(5) approximation, and 95% simultaneous confidence bands

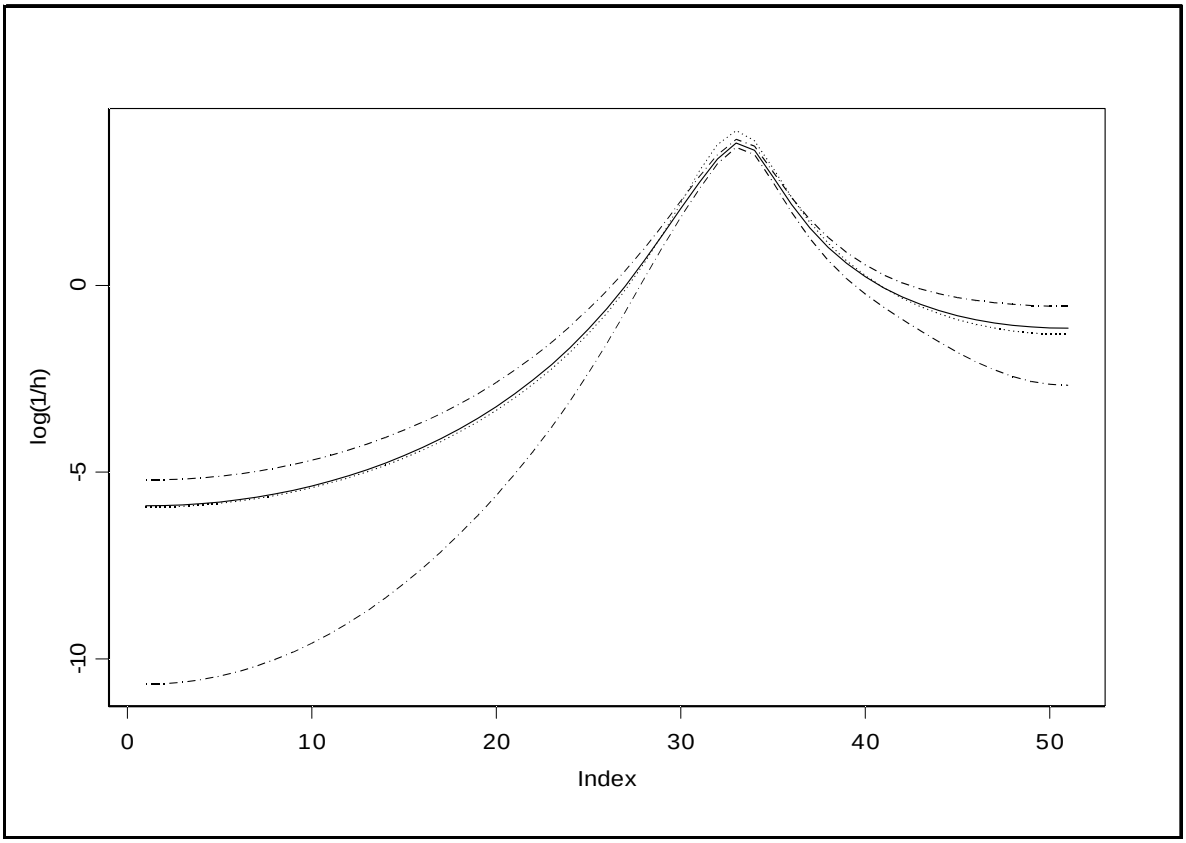

Table 1

a) Frequency distributions of autoregressive orders chosen by S-PLUS using Burg's estimation algorithm, in 1000 replications (p ≤ 5).

Model	True order p	Sample size T	Chosen Order				
			1	2	3	4	5
I	1	100	733	113	68	46	40
		200	744	121	61	44	30
		400	722	146	55	38	39
II	2	100	1	745	134	69	51
		200	0	756	133	53	58
		400	0	766	109	58	57
III	4	100	0	0	0	819	181
		200	0	0	0	836	164
		400	0	0	0	850	150
IV	5	100	0	0	0	428	572
		200	0	0	0	207	793
		400	0	0	0	38	962

b) Frequency distributions of autoregressive orders chosen by S-PLUS using Burg's estimation algorithm, in 1000 replications (p ≤ 10).

Model	True order p	Sample size T	Chosen Order									
			1	2	3	4	5	6	7	8	9	10
I	1	100	699	108	63	37	30	18	11	11	11	12
		200	702	115	57	36	21	21	20	12	7	9
		400	691	138	49	34	30	19	14	10	8	7
II	2	100	1	694	118	57	29	32	25	13	17	14
		200	0	719	118	49	38	26	18	15	8	9
		400	0	730	101	53	42	23	14	18	11	8
III	4	100	0	0	0	665	135	58	46	44	24	28
		200	0	0	0	712	123	63	34	36	16	16
		400	0	0	0	742	107	54	33	25	18	21
IV	5	100	0	0	0	365	399	105	52	38	26	25
		200	0	0	0	171	569	113	49	41	32	25
		400	0	0	0	33	744	98	46	38	20	21

Table 2

Average estimates in 1000 replications

a) Model I, p = 1, $\alpha_1 = 0.60$, MLE

T	α_1	$Var(\alpha_1)^{1/2}$	σ^2
100	0.5895	0.0805	0.9818
200	0.5944	0.0568	0.9932
400	0.5971	0.0401	0.9958

b) Model I, p = 1, $\alpha_1 = 0.60$, p_{est} Burg's Algorithm

T	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	σ^2
100	0.5907	-0.0084	0.0091	-0.0494	-0.0178	0.0296	-0.0117	-0.0545	0.0299	-0.0136	0.9589
200	0.5964	-0.0048	-0.0082	-0.0224	0.0220	-0.0118	-0.0062	-0.0460	0.0422	0.0286	0.9818
400	0.5977	-0.0002	-0.0088	-0.0013	0.0168	-0.0292	-0.0144	0.0038	0.0498	0.0792	0.9899

c) Model II, p = 2, $\alpha_1 = 0.40$, $\alpha_2 = 0.45$, MLE

T	α_1	α_2	$Var(\alpha_1)^{1/2}$	σ^2
100	0.3966	0.4255	0.0909	0.9707
200	0.3978	0.4369	0.0637	0.9884
400	0.3986	0.4446	0.0448	0.9934

d) Model II, p = 2, $\alpha_1 = 0.40$, $\alpha_2 = 0.45$, p_{est} Burg's Algorithm

T	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	σ^2
100	0.3986	0.4287	0.0050	-0.0364	-0.0052	0.0105	-0.0152	-0.0260	0.0695	-0.0465	0.9499
200	0.3991	0.4395	-0.0015	-0.0254	0.0058	-0.0095	-0.0334	-0.0100	0.0445	0.0623	0.9776
400	0.3985	0.4445	-0.0010	-0.0004	0.0071	-0.0042	0.0030	-0.0102	0.0256	-0.0278	0.9881

Table 2 (continued)

e) Model III, p = 4, $\alpha_1 = 2.760$, $\alpha_2 = -3.8106$, $\alpha_3 = 2.6535$, $\alpha_4 = -0.9238$, MLE

T	α_1	α_2	α_3	α_4	$Var(\alpha_1)^{1/2}$	σ^2
100	2.7370	-3.7478	2.5892	-0.8944	0.0442	0.9493
200	2.7481	-3.7774	2.6190	-0.9077	0.0295	0.9786
400	2.7539	-3.7935	2.6356	-0.9155	0.0201	0.9880

f) Model III, p = 4, $\alpha_1 = 2.760$, $\alpha_2 = -3.8106$, $\alpha_3 = 2.6535$, $\alpha_4 = -0.9238$, p_{est} Burg's Algorithm

T	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	σ^2
100	2.7397	-3.7579	2.6074	-0.9192	0.0802	-0.1246	0.1192	-0.0932	0.0549	-0.0338	0.9335
200	2.7484	-3.7806	2.6278	-0.9231	0.0617	-0.0990	0.1112	-0.0949	0.0753	-0.0455	0.9711
400	2.7540	-3.7962	2.6420	-0.9237	0.0272	-0.0517	0.0963	-0.1323	0.1048	-0.0483	0.9839

g) Model IV, p = 5, $\alpha_1 = -1.7$, $\alpha_2 = -2.4$, $\alpha_3 = -1.634$, $\alpha_4 = -0.872$, $\alpha_5 = -0.168$, MLE

T	α_1	α_2	α_3	α_4	α_5	$Var(\alpha_1)^{1/2}$	σ^2
100	-1.6802	-2.3508	-1.5788	-0.8355	-0.1558	0.1008	0.9432
200	-1.6906	-2.3763	-1.6061	-0.8537	-0.1606	0.0705	0.9724
400	-1.6957	-2.3879	-1.6208	-0.8626	-0.1650	0.0496	0.9848

h) Model IV, p = 5, $\alpha_1 = -1.7$, $\alpha_2 = -2.4$, $\alpha_3 = -1.634$, $\alpha_4 = -0.872$, $\alpha_5 = -0.168$, p_{est} Burg's Algorithm

T	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	σ^2
100	-1.6680	-2.3278	-1.5388	-0.8103	-0.2139	-0.0014	0.0319	0.0255	0.0272	-0.0161	0.9308
200	-1.6844	-2.3641	-1.5839	-0.8383	-0.1833	-0.0208	-0.0400	-0.0562	-0.0392	-0.0296	0.9666
400	-1.6948	-2.3860	-1.6174	-0.8604	-0.1696	-0.0076	-0.0146	-0.0230	-0.0203	-0.0103	0.9818

Table 3

Averages of sample points falling outside the 95% simultaneous confidence bands.

(In parenthesis, number of series with $\hat{p} = p$)

Model	True Order p	Bands Computed Using p			Bands Computed Using $\hat{p} \leq 5$			Bands Computed Using $\hat{p} \leq 10$		
		T = 100	T = 200	T = 400	100	200	400	100	200	400
I	1	2.25	5.46	18.13	3.05 (267)	6.54 (256)	10.03 (278)	7.52 (301)	7.15 (298)	9.82 (309)
II	2	3.09	6.17	14.49	5.49 (255)	6.13 (244)	9.42 (224)	7.37 (306)	7.04 (281)	9.37 (270)
III	4	14.73	30.57	65.21	17.87 (181)	36.16 (164)	74.57 (150)	18.16 (335)	32.32 (288)	64.42 (258)
IV	5	8.99	17.73	38.01	12.47 (428)	27.95 (207)	64.82 (38)	11.87 (597)	20.29 (427)	32.09 (249)

Table 4

Averages of sample points falling outside the 95% simultaneous confidence bands.

Model	True AR Order	Sample Size T	This Study		Previous Study			
			p	$\hat{p}$	(1)	(2)	(3)	(4)
II	2	100	3.09	7.37	0.716	1.171	1.962	0.84
		200	6.17	7.04	0.471	0.971	2.262	0.739
		400	14.49	9.37	0.306	0.773	2.797	0.763
III	4	100	14.73	18.16	14.401	13.954	11.669	11.081
		200	30.57	32.32	4.052	3.953	3.579	2.293
		400	65.21	20.29	0.693	0.945	2.512	0.657
IV	5	100	8.99	11.87	3.141	3.014	2.66	1.987
		200	17.73	20.29	1.451	1.679	2.423	1.2
		400	38.01	32.09	0.631	0.934	2.62	0.753

References

Anderson, T. W. (1971) *The Statistical Analysis of Time Series*, New York, John Wiley.

Beamish, N. and M. B. Priestley (1981) A study of autoregressive and window spectral estimation. *Appl. Statist.* **30**, 41-58.

Brockwell, P. J. and R. A. Davis (1987) *Time Series: Theory and Methods*. New York, Springer-Verlag.

Harvey, A. C. (1994) *Time Series Models, Second Edition*. Cambridge, Massachusetts, The MIT Press.

Koopmans, L. H. (1974) *The Spectral Analysis of Time Series*. New York, Academic Press.

Mentz, R. P., A. J. Viollaz and C. I. Martinez (2003) Confidence bands for time series problems in the time and frequency domains, *Brazilian Journal of Probability and Statistics*, **17**, 179-211.

Newton, H. J. and M. Pagano (1984) Simultaneous confidence bands for autoregressive spectra. *Biometrika*. **71,** 1, 197-202.

Priestley, M. B. (1981) *Spectral Analysis and Time Series, Vol. 1: Univariate Series*. London, Academic Press.

Scheffé, H. (1959) *The Analysis of Variance*. New York, John Wiley.

Shibata, R. (1976) Selection of the order of an autoregressive model by Akaike's information criterion, Biometrika, **63**, No. 1, 117-126.

Woodroofe, M. and J. Van Ness (1967) The maximum deviation of spectral densities, *Ann. Math. Stat.*,**36**, 1558-1569.