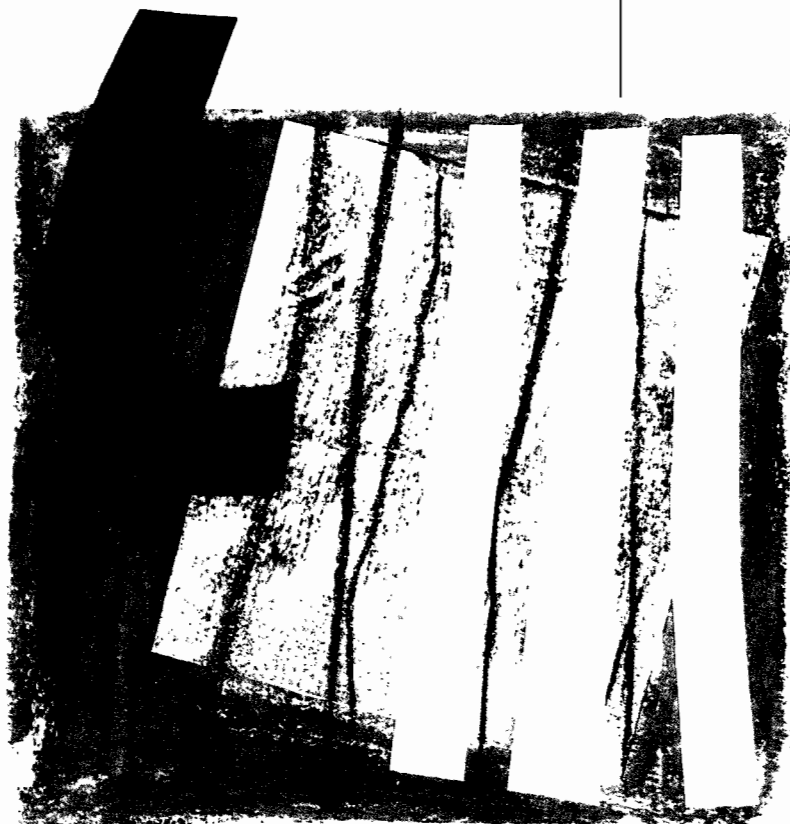


**TRIMMING FREQUENCIES IN  
LOG-PERIODOGRAM REGRESSION  
OF LONG-MEMORY TIME SERIES**

Cristina Martínez  
and Santiago Velilla

96-14



WORKING PAPERS

Working Paper 96-14  
Statistics and Econometrics Series 04  
March 1996

Departamento de Estadística y Econometría  
Universidad Carlos III de Madrid  
Calle Madrid, 126  
28903 Getafe (Spain)  
Fax (341) 624-9849

TRIMMING FREQUENCIES IN LOG-PERIODOGRAM REGRESSION  
OF LONG-MEMORY TIME SERIES

Cristina Martínez and Santiago Velilla\*

Abstract

---

Previous work on log-periodogram regression in time series with long range dependence is reviewed. The effect of both low and large frequencies on the estimate of the fractional difference parameter is analyzed. Some new simulation results are presented.

---

Key Words

Autocorrelation function; Fractional ARIMA models; Spectral density.

\* Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. Research partially supported by grant PB-930232 (Spain).

## 1. INTRODUCTION

There is strong evidence that long - memory time series occur quite frequently in practice. A characteristic indication of long range dependence is the appearance of unbounded spectral densities in a neighborhood of the origin  $\lambda = 0$ . The class of fractional ARIMA(p,d,q) models allow for this situation. This paper treats some aspects of estimation of the parameter  $d$ . Some background is given in section 2, where notation is also established. Log - periodogram regression, a fundamental technique of estimation, is introduced in section 3. Advantages and disadvantages of log - periodogram regression are discussed next and alternative techniques are discussed. Some new simulation results are presented in section 4 and section 5 contains final comments.

## 2. BACKGROUND

Let  $\{X_t: t \in \mathbb{Z}\}$  be a second order stationary process with finite variance. An important summary of the stochastic features of  $\{X_t: t \in \mathbb{Z}\}$  is given by the autocorrelation function

$$\rho_k = \text{cov}[X_{t+k}, X_t] / \text{var}[X_t] = \gamma_k / \gamma_0, \quad (2.1)$$

evaluated at integer lags  $k = 0, \pm 1, \pm 2, \dots$ . Given a finite sample  $X_1, \dots, X_n$ , interest lies in finding a parsimonious parametric description of  $\{X_t: t \in \mathbb{Z}\}$ , particularly for reliable prediction of future values of the process. Since, properly speaking, the coordinates of the sequence  $\{\rho_k: k \in \mathbb{Z}\}$  are unknown parameters, practical modelling of  $\{X_t: t \in \mathbb{Z}\}$  should be based on the sample autocorrelation function

$$r_k = \hat{\gamma}_k / \hat{\gamma}_0, \quad k = 0, \pm 1, \pm 2, \dots, \pm m, \quad (2.2)$$

where  $\hat{\gamma}_k = n^{-1} \sum_{t=1}^{n-k} X_t X_{t+k}$ ,  $\hat{\gamma}_0 = n^{-1} \sum_{t=1}^n X_t^2$ , and  $m \ll n$  is an upper bound for the

lag in order to produce efficient estimates of  $\hat{\gamma}_k$ .

It is often the case that the information provided by the  $r_k$  suggests that  $\rho_k$  decays very fast. Broadly speaking, the process has "short - memory", a term that describes that, according to the autocorrelation function,  $X_{t+k}$  and  $X_t$  show little dependence for  $k$  large enough. In this situation, the causal and invertible ARMA(p,q) model for the process  $\{X_t: t \in \mathbb{Z}\}$ ,

$$\phi(B)X_t = \theta(B)\varepsilon_t, \quad (2.3)$$

where  $\phi(B)$  and  $\theta(B)$  are polynomial of degrees, respectively,  $p$  and  $q$  with roots outside the unit circle,  $B$  is the backward shift operator  $BX_t = X_{t-1}$ , and  $\{\varepsilon_t\}$  is a zero mean white noise with finite variance  $\sigma^2$ , provides a well established methodology for both inference and prediction. (see, e.g. Brockwell and Davis (1991)). Under (2.3), it can be seen that  $\rho_k$  is bounded exponentially

$$|\rho_k| \leq A \exp(-Bk), \quad k = 0, 1, 2, \dots, \quad (2.4)$$

for some positive constants  $A$  and  $B$ , a property that makes explicit the fast decay of the autocorrelation when the lag increases. Observe that (2.4) implies that

$$\sum_{k=-\infty}^{\infty} |\rho_k| < \infty, \quad (2.5)$$

so that the spectral density of the process

$$f(\lambda) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \gamma_k \exp(-ik\lambda), \quad -\pi \leq \lambda \leq \pi, \quad (2.6)$$

is bounded in a neighborhood of  $\lambda = 0$ . In fact, for an ARMA(p,q) process, since  $f(\lambda) = (\sigma^2/2\pi) |\phi(z)|^2 / |\theta(z)|^2$ , where  $z = \exp(-i\lambda)$ , the spectral density is continuous and bounded in  $[-\pi, \pi]$  by construction.

The flexibility of the ARMA(p,q) model (2.3) is not enough, however, to cover all practical cases. As observed by Granger (1966), the typical "shape" of the spectral density of an economic variable offers an unbounded

behaviour near the zero frequency. As described in Beran (1994, chaps. 1 and 2) there are many data situations in which the process, although stationary, has a sample autocorrelation that suggests a slow hyperbolic decay of the  $\rho_k$  in the form

$$|\rho_k| \sim ck^{-\alpha}, \quad (2.7)$$

as  $k \rightarrow \infty$  for some  $\alpha \in (0,1)$ . Equivalently, near the origin the spectral density is

$$f(\lambda) \sim m|\lambda|^{-\beta}, \quad (2.8)$$

for some  $\beta \in (0,1)$ . The convergence condition (2.5) is violated and the process is said to possess "long range" dependence or a "long - memory" pattern.

In the light of the comments of the previous paragraph, the need for new parametric models for time series with long - memory is apparent. The most common alternative to the classic ARMA(p,q) formulation is given by the family of fractionally differenced ARIMA (p,d,q) models. An stationary process  $\{X_t: t \in \mathbb{Z}\}$  follows an ARIMA (p,d,q) model when it can be written in the form

$$\phi(B)(1 - B)^d X_t = \theta(B)\varepsilon_t, \quad (2.9)$$

where  $-0.5 < d < 0.5$  and the rest of the elements are as in (2.3). (2.9) is a more flexible class of models than the traditional ARIMA (p,h,q) family for integer h. This avoids overdifferencing the series that leads to a zero spectral density at the null frequency, therefore missing perhaps some relevant qualitative features of the data. Near the origin, the spectral density of (2.9) is

$$f(\lambda) \sim m|\lambda|^{-2d}, \quad (2.10)$$

where  $m = (\sigma^2/2\pi)[\phi(1)]^2/[\theta(1)]^2$ , so there is long - range dependence for  $0 < d < 0.5$ . In this case, the asymptotic behaviour of  $\rho_k$  is

$$\rho_k \sim [\Gamma(1-d)/\Gamma(d)] |k|^{2d - 1}. \quad (2.11)$$

An equivalent form of writing (2.9), is

$$(1 - B)^d X_t = u_t, \quad (2.12)$$

where  $u_t = [\theta(B)/\phi(B)]\varepsilon_t$  is a linear process with bounded and continuous spectral density depending on the short memory part of (2.9). The problem of estimating the fractional difference parameter  $d$  is considered next.

### 3. LOG - PERIODOGRAM REGRESSION

Geweke and Porter - Hudak (1983) develop a method of estimation of  $d$  based on a least squares regression technique in the framework (2.12). Let  $f(\lambda)$  and  $g(\lambda)$  be, respectively, the spectral densities of the processes  $\{X_t; t \in \mathbb{Z}\}$  and  $\{u_t; t \in \mathbb{Z}\}$ . Observe that  $g(\lambda) = (\sigma^2/2\pi)|\phi(z)|^2/|\theta(z)|^2$ ,  $z = \exp(-i\lambda)$ . Since

$$f(\lambda) = |1 - \exp(-i\lambda)|^{-2d} g(\lambda), \quad (3.1)$$

taking logarithms in (3.1) gives

$$\log[f(\lambda)] = \log[g(0)] - d \log[|1 - \exp(-i\lambda)|^2] + \log[g(\lambda)/g(0)]. \quad (3.2)$$

Introduce now the periodogram of the series

$$I_n(\omega_j) = (2\pi n)^{-1} \left| \sum_{t=1}^n X_t \exp(-it\omega_j) \right|^2, \quad (3.3)$$

where  $\omega_j = 2\pi j/n$  is the  $j$ th Fourier frequency,  $1 \leq j < (n/2)$ . Adding  $\log[I_n(\omega_j)]$  to both sides of (3.2) leads to

$$\begin{aligned} \log[I_n(\omega_j)] &= \log[g(0)] - d \log[|1 - \exp(-i\omega_j)|^2] \\ &\quad + \log[I_n(\omega_j)/f(\omega_j)] + \log[g(\omega_j)/g(0)]. \end{aligned} \quad (3.4)$$

If  $\omega_j$  is small enough, by the continuity assumptions on the density  $g(\lambda)$ , the last summand in (3.4) is negligible in comparison to the others. Changing the notation, (3.4) suggests estimating  $d$  by least squares in the simple linear regression model

$$y_j = a - d x_j + \varepsilon_j, \quad (3.5)$$

where  $a = \log[g(0)]$ ,  $x_j = \log[|1 - \exp(-i\omega_j)|^2] = \log[4 \sin^2(\omega_j/2)]$ , and  $\varepsilon_j = \log[I_n(\omega_j)/f(\omega_j)]$ , for  $j = 1, \dots, m$ , where  $m$  is a threshold to be determined. Geweke and Porter - Hudak (1983) argue heuristically that there exists a sequence  $m = m_n$  such that the least squares estimator

$$\hat{d}_n = - \frac{\sum_{j=1}^m (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^m (x_j - \bar{x})^2}, \quad (3.6)$$

is, asymptotically,

$$N(d, \pi^2 / [6 \sum_{j=1}^m (x_j - \bar{x})^2]). \quad (3.7)$$

The method above has several attractive features. For example, it is computationally very simple and allows estimation of  $d$  without knowledge of the orders  $p$  and  $q$ . However, as some authors point out, the proposal of Geweke and Porter - Hudak (1983) presents several problems that can be summarized as follows:

*i)* For a long - memory process as (2.12), the usual asymptotic properties of the periodogram do not hold and, therefore, the i.i.d. assumption for the errors  $\{\varepsilon_j\}$  is untenable. This has been obtained in Künsch (1986), Hurvich and Beltrao (1993), and Robinson (1995). Although the effect of this deviation can be ignored asymptotically, it can affect the performance of the estimator  $\hat{d}_n$  in small samples. A possible solution is to carry out the least squares regression on small Fourier frequencies only;

*ii)* Beran (1994, chap. 4) and Robinson (1995) mention that the derivation of the asymptotic distribution (3.7) by Geweke and Porter - Hudak (1983), is not totally correct. Robinson (1995) suggests trimming frequencies not only from above but also from below. Change the notation slightly and introduce the parameter  $H$ , where

$$d = H - .5. \quad (3.8)$$

As stated in Beran (1994, p. 98), a key result is:

**Theorem 3.1** Let the process  $\{X_t; t \in \mathbb{Z}\}$  of (2.12) be Gaussian with long - memory. Let  $\hat{H}_n$  ( $\hat{d}_n + .5$ ) be the least squares estimate of  $H$  based on Fourier frequencies  $\omega_j$ , where  $l_n \leq j \leq m_n$ , and  $\{l_n\}$ ,  $\{m_n\}$  are integer sequences that tend to infinity and satisfy some growth restrictions, among others,  $l_n/m_n \rightarrow 0$ . Under some regularity conditions on the spectral density of  $\{X_t; t \in \mathbb{Z}\}$ ,

$$m_n^{1/2}(\hat{H}_n - H) \xrightarrow{D} N(0, \pi^2/24). \quad (3.9)$$

An important practical problem that remains to be solved is how to choose  $l_n$  and  $m_n$  in finite samples. According to Beran (1994, p. 99) increasing  $m_n$  reduces the variance of  $\hat{H}_n$  but increases the bias. On the other hand, reducing  $m_n$  increases the variance but reduces the bias. A simulation is performed by Agiakloglou, Newbold and Wohar (1993) to study the bias phenomenon in small samples. Hurvich and Beltrao (1994) discuss a data driven choice of  $l_n$ . Recently, Hassler (1995) studies the problem of choice of  $l_n$  and  $m_n$  in finite samples. Further comments on the choice of  $l_n$  and  $m_n$  are given in the next section.

#### 4. SIMULATION RESULTS

It seems as if all this previous work would not be at all conclusive in providing helpful guidelines for determining suitable values of the trimming constants  $l_n$  and  $m_n$  given a series of length  $n$ . This section reports some preliminary simulation results that illustrate the nature of the problem.

In table 1,  $N = 1000$  samples of length  $n = 400$  are generated through an ARIMA(0,d,0) model for values of  $d$  in the range  $-.4$  to  $.4$  separated by  $.1$ . The Geweke - Porter - Hudak estimator (GPH) is computed for each sample, and the average values and mean squared errors are reported along with the bias.



<b>d</b>	<b>mean</b>	<b>bias</b>	<b>mse</b>
.4	.722	.322	.008
.3	.543	.243	.009
.2	.361	.161	.009
.1	.184	.084	.009
-.1	-.175	-.074	.008
-.2	-.351	-.151	.009
-.3	-.528	-.228	.010
-.4	-.699	-.299	.010

**Table 1**

As observed from table 1, the bias in the **GPH** increases with the absolute value of **d** in quite a symmetric fashion. The mean squared errors are all uniformly small suggesting a quite accurate estimation. Tables 2, 3, 4, and 5 are formed using the both-sides trimming proposal of Robison (1995). The theoretical asymptotic variances offered by theorem 3.1, namely  $\pi^2/(24m_n)$  are given in parenthesis next to the table number. The values of the trimming thresholds appear below the table number.

<b>d</b>	<b>mean</b>	<b>bias</b>	<b>mse</b>
.4	.737	.337	.0378
.3	.555	.255	.0387
.2	.370	.170	.0394
.1	.186	.086	.0382
-.1	-.174	-.074	.0366
-.2	-.370	-.170	.0413
-.3	-.556	-.256	.0425
-.4	-.734	-.334	.0409

**Table 2 (.0032)**

$$l_n = 12; m_n = 128$$

<b>d</b>	<b>mean</b>	<b>bias</b>	<b>mse</b>
.4	.675	.275	.0267
.3	.512	.212	.0262
.2	.342	.142	.0276
.1	.170	.070	.0280
-.1	-.166	-.066	.0265
-.2	-.342	-.142	.0280
-.3	-.505	-.205	.0299
-.4	-.678	-.278	.0268

**Table 3 (.0023)**

$$l_n = 16; m_n = 180$$

<b>d</b>	<b>mean</b>	<b>bias</b>	<b>mse</b>
.4	.716	.316	.0143
.3	.537	.237	.0143
.2	.356	.156	.0153
.1	.180	.080	.0149
-.1	-.180	-.080	.0142
-.2	-.356	-.156	.0143
-.3	-.534	-.234	.0149
-.4	-.709	-.309	.0144

**Table 4 (.0023)**

$$l_n = 4; m_n = 180$$

<b>d</b>	<b>mean</b>	<b>bias</b>	<b>mse</b>
.4	.702	.302	.0177
.3	.528	.228	.0174
.2	.351	.151	.0192
.1	.176	.076	.0191
-.1	-.176	-.076	.0190
-.2	-.351	-.151	.0183
-.3	-.525	-.225	.0197
-.4	-.697	-.297	.0184

**Table 5 (.0023)**

$$l_n = 8; m_n = 180$$

Some general common features of the tables can be now extracted. Increasing  $m$  reduces the variance while, for fixed  $m$ , decreasing  $l$  reduces the variance as well. The effect on the bias of the estimate is more difficult to describe. Notice, however, the remarkable distance between the

theoretical limit variance and the relatively stable values of the column of mean squared errors. Since the conditions of theorem 3.1 are asymptotic in nature, the discrepancy might be attributed, in first instance, to the relatively small sample size  $n = 400$ . To check this point further,  $N = 500$  samples of size  $n = 1000$  are again generated from the same class of ARIMA(0,d,0) models as before. Results are displayed in tables 6 and 7.

d	mean	bias	mse
.4	.728	.328	.0078
.3	.550	.250	.0071
.2	.368	.168	.0070
.1	.182	.082	.0069
-.1	-.183	-.083	.0068
-.2	-.366	-.166	.0070
-.3	-.546	-.246	.0067
-.4	-.728	-.328	.0075

**Table 6 (.0010)**

$$l_n = 12; m_n = 400$$

--

d	mean	bias	mse
.4	.696	.296	.0056
.3	.526	.226	.0052
.2	.351	.151	.0053
.1	.170	.070	.0053
-.1	-.172	-.072	.0054
-.2	-.348	-.148	.0053
-.3	-.520	-.220	.0048
-.4	-.697	-.297	.0058

Table 7 (.0008)

$$l_n = 12; m_n = 490$$

Although in table 7 nothing apparently seems to have changed, at least in table 6 the theoretical asymptotic variance and the column of mean squared errors are of the same order.

## 5. FINAL COMMENTS

The simulation results presented in section 4 are, as previous work on trimming, non conclusive. In the light of tables 2, 3, 4, 5, 6, and 7, it could be conjectured that the sample size  $n$  needed for finding the trimming constants that fit with the statement of theorem 3.1 might be too large for flexible computing work. Further research along these lines is in progress.

## REFERENCES

Agiakloglou, Ch., Newbold, P., and Wohar, M. (1993), "Bias in an Estimator of the Fractional Difference Parameter," *Journal of Time Series*

*Analysis*, 14, 235 - 246.

Beran, J. (1994), *Statistics for Long - Memory Processes*, New York: John Wiley.

Brockwell, P.J., and Davis, R.A. (1991), *Time Series: Theory and Methods*, New York: Springer Verlag.

Geweke, J., and Porter-Hudak, S. (1983), "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221 - 238.

Granger, C.W.J. (1966), "The Typical Spectral Shape of an Economic Variable," *Econometrica*, 34, 150 - 161.

Hassler, U. (1995), "The Periodogram Regression: Correction and Comments," *Communications in Statistics - Theory and Methods*, 24, 3137 - 3146.

Hurvich, C.M., and Beltrao, K.I. (1993): "Asymptotics for the Low-Frequency Ordinates of the Periodogram of a Long-Memory Time Series," *Journal of Time Series Analysis*, 14, 455 - 472.

Hurvich, C.M., and Beltrao, K.I. (1994): "Automatic Semiparametric Estimation of the Memory Parameter of a Long-Memory Time Series," *Journal of Time Series Analysis*, 15, 285 - 302.

Künsch, H. (1986), "Discrimination Between Monotonic Trends and Long - Range Dependence," *Journal of Applied Probability*, 23, 1025 -1030.

Robinson, P.M. (1995), "Log - Periodogram Regression for Time Series with Long - Range Dependence," *The Annals of Statistics*, 23, 1048 - 1072.