98-F-1

Estimation of the Autocorrelation Function of a Stationary Time Series with Missing Observations

Yoshihiro Yajima
Faculty of Economics, University of Tokyo

Haruhisa Nishino
Faculty of Economics, University of Tokyo

January 1998

ESTIMATION OF THE AUTOCORRELATION FUNCTION OF A STATIONARY TIME SERIES * WITH MISSING OBSERVATIONS †

Yoshihiro Yajima

Faculty of Economics, University of Tokyo ‡

Haruhisa Nishino

Faculty of Economics, University of Tokyo §

January 9, 1998

*AMS(1991) *subject classifications.* 62M10, 90A20.

†*Key words*, autocorrelation, missing observations, long-memory time series

‡Hongo 7-3-1, Bunkyo-ku Tokyo JAPAN 113, yajima@e.u-tokyo.ac.jp

§Hongo 7-3-1, Bunkyo-ku Tokyo JAPAN 113, nishino@e.u-tokyo.ac.jp

Abstract

We consider three estimators of the autocorrelation function for a stationary process with missing observations. The first estimator is linked with the Yule-Walker estimator, the second one the least squares estimator, and the third one the sample correlation coefficient. We clarify their asymptotic differences and derive asymptotic distributions for both short-memory and long-memory models. In most of short-memory models the third estimator has smaller asymptotic variance than the others. On the other hand, if a process follows a long-memory model and its spectral density function is not square integrable, then the asymptotic distributions of the three estimators are all the same and the distribution is still the same as that of estimator under complete sampling too.

1 Introduction

Estimation of the autocorrelation function of a stationary time series is one of the important aspects throughout time series model building procedure. An estimator of the autocorrelation function is a fundamental tool in the identification stage to clarify the dependence structure of a time series. Also if we adopt a Newton-Raphson procedure to obtain the maximum likelihood estimator of a parametric time series model, we usually calculate the initial value by the method of moments based on it.

Hence asymptotic properties of estimators of the autocorrelation function, especially the sample autocorrelation function, the ratio of the sample covariance and the sample variance have been extensively studied in the literature. Bartlett (1946) and Anderson and Walker (1964) derived the asymptotic normality of the sample autocorrelation function under the assumption that the coefficients in the infinite order moving-average representation of a linear stationary process are absolutely summable. And recently Cavazos-Cadena (1994) and He (1996) extended their results to the case that the autocorrelation function is square summable, which is equivalent to the assumption that the spectral density function of a process is square integrable. Furthermore Hosking (1996) investigated the case that a stationary process follows a long-memory model and its spectral density function is not square integrable.

On the other hand, missing observations arise from a variety of causes such as machinery disorder, clerical error or the inability to observe data in a bad weather or on holidays [see, e.g., Robinson (1985)]. Many authors attempted to extend estimation procedures suggested for time series data under complete sampling to those with missing observations. In this note we focus on estimation of the autocorrelation function among them for the sake of motivation mentioned in the beginning.

We shall consider three estimators, two of which do not seem to have been discussed fully as yet and clarify the difference in their asymptotic properties. The first one is the sample autocorrelation function extended to the case with missing observations. Originally this estimator was proposed by Parzen (1963); its asymptotic properties were investigated in further details by Dunsmuir and Robinson (1981). The second one is a least-squares type estimator. Finally the third one is a sample correlation coefficient type estimator. The latter two ones were proposed by Takeuchi (1995). And Shin and Sarker (1995) adopted the second estimator as the initial value of a Newton-Raphson procedure for the maximum likelihood estimator in an AR(1) model. These estimators have the same asymptotic properties under complete sampling but as is shown later, they can behave asymptotically differently from each other in the presence of missing observations.

The rest of this note is organized as follows. Section 2 contains our model, notations, basic assumptions and introduces the three estimators. Section 3 presents asymptotic normality and asymptotic variances of the estimators under the assumption that the true underlying process follows a short-memory model or a long-memory model with a square integrable spectral density function. It is rather difficult to compare the differences among the asymptotic variances of the estimators in a general situation. In Section 4 we distinguish the differences among the estimators clearer in more specified models for a stationary process and a missing structure.

The results on the case that a process follows a long-memory model and its spectral density function is not square integrable, are stated in Section 5. In Section 6 some consideration is given to the properties of the estimators when the mean of a process is unknown. Finally we shall reinforce the theoretical results by computational experiments in Section 7.

Our main results are as follows. First the third estimator is the best one in most of short-memory models from the viewpoint of asymptotic variance. Next under the assumption that a process follows a long-memory model and its spectral density function is not square integrable, the three estimators have the same limiting distribution, more interestingly which is also identical with the limiting distributions of those under complete sampling. These results imply that there is no loss of information asymptotically even in the presence of missing observations if a process follows a long-memory model since there still exists a strong dependence between distant observations.

2 Model and Estimators

Let $\{X(t)\}$ be a stationary process of the form

$$X(t) = \sum_{j=0}^{\infty} \beta(j)\varepsilon(t-j), \quad \sum_{j=0}^{\infty} \beta(j)^2 < \infty,$$

where the white noise process $\{\varepsilon(t)\}$ consists of uncorrelated random variables with mean zero and variance σ_ε^2. Parzen (1963) introduced the time series model with missing observations as a specific case of an amplitude modulated stationary process. Following him we express observed data $\{Y(n)\}$ $(n = 1, 2, \ldots, N)$ by $Y(n) = a(n)X(n)$ where $\{a(n), n = 1, 2, \ldots\}$ represents the state of observation,

$$\begin{cases} a(n) = 1, & X(n) \text{ observed}, \\ a(n) = 0, & X(n) \text{ missing}. \end{cases} \tag{1}$$

We assume the same conditions as those in Parzen (1963) that $\{a(n)\}$ follows

$$\mu_a = \lim_{N\to\infty} \overline{a}, a.s., \tag{2}$$

$$\gamma_a(l) = \lim_{N\to\infty} (C_a(l) - \overline{a}^2), \ a.s., \tag{3}$$

$$\nu(r,s,u) = \lim_{N\to\infty} C_a(r,s,u), \ a.s., \tag{4}$$

where

$$\begin{aligned} \overline{a} &= \frac{1}{N}\sum_{n=1}^{N} a(n), \\ C_a(l) &= \frac{1}{N}\sum_{n=1}^{N-l} a(n)a(n+l), \ 0 \le l \le N-1, \\ C_a(r,s,u) &= \frac{1}{N}\sum_{n=1}^{N-max(r,s,u)} a(n)a(n+r)a(n+s)a(n+u), \ 0 \le r,s,u \le N-1. \end{aligned}$$

We call a process which satisfies (2), (3), and (4) an asymptotically 4th order stationary process in this paper. $\bar{X}, \bar{Y}, C_X(l)$, and $C_Y(l)$ are defined in the same way as $\bar{a}$ and $C_a(l)$ respectively. And we assume that $\{X(n)\}$ and $\{a(n)\}$ are independent hereafter.

Next we put $\gamma_X(l) = Cov(X(t), X(t+l))$, $\rho_X(l) = Cor(X(t), X(t+l)) = \gamma_X(l)/\gamma_X(0)$ and $\tilde{\gamma}_a(l) = \gamma_a(l) + \mu_a^2$. And let $f(\lambda)$ be the spectral density function of $\{X(n)\}$, $\gamma_X(l) = \int_{-\pi}^{\pi} e^{i\lambda l} f(\lambda) d\lambda$. $\gamma_Y(l)$ and $\rho_Y(l)$ are defined in the same way if $\{Y(n)\}$ is a stationary process. C implies a constant which is independent of N but is not always the same one in each context.

Now we introduce the three estimators of the autocorrelation function $\{\rho_X(l)\}$. The first one is proposed by Parzen (1963) and, later, investigated its asymptotic properties under various assumptions on $\{\varepsilon(n)\}$ by Dunsmuir and Robinson (1981). We denote it by $\hat{\rho}_{PDR}(l)$,

$$\hat{\rho}_{PDR}(l) \quad = \quad \frac{C_Y(l)/C_a(l)}{C_Y(0)/C_a(0)} = \frac{\sum_{n=1}^{N-l} Y(n)Y(n+l) / \sum_{n=1}^{N-l} a(n)a(n+l)}{\sum_{n=1}^{N} Y(n)^2 / \sum_{n=1}^{N} a(n)^2}.$$

Since the numerator and denominator are estimators of $\gamma_X(l)$ and $\gamma_X(0)$ respectively. $\hat{\rho}_{PDR}(l)$ is interpreted as a kind of the Yule-Walker estimator,

The second one is proposed by Takeuchi (1995) and is adopted independently by Shin and Sarker (1995) as the initial value of a Newton-Raphson procedure to obtain the maximum likelihood estimator in an AR(1) model. This estimator is defined by

$$\hat{\rho}_{SST}(l) = \frac{\sum_{n=1}^{N-l} Y(n)Y(n+l)}{\sum_{n=1}^{N-l} a(n+l)Y(n)^2}.$$

Noting $\hat{\rho}_{SST}(l) = \frac{\sum_{n=1}^{N-l} a(n)a(n+l)X(n)X(n+l)}{\sum_{n=1}^{N-l} a(n)a(n+l)X(n)^2}$, we see that $\hat{\rho}_{SST}(l)$ is the least squares estimator for the regression model which consists of all of the observed pairs $(X(n+l), X(n))$

with $X(n+l)$ a dependent variable and X(n) an independent variable respectively.

Finally the third estimator is also proposed by Takeuchi (1995) and defined by

$$\hat{\rho}_T(l) = \frac{\sum_{n=1}^{N-l} Y(n)Y(n+l)}{\sqrt{\sum_{n=1}^{N-l} a(n+l)Y(n)^2}\sqrt{\sum_{n=1}^{N-l} a(n)Y(n+l)^2}}.$$

$\hat{\rho}_T(l)$ is the sample correlation coefficients based on all of the observed pairs $(X(n+l), X(n))$.

It is easily seen that these estimators have the same limiting distribution under complete sampling ($a(n) \equiv 1$). However as is shown later, though they are asymptotically normally distributed, the variances of their limiting distributions are different from each other if the spectral density function is square integrable. While if the spectral density function is no longer square integrable, their limiting distributions are all the same, and, more interestingly, are identical with that of estimators under complete sampling.

3 The limiting distribution (The spectral density is square integrable)

We impose the same assumptions on $\{\varepsilon(n)\}$ and $f(\lambda)$ as those in Dunsmuir and Robinson (1981) throughout this section.

$\{\varepsilon(n)\}$ is a strictly stationary and ergodic process. Let $\mathcal{F}_\epsilon(n)$ be the σ-algebra of events generated by $\{\varepsilon(m); m \leq n\}$. And we introduce the following assumptions,

Assumption 1. (i)

$$\begin{aligned} E(\varepsilon(n)|\mathcal{F}_\epsilon(n-1)) &= 0, \ a.s., \\ E(\varepsilon(n)^2|\mathcal{F}_\epsilon(n-1)) &= \sigma^2, \ a.s., \end{aligned}$$

$$E(\varepsilon(n)^4|\mathcal{F}_\epsilon(n-1)) = 3\sigma^4 + \kappa, \ a.s.,$$

and $E(\varepsilon(n)^3|\mathcal{F}_\epsilon(n-1))$ is a constant, a.s. Hence κ is the fourth-order cumulant of $\{\varepsilon(n)\}$.

(ii) $f(\lambda)$ is square integrable.

Here we note a relation between Assumption 1 (ii) and a correlation structure of a stationary process.

Assumption 1 (ii) is equivalent that the autocovariance function is square summable, $\sum_{h=0}^{\infty} \gamma_X(h)^2 < \infty$. If the autocovariance function of a time series model satisfies absolute summability $\sum_{h=0}^{\infty} |\gamma_X(h)| < \infty$, this model is called a short-memory model. Clearly the spectral density function of a short-memory model is square integrable. A typical example is an ARMA model since its autocovariance function satisfies $|\gamma_X(h)| \leq C\tau^{|h|}$ with a constant τ $(0 < \tau < 1)$.

On the other hand there are many actual time series in which the dependence between distant observations are not negligible.

Long-memory models have been developed recently in order to explain such a series. Two popular models are proposed. One is a fractional ARIMA model by Granger and Joyeux (1980) and Hosking (1981). The other is a fractional Gaussian noise by Mandelbrot and Van Ness (1968). Their autocovariance function is characterized by

$$\gamma_X(h) \sim L(h)h^{2d-1}, \quad 0 < d < \frac{1}{2}, \tag{5}$$

where $\sim$ means that the ratio of the left and right hand side terms converges to 1 as $h \to \infty$ and $L(h)$ is a slowly varying function of infinity,

$$\frac{L(ht)}{L(h)} \to 1, \qquad as\ h \to \infty\ for\ all\ t > 0.$$

It follows from (5) that $\gamma_X(h)$ is no longer absolute summable. Hence Assumption 1 (ii) includes short-memory models and long-memory models with $0 < d < \frac{1}{4}$. We shall consider the case of long-memory models with $\frac{1}{4} \le d < \frac{1}{2}$ in the following section.

Now we consider asymptotic properties of the estimators. The limiting distribution of $\hat{\rho}_{PDR}(l)$ has already been derived by Dunsmuir and Robinson (1981). We shall extend their procedure to obtain the limiting distributions of $\hat{\rho}_{SST}(l)$ and $\hat{\rho}_T(l)$. For this purpose we define the random variables $b(k_a, k_{1x}, k_{2x})$ $(k_a, k_{1x}, k_{2x} = 0, 1, 2, \ldots)$ by

$$b(k_a, k_{1x}, k_{2x}) = N^{-1/2} \sum_{n=1}^{N} a(n)a(n+k_a)\Big(X(n+k_{1x})X(n+k_{2x}) - \gamma_X(k_{2x} - k_{1x})\Big)$$

And we note that similar to Theorem 1 of Dunsmuir and Robinson (1981), we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} a(n+k_1)a(n+k_2)X(n+l_1)X(n+l_2) = \tilde{\gamma}_a(k_2 - k_1)\gamma_X(l_2 - l_1), \; a.s. \qquad (6)$$

Now we shall express $\hat{\rho}_{SST}(l)$ and $\hat{\rho}_T(l)$ in terms of $b(k_a, k_{1x}, k_{2x})$, $\tilde{\gamma}_a(l)$, $\gamma_X(0)$, and $\rho_X(l)$. Let $\tilde{\rho}_{SST}(l)$ and $\tilde{\rho}_T(l)$ be defined in the same way as $\hat{\rho}_{SST}(l)$ and $\hat{\rho}_T(l)$ except for the summation $\sum_{n=1}^{N-l}$ being replaced by $\sum_{n=1}^{N}$.

Then

$$\begin{aligned} N^{1/2}(\hat{\rho}_T(l) - \rho_X(l)) &= N^{1/2}(\tilde{\rho}_T(l) - \rho_X(l)) + o_p(1) \\ &= \frac{N^{-1/2}[\sum_{n=1}^{N} Y(n)Y(n+l) - \rho_X(l)\sqrt{\sum_{n=1}^{N} a(n+l)Y(n)^2}\sqrt{\sum_{n=1}^{N} a(n)Y(n+l)^2}]}{\sqrt{\frac{1}{N}\sum_{n=1}^{N} a(n+l)Y(n)^2}\sqrt{\frac{1}{N}\sum_{n=1}^{N} a(n)Y(n+l)^2}} + o_p(1). \end{aligned} \qquad (7)$$

The numerator of (7) is expanded by

$$N^{-1/2} \sum_{n=1}^{N} a(n)a(n+l)\left(X(n)X(n+l) - \gamma_X(l)\right)$$

$$
\begin{aligned}
&-N^{-1/2}\rho_X(l)\left[\sqrt{\Sigma_{n=1}^N a(n+l)Y(n)^2}\sqrt{\Sigma_{n=1}^N a(n)Y(n+l)^2} - \gamma_X(0)\sum_{n=1}^N a(n)a(n+l)\right] \\
= \;& b(l,0,l) \\
&-N^{-1/2}\rho_X(l)\left[\gamma_X(0)\sum_{n=1}^N a(n)a(n+l)\sqrt{\left(1+\frac{\sum_{n=1}^N a(n)a(n+l)(X(n)^2-\gamma_X(0))}{\gamma_X(0)\sum_{n=1}^N a(n)a(n+l)}\right)}\right. \\
&\left.\times\sqrt{\left(1+\frac{\sum_{n=1}^N a(n)a(n+l)(X(n+l)^2-\gamma_X(0))}{\gamma_X(0)\sum_{n=1}^N a(n)a(n+l)}\right)}-1\right] \\
= \;& b(l,0,l) \\
&-\frac{N^{-1/2}\rho_X(l)}{2}\left[\sum_{n=1}^N a(n)a(n+l)\{(X(n)^2-\gamma_X(0))+(X(n+l)^2-\gamma_X(0))\}\right]+o_p(1) \\
= \;& b(l,0,l)-\frac{1}{2}\rho_X(l)(b(l,l,l)+b(l,0,0))+o_p(1). \qquad (8)
\end{aligned}
$$

The second equality is derived by noting that $\sqrt{1+x} = 1+\frac{x}{2}+o(x^2)$ and $\left(\sum_{n=1}^N a(n)a(n+l)(X(n+k)^2-\gamma_X(0))\right)/\sum_{n=1}^N a(n)a(n+l)$ $(k=0,l)$ is $O_p(1/\sqrt{N})$.

Then it follows from (6), (7), and (8) that

$$
N^{1/2}(\hat{\rho}_T(l)-\rho(l)) = \frac{b(l,0,l)-\rho_X(l)(b(l,l,l)+b(l,0,0))/2}{\tilde{\gamma}_a(l)\gamma_X(0)}+o_p(1). \qquad (9)
$$

Similarly

$$
N^{1/2}(\hat{\rho}_{SST}(l)-\rho(l)) = \frac{b(l,0,l)-\rho_X(l)b(l,0,0)}{\tilde{\gamma}_a(l)\gamma_X(0)}+o_p(1). \qquad (10)
$$

Finally we express $\hat{\rho}_{PDR}(l)$ in terms of $\hat{\rho}_{SST}(l)$ and $b(k_a,k_{1x},k_{2x})$ to compare the variance of the limiting distributions of $\hat{\rho}_{PDR}(l)$ with that of $\hat{\rho}_{SST}(l)$. We have

$$
\begin{aligned}
&N^{1/2}\left(\hat{\rho}_{PDR}(l)-\rho_X(l)\right) \\
= \;& N^{1/2}\left(\hat{\rho}_{SST}(l)-\rho_X(l)\right)\frac{\sum_{n=1}^{N-l}a(n)a(n+l)X(n)^2}{\sum_{n=1}^N a(n)^2X(n)^2}\frac{\sum_{n=1}^N a(n)^2}{\sum_{n=1}^{N-l}a(n)a(n+l)} \\
&+N^{1/2}\rho_X(l)\left(\frac{\sum_{n=1}^{N-l}a(n)a(n+l)X(n)^2}{\sum_{n=1}^N a(n)^2X(n)^2}\frac{\sum_{n=1}^N a(n)^2}{\sum_{n=1}^{N-l}a(n)a(n+l)}-1\right). \qquad (11)
\end{aligned}
$$

The second term of (11) is equal to

$$N^{1/2}\rho_X(l)\left[\frac{\sum_{n=1}^{N} a(n)^2 \sum_{n=1}^{N-l} a(n)a(n+l)\Big(X(n)^2-\gamma_X(0)\Big)}{\sum_{n=1}^{N} a(n)^2 X(n)^2 \sum_{n=1}^{N-l} a(n)a(n+l)}\right]$$

$$-N^{1/2}\rho_X(l)\left[\frac{\sum_{n=1}^{N} a(n)^2\Big(X(n)^2-\gamma_X(0)\Big)}{\sum_{n=1}^{N} a(n)^2 X(n)^2}\right]$$

$$= \frac{\rho_X(l)[\tilde{\gamma}_a(0)b(l,0,0)-\tilde{\gamma}_a(l)b(0,0,0)]}{\tilde{\gamma}_a(0)\tilde{\gamma}_a(l)\gamma_X(0)} + o_p(1). \quad (12)$$

The last equality follows from (2), (3), and (6). Hence from (2), (3), (6), (11), and (12) we have

$$N^{1/2}\Big(\hat{\rho}_{PDR}(l)-\rho_X(l)\Big)$$

$$= N^{1/2}\Big(\hat{\rho}_{SST}(l)-\rho_X(l)\Big) + \frac{\rho_X(l)[\tilde{\gamma}_a(0)b(l,0,0)-\tilde{\gamma}_a(l)b(0,0,0)]}{\tilde{\gamma}_a(0)\tilde{\gamma}_a(l)\gamma_X(0)} + o_p(1). \quad (13)$$

The asymptotic distribution of $\hat{\rho}_{PDR}(l)$ has already been derived by Dunsmuir and Robinson (1981). However the relation (13) is useful to distinguish a difference between the asymptotic variances of $\hat{\rho}_{PDR}(l)$ and of $\hat{\rho}_{SST}(l)$ clearer. Since if the sum of the variance of the second term and the two times covariance of the first and second terms of (13) is positive (negative) asymptotically, it implies that the variance of $N^{1/2}(\hat{\rho}_{PDR}(l)-\rho_X(l))$ is greater (smaller) than that of $N^{1/2}(\hat{\rho}_{SST}(l)-\rho_X(l))$ asymptotically.

Now from (9), (10), and (13), it suffices to derive the limiting distribution of $b(k_a, k_{1x}, k_{2x})$ and to evaluate its covariances matrices in order to consider asymptotic behaviors of the estimators.

Proposition 1 *Let $\{X(n)\}$ satisfy Assumption 1. And let $\{a(n)\}$ be an asymptotically 4th order stationary process which satisfies* (2), (3), *and* (4). *Then any finite set of the $b(k_a, k_{1x}, k_{2x})$ with $\tilde{\gamma}_a(k_a) \neq 0$ are asymptotically normally distributed with zero means and* *covariances*

$$\lim_{N\to\infty} Cov(b(k_a, k_{1x}, k_{2x}), b(l_a, l_{1x}, l_{2x}))$$
$$= \sum_{u=-\infty}^{\infty} \nu(k_a, u + k_{1x} - l_{1x}, u + k_{1x} - l_{1x} + l_a)\gamma_X(u)\gamma_X(u + k_{1x} - k_{2x} - l_{1x} + l_{2x})$$
$$+ \sum_{u=-\infty}^{\infty} \nu(k_a, u + k_{1x} - l_{2x}, u + k_{1x} - l_{2x} + l_a)\gamma_X(u)\gamma_X(u + k_{1x} - k_{2x} + l_{1x} - l_{2x})$$
$$+\kappa \sum_{u=-\infty}^{\infty} \nu(k_a, u + k_{1x} - l_{1x}, u + k_{1x} - l_{1x} + l_a)$$
$$\times \sum_{r=0}^{\infty} \beta(r)\beta(r - k_{1x} + k_{2x})\beta(r + u)\beta(r + u - l_{1x} + l_{2x}). \tag{14}$$

The proof of Proposition 1 is omitted since the result is derived in the same way as Theorem 2 of Dunsmuir and Robinson (1981) with a slight modification.

Then the following theorem follows immediately from Proposition 1, (9), (10), and (13).

Theorem 1 *$N^{1/2}(\hat{\rho}_{SST}(l) - \rho_X(l))$, $N^{1/2}(\hat{\rho}_T(l) - \rho_X(l))$, and $N^{1/2}(\hat{\rho}_{PDR}(l) - \rho_X(l))$ are asymptotically normally distributed with zero mean and variance $AV_{SST}(l)$, $AV_T(l)$ and $AV_{PDR}(l)$ respectively where*

$$AV_{SST}(l) = \frac{1}{\tilde{\gamma}_a(l)^2\gamma_X(0)^2} \lim_{N\to\infty} Var(b(l,0,l) - \rho_X(l)b(l,0,0)),$$
$$AV_T(l) = \frac{1}{\tilde{\gamma}_a(l)^2\gamma_X(0)^2} \lim_{N\to\infty} Var\Big(b(l,0,l) - \rho_X(l)\big(b(l,0,0) + b(l,l,l)\big)/2\Big)$$

and

$$AV_{PDR}(l) = AV_{SST}(l)$$
$$+\frac{2\rho_X(l)}{\tilde{\gamma}_a(0)\tilde{\gamma}_a(l)^2\gamma_X(0)^2} \lim_{N\to\infty} Cov\Big[\Big(b(l,0,l) - \rho_X(l)b(l,0,0)\Big), \Big(\tilde{\gamma}_a(0)b(l,0,0) - \tilde{\gamma}_a(l)b(0,0,0)\Big)\Big]$$
$$+\frac{\rho_X(l)^2}{\tilde{\gamma}_a(0)^2\tilde{\gamma}_a(l)^2\gamma_X(0)^2} \lim_{N\to\infty} Var\Big[\tilde{\gamma}_a(0)b(l,0,0) - \tilde{\gamma}_a(l)b(0,0,0)\Big].$$

4 Examples (The spectral density is square integrable)

The asymptotic variances of the estimators in Theorem 1, $AV_{SST}(l)$, $AV_T(l)$ and $AV_{PDR}(l)$, are rather complicated in a general situation of $\{X(n)\}$ and $\{a(n)\}$. Hence we shall impose more specific assumptions at least on $\{X(n)\}$ or $\{a(n)\}$ to derive a clear-cut ranking among the estimators. Throughout this section $\{X(n)\}$ is assumed to be a Gaussian stationary process, which implies $\kappa = 0$. Hence the third term of (14) vanishes.

Now we shall consider two examples. In the first example we assume that $\{a(n)\}$ is a sequence of Bernoulli trials. However we do not need any additional condition on $\{X(n)\}$. In the second one we specify that $\{X(n)\}$ is an $AR(1)$ process, $X(n) - \phi X(n-1) = \varepsilon(n)$ but do not impose any other condition on $\{a(n)\}$ except that $\{a(n)\}$ is asymptotically 4th order stationary. First we have the following theorem.

Theorem 2 *Let $\{X(n)\}$ be a Gaussian stationary process and $\{a(n)\}$ a sequence of Bernoulli trials with $p = Pr\{a(n) = 1\}$. Then*

(i) *$AV_T(l) \leq AV_{SST}(l)$ for any l ($l = 1, 2, \ldots$).*

(ii) *$AV_{SST}(l) \leq AV_{PDR}(l)$ holds if and only if $0 \leq \frac{1}{p} + 2(\rho_X(2l) - \rho_X(l)^2)$ or $\rho_X(l) = 0$.*

Proof. (i) First we note that $\tilde{\gamma}_a(0) = E[a(n)^2] = p$ and $\tilde{\gamma}_a(l) = E[a(n)a(n+l)] = p^2$ for $l \geq 1$. Then it follows from Proposition 1 and Theorem 1 that

$$\begin{aligned} p^4 AV_{SST}(l) &= \sum_u \nu(l,u,u+l)\rho_X(u)^2 + \sum_u \nu(l,u,u-l)\rho_X(u)\rho_X(u-2l) \\ &\quad -4\rho_X(l)\sum_u \nu(l,u,u-l)\rho_X(u)\rho_X(u-l) \\ &\quad +2\rho_X(l)^2 \sum_u \nu(l,u,u+l)\rho_X(u)^2. \end{aligned} \tag{15}$$

Similarly

$$
\begin{aligned}
p^4 AV_T(l) \;=\; & \sum_u \nu(l,u,u+l)\rho_X(u)^2 + \sum_u \nu(l,u,u-l)\rho_X(u)\rho_X(u-2l) \\
& -2\rho_X(l)\sum_u \nu(l,u,u+l)\rho_X(u)\rho_X(u-l) \\
& -2\rho_X(l)\sum_u \nu(l,u,u-l)\rho_X(u)\rho_X(u-l) \\
& +\rho_X(l)^2\sum_u \nu(l,u,u+l)\rho_X(u)^2 \\
& +\rho_X(l)^2\sum_u \nu(l,u+l,u+2l)\rho_X(u)^2. \qquad (16)
\end{aligned}
$$

Next we have

$$
\nu(l,u,u+l) = \begin{cases} p^2, & u=0, \\ p^3, & u=-l,l, \\ p^4, & otherwise, \end{cases} \qquad (17)
$$

$$
\nu(l,u,u-l) = \begin{cases} p^2, & u=l, \\ p^3, & u=0,2l, \\ p^4, & otherwise, \end{cases} \qquad (18)
$$

and

$$
\nu(l,u+l,u+2l) = \begin{cases} p^2, & u=-l, \\ p^3, & u=-2l,0, \\ p^4, & otherwise. \end{cases} \qquad (19)
$$

And note that $\nu(j,k,l)$ are symmetric with respect to their arguments. Then it follows from (15), (16), (17), (18), and (19) with an elementary calculation that

$$
\begin{aligned}
p^4\Big(AV_T(l) - AV_{SST}(l)\Big) \;&=\; \rho_X(l)^2 p^2(1-p)\Big[p\rho_X(2l)^2 + (1-p)\rho_X(l)^2 - 1\Big] \\
&\le\; \rho_X(l)^2 p^2(1-p)\Big[p + (1-p) - 1\Big] = 0. \qquad (20)
\end{aligned}
$$

Hence we have the result. ∎

(ii) It follows from Proposition 1 and Theorem 1 that

$$\begin{aligned} &p^5\Big(AV_{PDR}(l) - AV_{SST}(l)\Big) \\ &= 4\rho_X(l)\Big[p\sum_u \nu(l,u,u+l)\rho_X(u)\rho_X(u-l) - p^2\sum_u \nu(l,u,u)\rho_X(u)\rho_X(u-l) \\ &\quad -p\rho_X(l)\sum_u \nu(l,u,u+l)\rho_X(u)^2 + p^2\rho_X(l)\sum_u \nu(l,u,u)\rho_X(u)^2\Big] \\ &\quad +\frac{2}{p}\rho_X(l)^2\Big[p^2\sum_u \nu(l,u,u+l)\rho_X(u)^2 - 2p^3\sum_u \nu(l,u,u)\rho_X(u)^2 \\ &\quad +p^4\sum_u \nu(0,u,u)\rho_X(u)^2\Big]. \end{aligned} \tag{21}$$

And

$$\nu(l,u,u) = \begin{cases} p^2, & u = 0, l, \\ p^3, & otherwise, \end{cases} \tag{22}$$

and

$$\nu(0,u,u) = \begin{cases} p, & u = 0, \\ p^2, & u \neq 0. \end{cases} \tag{23}$$

Then from (17), (21), (22), and (23) we have

$$\begin{aligned} p^5(AV_{PDR}(l) - AV_{SST}(l)) &= 4\rho_X(l)^2(p^4-p^5)\Big(\rho_X(2l) - \rho_X(l)^2\Big) + \frac{2}{p}\rho_X(l)^2(p^4-p^5) \\ &= 2\rho_X(l)^2(p^4-p^5)\Big(\frac{1}{p} + 2\Big(\rho_X(2l) - \rho_X(l)^2\Big)\Big). \end{aligned} \tag{24}$$

Hence we have the result. ∎

The following corollary shows a few sufficient conditions which ensure that $AV_{SST}(l) \leq AV_{PDR}(l)$.

Corollary 1 *If at least one of the conditions:* (i) $p \leq \frac{1}{2}$, (ii) $\frac{1}{2} \leq \rho_X(l)^2$, *or* (iii) $\rho_X(l)^2 < \frac{1}{2}$ *and* $0 \leq \rho_X(2l)$ *holds, then* $AV_{SST}(l) \leq AV_{PDR}(l)$.

Proof. First we note that the partial correlation between $X(1)$ and $X(2l+1)$ given $X(l+1)$ is $\Big(\rho_X(2l)-\rho_X(l)^2\Big)/\Big(1-\rho_X(l)^2\Big)$. Hence

$$\Big|\rho_X(2l)-\rho_X(l)^2\Big| \le 1-\rho_X(l)^2 \le 1. \tag{25}$$

Hence clearly $0 \le \frac{1}{p} + 2(\rho_X(2l)-\rho(l)^2)$ if $p \le \frac{1}{2}$. Next if $\frac{1}{2} \le \rho_X(l)^2$, from (25)

$$0 \le 2\rho_X(l)^2 - 1 \le \frac{1}{p} - 2\Big(1-\rho_X(l)^2\Big) \le \frac{1}{p} + 2\Big(\rho_X(2l)-\rho_X(l)^2\Big).$$

Finally if $\rho_X(l)^2 < \frac{1}{2}$ and $0 < \rho_X(2l)$, we have

$$0 \le 2\rho_X(2l) = 1 + 2\Big(\rho_X(2l) - \frac{1}{2}\Big) < \frac{1}{p} + 2\Big(\rho_X(2l)-\rho_X(l)^2\Big).$$

Hence the proof is completed. ∎

A definite inequality between $AV_{SST}(l)$ and $AV_{PDR}(l)$ does not hold for $\rho_X(2l) < 0$ and $0 < \rho_X(l)^2 < \frac{1}{2}$. However we do not have any example in which $\{X(n)\}$ satisfies the assumption in this note and $AV_{SST}(l) > AV_{PDR}(l)$ is true at the moment.

Next we show the result on the second example.

Theorem 3 *Let $\{X(n)\}$ be a Gaussian $AR(1)$ process and $\{a(n)\}$ an asymptotically 4th order stationary process which satisfies* (2), (3), *and* (4).

$$AV_T(1) \le AV_{SST}(1).$$

Proof. Put $l=1$ and $\rho_X(u)=\phi^{|u|}$ and replace p^4 by $\tilde{\gamma}_a(1)^2$ in (15) and (16). Then

$$\begin{aligned} &\tilde{\gamma}_a(1)^2\Big(AV_T(1) - AV_{SST}(1)\Big) \\ &= 2\phi\sum_u\Big\{\nu(1,u,u+1)-\nu(1,u-1,u)\Big\}\phi^{|u|+|u-1|} \\ &\quad +\phi^2\sum_u\Big\{\nu(1,u+1,u+2)-\nu(1,u,u+1)\Big\}\phi^{2|u|}. \end{aligned} \tag{26}$$

Noting that $\sum_u \nu(1,u-1,u)\phi^{|u|+|u-1|} = \sum_u \nu(1,u,u+1)\phi^{|u|+|u+1|}$, $\sum_u \nu(1,u+1,u+2)\phi^{2|u|} = \sum_u \nu(1,u,u+1)\phi^{2|u-1|}$ and $\nu(1,-u+1,-u+1) = \nu(1,u,u+1)$ for $u > 0$, we have

$$\begin{aligned}&\tilde{\gamma}_a(1)^2\Big(AV_T(1) - AV_{SST}(1)\Big)\\ &= (1-\phi^2)\Big[(1-\phi^2)\sum_{u=1}^{\infty}\nu(1,u,u+1)\phi^{2u} - \nu(1,0,1)\phi^2\Big]. \qquad (27)\end{aligned}$$

From the definition and Schwarz's inequality we have

$$\begin{aligned}&\nu(1,u,u+1)\\ &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N-(u+1)} a(n)a(n+1)a(n+u)a(n+u+1)\\ &\leq \lim_{N\to\infty}\Big[\frac{1}{N}\sum_{n=1}^{N-(u+1)} a(n)^2a(n+1)^2\Big]^{1/2}\Big[\frac{1}{N}\sum_{n=1}^{N-(u+1)} a(n+u)^2a(n+u+1)^2\Big]^{1/2}\\ &= \nu(0,1,1).\end{aligned}$$

Hence the right-hand side of (27) is bounded by

$$(1-\phi^2)\nu(0,1,1)\Big[(1-\phi^2)\sum_{u=1}^{\infty}\phi^{2u} - \phi^2\Big] = 0.$$

Then the proof is completed. ∎

We are not able to find a clear-cut relation between $AV_{SST}(1)$ and $AV_{PDR}(1)$. Though we omit a laborsome detail, it suffices to evaluate by the same argument as in Theorem 2 (ii) that

$$\begin{aligned}&AV_{PDR}(1) - AV_{SST}(1)\\ &= 2\tilde{\gamma}_a(1)^{-2}\Big[2\sum_{u=1}^{\infty}\nu(1,u,u+1)\phi^{2u} + \nu(1,0,1)\phi^2\Big]\\ &\quad -4\Big(\tilde{\gamma}_a(0)\tilde{\gamma}_a(1)\Big)^{-1}\Big[\sum_{u=0}^{\infty}\nu(0,u,u+1)\phi^{2u+2} + \sum_{u=1}^{\infty}\nu(1,u,u)\phi^{2u}\Big]\end{aligned}$$

$$+\tilde{\gamma}_a(0)^{-2}\Big[4\sum_{u=1}^{\infty}\nu(0,u,u)\phi^{2u+2}\nu(0,0,0)\phi^2\Big]. \tag{28}$$

Later we evaluate (28) by considering a more specific example of $\{a(n)\}$.

Remark 1. (i) Let $\{X(n)\}$ be an $AR(1)$ process. Then $\rho(2l)=\rho(l)^2=\phi^{2l}$ $(l\geq 1)$. Hence it follows from Theorem 2 that

$$AV_T(l)\leq AV_{SST}(l)\leq AV_{PDR}(l),$$

for any $l(\geq 1)$. Especially if we put $l=1$, it follows from (15), (17), and (18), after some calculation that

$$AV_{SST}(1)=\frac{1}{p^2}(1-\phi^2). \tag{29}$$

Thus we have from (20) and (29)

$$\begin{aligned}AV_T(1) &= AV_{SST}(1)+\frac{\phi^2(1-p)(p\phi^4+(1-p)\phi^2-1)}{p^2}\\ &= \frac{(1-\phi^2)[1-(1-p)\phi^2(1+p\phi^2)]}{p^2},\end{aligned} \tag{30}$$

and from (24) and (29)

$$\begin{aligned}AV_{PDR}(1) &= AV_{SST}(1)+\frac{2\phi^2(1-p)}{p^2}\\ &= \frac{1+\phi^2(1-2p)}{p^2}.\end{aligned} \tag{31}$$

The relation (31) is equal to (6.9) of Dunsmuir and Robinson (1981). The relations (30) and (31) also follow from (26), (28), and (29).

(ii) In order to evaluate (28), we consider an $A-B$ sampling in which A consecutive values are observed and next B values are missing. Then the first term of the right-hand side of (28) is

$$\sum_{u=1}^{\infty} \nu(1,u,u+1)\phi^{2u} = \sum_{u=1}^{A+B} \nu(1,u,u+1)\phi^{2u}\left[\sum_{i=1}^{\infty}\phi^{2(A+B)(i-1)}\right] = \frac{1}{1-\phi^{2(A+B)}}\sum_{u=1}^{A+B}\nu(1,u,u+1)\phi^{2u}.$$

The other terms are obtained similarly. And $\tilde{\gamma}_a(0) = \frac{A}{(A+B)}$, $\tilde{\gamma}_a(1) = \frac{(A-1)}{(A+B)}$. And $\nu(i,u,u+1)$ and $\nu(i,u,u)$ $(i = 0,1)$, are derived straightforwardly. Here we put $A = 3,4$ to make the relation clearer. Then after an elementary calculation,

$$AV_{PDR}(1) - AV_{SST}(1) = \begin{cases} -\frac{4\phi^4(\phi^2-1)^2(3\phi^2+4)}{9(1-\phi^8)}, & A = 3, \\ \frac{5\phi^2(\phi^2-1)^2(-6\phi^6-9\phi^4-\phi^2+2)}{36(1-\phi^{10})}, & A = 4. \end{cases}$$

Clearly $AV_{PDR}(1) \leq AV_{SST}(1)$ for $A = 3$. While if we put $f(x) = -6x^3 - 9x^2 - x + 2$ with $x = \phi^2$, then $f(0) = 2$, $f(1) = -14$ and $f(x)$ is a decreasing function of x in $[0,1]$. Hence if ϕ^2 is closer to 0, $AV_{PDR}(1) \geq AV_{SST}(1)$ and contrarily if ϕ^2 is closer to 1, $AV_{PDR}(1) \leq AV_{SST}(1)$ for $A = 4$. In Section 7 we shall show by computational experiments that the larger A is, $AV_{SST}(1)$ achieves the better performance than $AV_{PDR}(1)$.

5 The limiting distribution (The spectral density is not square integrable)

We shall consider the case that $f(\lambda)$ is not square integrable in this section. Hence we assume that $\gamma_X(h)$ satisfies (5) with $\frac{1}{4} \leq d < \frac{1}{2}$ and that $L(h) = \alpha$ with a constant α for simplicity since the result in this section can be generalized to any slowly varying function $L(h)$. And let $\{X(n)\}$ be a Gaussian process. The consideration for a non-Gaussian process is given in the end of this section.

Hence we impose the following assumption on $\{X(n)\}$.

Assumption 2. $\{X(n)\}$ is a Gaussian stationary process and $\gamma_X(h)$ satisfies

$$\gamma_X(h) \sim \alpha h^{2d-1}, \quad \frac{1}{4} \leq d < \frac{1}{2}. \tag{32}$$

Assumption 3. $\{a(n)\}$ is a 4th-order stationary ergodic process and if we define $\gamma_{aak}(h)$ by

$$\gamma_{aak}(h) = Cov(a(t)a(t+k), a(t+h)a(t+k+h)),$$

then

$$\gamma_{aak}(h) \longrightarrow 0 \quad (h \to \infty),$$

for any fixed $k(\geq 0)$.

Assumption 4. $\{a(n)\}$ is deterministic and periodical with period M, that is, $a(n) = a(n+M)$ for any $n(\geq 1)$.

Then the relations (2), (3), and (4) are satisfied under both Assumptions 3 and 4. We put the sample size $N = N^*M$ without loss generality under Assumption 4.

First we investigate the asymptotic behavior of $b(k_a, k_{1x}, k_{2x})$.

Lemma 1 *Let $\{X(n)\}$ be a Gaussian stationary process and $\{\gamma_X(h)\}$ satisfy (32). And let $\{a(n)\}$ satisfy Assumption 3 or 4.*

(i) *If $d = \frac{1}{4}$,*

$$\frac{b(k_a, k_{1x}, k_{2x})}{\sqrt{\log N}} = \frac{\tilde{\gamma}_a(k_a)}{\sqrt{N \log N}} \sum_{n=1}^{N} [X(n)^2 - \gamma_X(0)] + o_p(1), \quad as\ N \to \infty.$$

(ii) *If* $\frac{1}{4} < d < \frac{1}{2}$,

$$\frac{b(k_a, k_{1x}, k_{2x})}{N^{2d-1/2}} = \frac{\tilde{\gamma}_a(k_a)}{N^{2d}} \sum_{n=1}^{N} [X(n)^2 - \gamma_X(0)] + o_p(1), \quad as\ N \to \infty.$$

Proof. We only prove (i) since (ii) is shown in the same way. First we impose Assumption 3 on $\{a(n)\}$. Then

$$\begin{aligned} &E\Big[\sum_{n=1}^{N}\Big(a(n)a(n+k_a) - \tilde{\gamma}_a(k_a)\Big)\Big(X(n+k_{1x})X(n+k_{2x}) - \gamma_X(k_{2x}-k_{1x})\Big)\Big]^2 \\ &= \sum_n \sum_m \gamma_{aak_a}(n-m)\Big[\gamma_X(n-m)^2 + \gamma_X(n+k_{1x}-m-k_{2x})\gamma_X(n+k_{2x}-m-k_{1x})\Big]. \end{aligned} \tag{33}$$

For any $\epsilon > 0$, there exists H such that $|\gamma_{aak_a}(h)| < \epsilon$ for any $h \geq H$. Then for sufficiently large N,

$$\begin{aligned} |\sum_n \sum_m \gamma_{aak_a}(n-m)\gamma_X(n-m)^2| &= |\sum_{|h| \leq N-1} (N - |h|)\gamma_{aak_a}(h)\gamma_X(h)^2| \\ &\leq N\gamma_{aak_a}(0) \sum_{|h|<H} \gamma_X(h)^2 + N\epsilon \sum_{H \leq |h| \leq N-1} \gamma_X(h)^2 \leq \epsilon C(\log N)N. \end{aligned}$$

The second term of (33) is evaluated in the same way. Hence

$$\begin{aligned} &(\log N)^{-1/2} b(k_a, k_{1x}, k_{2x}) \\ &= (N\log N)^{-1/2}\tilde{\gamma}_a(k_a)\sum_n \Big[X(n+k_{1x})X(n+k_{2x}) - \gamma_X(k_{2x}-k_{1x})\Big] + o_p(1). \end{aligned} \tag{34}$$

While as in the proof of Theorem 2.1 (2) of Yajima (1992), we have

$$\begin{aligned} Var\Big[\sum_{n=1}^{N}\Big\{\Big(X(n+k_{1x})X(n+k_{2x}) - X(n)^2\Big) - \Big(\gamma_X(k_{2x}-k_{1x}) - \gamma_X(0)\Big)\Big\}\Big] \\ = o(N\log N). \end{aligned} \tag{35}$$

Then the assertion follows from (34) and (35).

Next we consider the case that $\{a(n)\}$ satisfies Assumption 4. And we note that $\tilde{\gamma}_a(k_a) = \frac{1}{M}\sum_{r=1}^{M} a(r)a(r+k_a)$. Now we show that

$$(\log N)^{-1/2}\Big[b(k_a, k_{1x}, k_{2x}) - b(k_a, k_{1x}, k_{2x}-1)\Big] = o_p(1). \tag{36}$$

Put $k_x = k_{2x} - k_{1x}$. Then

$$\begin{aligned}
&(\log N)^{-1/2}\Big[b(k_a, k_{1x}, k_{2x}) - b(k_a, k_{1x}, k_{2x}-1)\Big] \\
&= (N\log N)^{-1/2}\sum_{n=1}^{N} a(n)a(n+k_a) \\
&\quad \times\Big[X(n+k_{1x})\Big(X(n+k_{2x}) - X(n+k_{2x}-1)\Big) - \Big(\gamma_X(k_x) - \gamma_X(k_x-1)\Big)\Big] \\
&= (N\log N)^{-1/2}\sum_{j=0}^{N^*-1}\sum_{r=1}^{M} a(r)a(r+k_a) \\
&\quad \times\Big[X(jM+r+k_{1x})\Big(X(jM+r+k_{2x}) - X(jM+r+k_{2x}-1)\Big) \\
&\quad - \Big(\gamma_X(k_x) - \gamma_X(k_x-1)\Big)\Big].
\end{aligned} \tag{37}$$

For any r $(=1,\dots M)$,

$$\begin{aligned}
&Var\Big[(N\log N)^{-1/2}\sum_{j=0}^{N^*-1} X(jM+r+k_{1x})\Big(X(jM+r+k_{2x}) - X(jM+r+k_{2x}-1)\Big)\Big] \\
&= (N\log N)^{-1}\sum_i\sum_j \gamma_X\Big((j-i)M\Big) \\
&\quad \times\Big[2\gamma_X\Big((j-i)M\Big) - \gamma_X\Big((j-i)M+1\Big) - \gamma_X\Big((j-i)M-1\Big)\Big] \\
&\quad + (N\log N)^{-1}\sum_i\sum_j\Big[\gamma_X\Big((j-i)M+k_x\Big) - \gamma_X\Big((j-i)M+k_x-1\Big)\Big] \\
&\quad \times\Big[\gamma_X\Big((j-i)M-k_x\Big) - \gamma_X\Big((j-i)M-k_x+1\Big)\Big] \\
&= (N\log N)^{-1}\sum_{|h|\le N^*-1}(N^*-|h|)\gamma_X(hM)\Big[2\gamma_X(hM) - \gamma_X(hM+1) - \gamma_X(hM-1)\Big] \\
&\quad + (N\log N)^{-1}\sum_{|h|\le N^*-1}(N^*-|h|)\Big[\gamma_X(hM+k_x) - \gamma_X(hM+k_x-1)\Big] \\
&\quad \times\Big[\gamma_X(hM-k_x) - \gamma_X(hM-k_x+1)\Big]
\end{aligned}$$

$$= o(1). \tag{38}$$

The last relation follows from that $\gamma_X(h) - \gamma_X(h+1) = o(h^{-1/2})$ as $h \to \infty$. Then the relation (36) follows from (37) and (38).

Similarly

$$(\log N)^{-1/2}\Big[b(k_a, k_{1x}, k_{2x}) - b(k_a, k_{1x}-1, k_{2x})\Big] = o_p(1). \tag{39}$$

Then applying (36) and (39) repeatedly,

$$(\log N)^{-1/2} b(k_a, k_{1x}, k_{2x}) = (\log N)^{-1/2} b(k_a, 0, 0) + o_p(1).$$

While by the same argument,

$$\begin{aligned}
(\log N)^{-1/2} b(k_a, 0, 0) &= (N \log N)^{-1/2} \sum_{j=0}^{N^*-1} \sum_{r=0}^{M} a(r) a(r+k_a)\Big(X(jM+r)^2 - \gamma_X(0)\Big) \\
&= (N \log N)^{-1/2} \sum_{j=0}^{N^*-1} \sum_{r=0}^{M} a(r) a(r+k_a)\Big(X(jM)^2 - \gamma_X(0)\Big) + o_p(1) \\
&= (N \log N)^{-1/2} \tilde{\gamma}_a(k_a) M \sum_{j=0}^{N^*-1} \Big(X(jM)^2 - \gamma_X(0)\Big) + o_p(1) \\
&= (N \log N)^{-1/2} \tilde{\gamma}_a(k_a) \sum_{n=1}^{N} \Big(X(n)^2 - \gamma_X(0)\Big) + o_p(1).
\end{aligned}$$

Hence the proof is completed. ∎

Now we have the limiting distribution of the estimators.

Theorem 4 *Assume the same conditions on $\{X(n)\}$ and $\{a(n)\}$ as those in Lemma 1.*

(i) *If* $d = \frac{1}{4}$, $\sqrt{\frac{N}{\log N}}\frac{\hat{\rho}_T(l) - \rho_X(l)}{1-\rho_X(l)}$, $\sqrt{\frac{N}{\log N}}\frac{\hat{\rho}_{SST}(l) - \rho_X(l)}{1-\rho_X(l)}$, *and* $\sqrt{\frac{N}{\log N}}\frac{\hat{\rho}_{PDR}(l) - \rho_X(l)}{1-\rho_X(l)}$ *have the common limiting distribution,* $N(0, \frac{4\alpha^2}{\gamma_X(0)^2})$ *as* $N \to \infty$.

(ii) *If* $\frac{1}{4} < d < \frac{1}{2}$, $N^{1-2d}\frac{\hat{\rho}_T(l) - \rho_X(l)}{1-\rho_X(l)}$, $N^{1-2d}\frac{\hat{\rho}_{SST}(l) - \rho_X(l)}{1-\rho_X(l)}$, *and* $N^{1-2d}\frac{\hat{\rho}_{PDR}(l) - \rho_X(l)}{1-\rho_X(l)}$ *have the* *common limiting distribution as $N \to \infty$, which is identical with the distribution of $R(1)$ where $R(1)$ is the value of Rosenblatt process $\{R(t); 0 \le t \le 1\}$ at $t = 1$.*

Proof. (i) Similar to (9), we have that

$$\sqrt{\frac{N}{\log N}}(\hat{\rho}_T(l) - \rho_X(l)) = \frac{b(l,0,l) - \rho_X(l)\Big(b(l,l,l) + b(l,0,0)\Big)/2}{\tilde{\gamma}_a(l)\gamma_X(0)(\log N)^{1/2}} + o_p(1).$$

Then it follows from Lemma 1 (i) that

$$\sqrt{\frac{N}{\log N}}\Big(\hat{\rho}_T(l) - \rho_X(l)\Big) = \Big(1 - \rho_X(l)\Big)(N \log N)^{-1/2} \sum_{n=1}^{N} \Big[X(n)^2 - \gamma_X(0)\Big]/\gamma_X(0) + o_p(1).$$

From Theorem 1' of Breuer and Major (1983), the limiting distribution of $(N \log N)^{-1/2} \sum_{n=1}^{N} \Big[X(n)^2 - \gamma_X(0)\Big]$ is $N(0, 4\alpha^2)$. Hence we have the result.

Similarly the limiting distribution of $\sqrt{\frac{N}{logN}}\Big(\hat{\rho}_{SST}(l) - \rho_X(l)\Big)$ is obtained. Finally using Lemma (i), we can show in the same way as (13) that

$$\sqrt{\frac{N}{\log N}}\Big(\hat{\rho}_{PDR}(l) - \rho_X(l)\Big) = \sqrt{\frac{N}{\log N}}\Big(\hat{\rho}_{SST}(l) - \rho_X(l)\Big) + o_p(1).$$

Then the proof is completed. ∎

(ii) By the same argument as that in the proof of (i), we see that the limiting distribution of the estimators is equal to that of $\gamma_X(0)^{-1} N^{-2d} \sum_{n-1}^{N} \Big(X(n)^2 - \gamma_X(0)\Big)$, which in turn is identical with the distribution of $R(1)$ [see Proposition of Rosenblatt (1979)].

Remark 2. (i) Interestingly the limiting distribution of the estimators is independent of $\{a(n)\}$. Hence the limiting distribution is still the same as that of estimators under complete sampling ($a(n) \equiv 1$). According to the proof of (i), the asymptotic distributions of the three estimators are determined only by the limiting distribution of the term

$$\Big(1 - \rho_X(l)\Big)(N \log N)^{-1/2} \sum_{n=1}^{N} \Big[X(n)^2 - \gamma_X(0)\Big]/\gamma_X(0). \tag{40}$$

Therefore the 3-dimensional asymptotic distribution of $(\sqrt{\frac{N}{\log N}}\frac{\hat{\rho}_T(l)-\rho_X(l)}{1-\rho_X(l)}, \sqrt{\frac{N}{\log N}}\frac{\hat{\rho}_{SST}(l)-\rho_X(l)}{1-\rho_X(l)}, \sqrt{\frac{N}{\log N}}\frac{\hat{\rho}_{PDR}(l)-\rho_X(l)}{1-\rho_X(l)})'$ is the degenerated distribution whose covariance matrix has rank 1 apparently. Moreover the asymptotic distribution of the estimator under complete sampling also depends solely on the term (40) as Hosking (1996) proved. A similar fact holds for Theorem 4 (ii) too.

We interpret this result as the correlation between distant observations is so strong that the presence of missing values between these observations cause no loss of information asymptotically. We shall consider the following example to make the reason more understandable.

Let X be a random variable. And suppose

$$X(n) \equiv X, \quad for\ any\ n.$$

We can regard $\{X(n)\}$ as a most extreme long-memory time series. Then $\rho_X(l) = 1$ for any l. If we assume that we observe only two samples at k and $k+l$, that is,

$$a(n) = \begin{cases} 1, & n = k, k+l, \\ 0, & otherwise, \end{cases}$$

and consider $\hat{\rho}_T(l)$ for example, then

$$\hat{\rho}_T(l) = \frac{\sum_{n=1}^{N-l} a(n)a(n+l)X(n)X(n+l)}{\sum_{n=1}^{N-l} a(n)a(n+l)X(n)^2} = \frac{X(k)X(k+l)}{X(k)^2} = \frac{X^2}{X^2} = 1.$$

On the other hand, under complete sampling ($a(n) \equiv 1$),

$$\hat{\rho}_T(l) = \frac{\sum_{n=1}^{N-l} X(n)X(n+l)}{\sum_{n=1}^{N-l} X(n)^2} = \frac{X^2}{X^2} = 1.$$

Therefore only two samples can supply all of the information about the autocorrelations of the time series. A similar situation emerges asymptotically for a long-memory model with $\frac{1}{4} \leq d < \frac{1}{2}$.

(ii) We assumed Gaussianity of $\{X(t)\}$ throughout this section. However Theorem 4 (ii) remains true even if $\{X(n)\}$ is a non-Gaussian process. Assume that $\{\varepsilon(t)\}$ consists of independently and identically distributed (*i.i.d.*) random variables with $E\Big[\varepsilon(t)^4\Big] < \infty$ and

$$\beta(j) \sim \gamma j^{d-1}, \quad \gamma > 0, \quad \frac{1}{4} < d < \frac{1}{2}$$

as $j \to \infty$. Hence $\gamma_X(h)$ satisfies (32). Then it is shown in the same way as in Theorem 3 of Hosking (1996) that Lemma 1 (ii) still holds since the fourth order cumulant terms are asymptotically negligible. Then it follows from Theorem 2 of Avram and Taqqu (1987) that the limiting distribution corresponding to the right-hand side of Lemma 1 (ii) is the same as that of a Gaussian case.

6 Mean Correction

Here we consider the case that the mean, $EX(n)$, is unknown. Then we can put $EX(n) = 0$ without loss of generality. And for simplicity we assume that $\{X(n)\}$ is a Gaussian process and

$$\gamma_X(h) \sim \alpha h^{2d-1}, \quad 0 < d < \frac{1}{2},$$

since the short-memory case can be treated in the same way. And let $\{a(n)\}$ satisfy Assumption 3 or 4 and $\gamma_a(h) \to 0$ as $h \to \infty$ under Assumption 3.

Define $\hat{X} = \bar{Y}/\bar{a}$. Then similar to Theorem 1 of Dunsmuir and Robinson (1981), we have

$$\lim_{N\to\infty} \hat{X} = E\Big[X(n)\Big], \qquad a.s.,$$

if $\mu_a \neq 0$. Hence we adopt $\hat{X}$ as an estimator of $EX(n)$.

First we shall show

$$\hat{X} = O_p(N^{d-1/2}) \tag{41}$$

and

$$\bar{X} - \hat{X} = o_p(N^{d-1/2}). \tag{42}$$

We express $\hat{X}$ by

$$\hat{X} = \frac{\sum_{n=1}^{N} a(n)X(n)/N}{\sum_{n=1}^{N} a(n)/N}. \tag{43}$$

First we impose Assumption 3 on $\{a(n)\}$. Then $\{Y(n)\}$ is a stationary process and

$$\gamma_Y(h) = \tilde{\gamma}_a(h)\gamma_X(h) = \gamma_a(h)\gamma_X(h) + \mu_a^2\gamma_X(h).$$

Hence

$$Var\Big[\sum_{n=1}^{N} Y(n)\Big] = Var\Big[\sum_{n=1}^{N} a(n)X(n)\Big] = O(N^{2d+1}). \tag{44}$$

The dominator of (43) converges to μ_a $a.s.$ as $N \to \infty$. Hence $\hat{X}$ satisfies the relation (41).

Next we impose Assumption 4 on $\{a(n)\}$. Then

$$\begin{aligned} Var\Big(\sum_{n=1}^{N} Y(n)\Big) &= Var\Big[\sum_{j=1}^{N^*}\sum_{r=1}^{M} Y(jM+r)\Big] \\ &\leq C\sum_{r=1}^{M} Var\Big[\sum_{j=1}^{N^*} Y(jM+r)\Big] \\ &\leq C\,M\,Var\Big[\sum_{j=1}^{N^*} X(jM+1)\Big] \\ &\leq CMN^* \sum_{|h|\leq N^*-1} |\gamma_X(Mh)| = O(N^{2d+1}). \end{aligned} \tag{45}$$

Hence the relation (41) is derived.

Next consider $\bar{X} - \hat{X}$. We express $\bar{X} - \hat{X}$ by

$$\bar{X} - \hat{X} = \frac{1}{N\mu_a}\sum_{n=1}^{N}\Big(\mu_a - a(n)\Big)X(n) + \bar{Y}\frac{\sum_{n=1}^{N}\Big(a(n) - \mu_a\Big)/N}{\mu_a \bar{a}}. \tag{46}$$

From (2), (44), and (45), the second term of (46) is $o_p(N^{d-1/2})$ under both Assumptions 3 and 4.

Next consider the first term of (46). Let $Z(n) = \Big(a(n) - \mu_a\Big)X(n)$. Then $\{Z(n)\}$ is a stationary process under Assumption 3 and $\gamma_Z(h) = \gamma_a(h)\gamma_X(h)$. Hence $Var\Big[\sum_{n=1}^{N}\Big(a(n) - \mu_a\Big)X(n)\Big]$ is $o(N^{2d+1})$ which implies that the first term is $o_p(N^{d-1/2})$.

Next we impose Assumption 4 on $\{a(n)\}$. Noting $\sum_{r=1}^{M}\Big[a(r) - \mu_a\Big] = 0$, we have

$$\sum_{n=1}^{N}\Big(a(n) - \mu_a\Big)X(n) = \sum_{j=1}^{N^*}\sum_{r=1}^{M}\Big(a(r) - \mu_a\Big)\Big(X(jM + r) - X(jM)\Big).$$

Hence

$$Var\Big[\sum_{n=1}^{N}\Big(a(n) - \mu_a\Big)X(n)\Big] \le C\sum_{r=1}^{M}|a(r) - \mu_a| Var\Big[\sum_{j=1}^{N^*}\Big(X(jM + r) - X(jM)\Big)\Big].$$

By the same argument as that in Lemma 1,

$$Var\Big[\sum_{j=1}^{N^*}\Big(X(jM + r) - X(jM)\Big)\Big] = o(N^{2d+1}), \quad (r = 1, 2, \ldots, M).$$

Hence the first term of (46) is $o_p(N^{d-1/2})$. Then the relation (42) is obtained.

Now we define $\bar{b}(k_a, k_{1x}, k_{2x})$ and $\hat{b}(k_a, k_{1x}, k_{2x})$ by replacing $X(n)$ in $b(k_a, k_{1x}, k_{2x})$ by $X(n) - \bar{X}$, $X(n) - \hat{X}$ respectively. Then though we omit the details, it follows from (41) and (42) that

$$\hat{b}(k_a, k_{1x}, k_{2x}) = b(k_a, k_{1x}, k_{2x}) + o_p(1), \quad 0 < d < \frac{1}{4}, \tag{47}$$

$$(\log N)^{-1/2}\hat{b}(k_a, k_{1x}, k_{2x}) = (\log N)^{-1/2}b(k_a, k_{1x}, k_{2x}) + o_p(1), \quad d = \frac{1}{4}, \tag{48}$$

$$N^{1/2-2d}\hat{b}(k_a, k_{1x}, k_{2x}) = N^{1/2-2d}\bar{b}(k_a, k_{1x}, k_{2x}) + o_p(1), \quad \frac{1}{4} < d < \frac{1}{2}, \tag{49}$$

and

$$N^{1/2-2d}\bar{b}(k_a, k_{1x}, k_{2x}) = N^{-2d}\tilde{\gamma}_a(k_a)\sum[(X(n)-\bar{X})^2 - \gamma_X(0)] + o_p(1), \quad \frac{1}{4} < d < \frac{1}{2}. \tag{50}$$

We have from (47) and (48) that Theorem 1, 2, 3, and 4 (i) still hold even if $EX(n)$ is unknown. While the relations (49) and (50) imply that the estimators have the same limiting distribution for $\frac{1}{4} < d < \frac{1}{2}$, which is also identical with that of estimators under complete sampling. However this distribution is different from that of $R(1)$ in Theorem 4 (ii) in the case that $EX(n)$ is known [see Yajima (1992) and Hosking (1996)]. The characteristic function of the distribution is given by Hosking (1996).

7 Computational Experiments

Here we shall reinforce the results in the previous sections by computational experiments. To generate random numbers, we use "ran2.c" in *Numerical recipes in C* [Press et al (1988)]. we suppose that $\{\varepsilon(n)\}$ are independently and identically distributed (*i.i.d.*) normal variables $N(0,1)$ in the first part of this section and then *i.i.d.* non-Gaussian random variables in the second part. The sample size is $N = 1000$ from Table 1 to 4 and Table 8, and $N = 1000$ and 5000 from Table 5 to 7. The number of replications is $K = 5000$ for all of the tables.

$\{a(n)\}$ is a sequence of Bernoulli trials and follows A-B sampling. We use

$$\hat{\rho}_{comp}(l) = \frac{\sum_{n=1}^{N-l} X(n)X(n+l)}{\sum_{n=1}^{N} X(n)^2},$$

as the estimator under complete sampling in this section and then consider the empirical ratios of the sample variances of the three estimators to those of the estimator under complete sampling, $\hat{\rho}_{comp}(l)$. For notational convenience, we denote them by $ER_{PDR}(l)$, $ER_{SST}(l)$, and $ER_T(l)$. We also denote the three theoretical asymptotic ratios of the variances by $AR_{PDR}(l)$, $AR_{SST}(l)$, and $AR_T(l)$ and give them in Table 1 and 3.

For example if we consider an AR(1) process, $X(n) = \phi X(n-1) + \varepsilon(n)$ and a sequence of Bernoulli trials $\{a(n)\}$, the asymptotic variances: $AV_{PDR}(1)$, $AV_{SST}(1)$, and $AV_T(1)$, are given by (29), (30), and (31). On the other hand if we denote the asymptotic variance under complete sampling $AV_{comp}(1)$, we have $AV_{comp}(1) = 1 - \phi^2$. Hence the asymptotic ratio $AR_{SST}(1)$ is $AV_{SST}(1)/AV_{comp}(1) = 1/p^2$ for $\hat{\rho}_{SST}(1)$ and any ϕ.

Table 1 show the result of the case that $\{X(n)\}$ is an AR(1) process from $\phi = -0.9$ to 0.9 and $\{a(n)\}$ is a sequence of Bernoulli trials with $p = 0.9$. On the other hand, the autocorrelation $\rho_X(1)$ of the AR(1) process is identical with the parameter ϕ. And Reinsel and Wincek (1987) (Theorem 2) and Qian (1988) gave theoretical asymptotic variances for the Gaussian MLE for ϕ with missing observations. Thus we also listed ratios of the asymptotic variances of $\hat{\phi}_{MLE}$ in reference to the other ones. We can find from Table 1 that the empirical values and the theoretical values are nearly equal to each other, and that $ER_T(1) \leq ER_{SST}(1) \leq ER_{PDR}(1)$ holds as Corollary 1 showed. Additionally, $AR_T(1)$ and $ER_T(1)$ are relatively small and close to those of $\hat{\phi}_{MLE}$. $\hat{\rho}_T(1)$ is therefore as competitive as $\hat{\phi}_{MLE}$.

Table 2 shows the result for A-B sampling. $ER_T(1)$ is always the smallest one. And $ER_{PDR}(1)$ is always smaller than $ER_T(1)$ with A-B=3-1. However for A-B=4-1,

$ER_{PDR}(1)$ is smaller than $ER_{SST}(1)$ only when $|\phi|$ is large. Conversely, $ER_{SST}(1)$ is smaller than $ER_{PDR}(1)$ when $|\phi|$ is small as Remark 1 (ii) assured. Further if A is larger with A=5, $ER_{SST}(1)$ is smaller than $ER_{PDR}(1)$ except for the case $|\phi| = 0.9$.

Next we consider a MA(1) process, $X(n) = \theta\varepsilon(n-1) + \varepsilon(n)$. Table 3 shows the result of the case of Bernoulli trials. We calculate the asymptotic variances by Theorem 2 in the following way. Substituting (17), (18), and (19) into (15) and (16), we have

$$AV_{SST}(1) = \frac{1}{p^2}[1 - (1+2p)\rho_X(1)^2 + 4\rho_X(1)^4],$$

and

$$AV_T(1) = \frac{1}{p^2}[1 - (2+p)\rho_X(1)^2 + (1+p)^2\rho_X(1)^4], \tag{51}$$

where $\rho_X(1)$ is $\frac{\theta}{1+\theta^2}$ since $\{X(n)\}$ follows the MA(1) process. And we obtain

$$AV_{PDR}(1) = \frac{1}{p^2}[1 + (1-4p)\rho_X(1)^2 + 4p^2\rho_X(1)^4],$$

by (24) and (51) and

$$AV_{comp}(1) = 1 - 3\rho_X(1)^2 + 4\rho_X(1)^4,$$

by Bartlett's formula [see e.g., Brockwell and Davis (1991)].

We listed the theoretical asymptotic ratios and the empirical ratios in Table 3 as in Table 1. We can find that the empirical values and the theoretical values are almost similar and $AR_T(1)$ and $ER_T(1)$ are the smallest for each θ. On the other hand, Table 4 contains the ratios of the empirical variances under A-B sampling to those under complete sampling as in Table 2. $ER_{SST}(1)$ are always larger than $ER_{PDR}(1)$ with A-B=3-1 and $ER_{SST}(1)$ are always smaller than $ER_{PDR}(1)$ with A-B=4-1 and A-B=5-1 in contrast to the result in Table 2.

Table 1: Ratios of variances to complete sampling, AR(1), Bernoulli trials p=0.90

ϕ	$\hat{\rho}_{PDR}$ empir	$\hat{\rho}_{PDR}$ theor	$\hat{\rho}_{SST}$ empir	$\hat{\rho}_{SST}$ theor	$\hat{\rho}_T$ empir	$\hat{\rho}_T$ theor	$\hat{\phi}_{MLE}$ theor
-0.90	2.1955	2.2872	1.2102	1.2346	1.0574	1.0617	1.0131
-0.70	1.4646	1.4718	1.2567	1.2346	1.1585	1.1474	1.0529
-0.50	1.3087	1.3169	1.2386	1.2346	1.1966	1.1968	1.1111
-0.30	1.2545	1.2590	1.2282	1.2346	1.2150	1.2226	1.1780
-0.10	1.2173	1.2371	1.2145	1.2346	1.2127	1.2333	1.2273
0.00	1.2296	1.2346	1.2304	1.2346	1.2299	1.2346	1.2346
0.10	1.2256	1.2371	1.2265	1.2346	1.2238	1.2333	1.2273
0.30	1.2705	1.2590	1.2466	1.2346	1.2323	1.2226	1.1780
0.50	1.3166	1.3169	1.2450	1.2346	1.2035	1.1968	1.1111
0.70	1.4713	1.4718	1.2330	1.2346	1.1510	1.1474	1.0529
0.90	2.2860	2.2872	1.2244	1.2346	1.0616	1.0617	1.0131

sample size 1000, replications 5000.

Table 2: Ratios of variances to complete sampling, AR(1), A-B sampling

ϕ	A-B=3-1			A-B=4-1			A-B=5-1		
	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$
-0.90	1.3726	1.9827	1.1026	1.4203	1.6898	1.0836	1.4508	1.4562	1.0541
-0.70	1.6307	1.9301	1.3783	1.5816	1.6499	1.2908	1.5345	1.4872	1.2350
-0.50	1.8996	2.0019	1.7096	1.7000	1.6586	1.4978	1.5598	1.4943	1.3883
-0.30	1.9496	1.9671	1.8663	1.6562	1.6450	1.5880	1.5010	1.4754	1.4413
-0.10	1.9795	1.9817	1.9689	1.6663	1.6689	1.6561	1.4926	1.4900	1.4863
0.00	2.0410	2.0508	2.0403	1.6929	1.6927	1.6936	1.5316	1.5323	1.5324
0.10	2.0483	2.0549	2.0354	1.7159	1.7069	1.7036	1.5039	1.5030	1.4982
0.30	1.9657	1.9812	1.8970	1.6908	1.6556	1.6153	1.5175	1.5069	1.4592
0.50	1.9069	2.0277	1.7091	1.7798	1.7478	1.5637	1.6009	1.5419	1.4105
0.70	1.6349	1.9081	1.3708	1.6289	1.6640	1.3005	1.5251	1.4617	1.2210
0.90	1.3863	1.9370	1.1106	1.3829	1.7001	1.0734	1.4059	1.4670	1.0514

sample size 1000, replications 5000.

Table 3: Ratios of variances to complete sampling, MA(1), Bernoulli trials p=0.90

θ	$\hat{\rho}_{PDR}$		$\hat{\rho}_{SST}$		$\hat{\rho}_{T}$	
	empir	theor	empir	theor	empir	theor
-0.90	1.3923	1.3633	1.3232	1.2959	1.2579	1.2367
-0.70	1.3514	1.3534	1.3116	1.2917	1.2510	1.2417
-0.50	1.3227	1.3229	1.2911	1.2777	1.2536	1.2465
-0.30	1.2709	1.2748	1.2474	1.2545	1.2352	1.2428
-0.10	1.2200	1.2395	1.2172	1.2370	1.2155	1.2358
0.10	1.2277	1.2395	1.2286	1.2370	1.2259	1.2358
0.30	1.2871	1.2748	1.2673	1.2545	1.2535	1.2428
0.50	1.3234	1.3229	1.2895	1.2777	1.2511	1.2465
0.70	1.3672	1.3534	1.2946	1.2917	1.2505	1.2417
0.90	1.3740	1.3633	1.3095	1.2959	1.2483	1.2367

sample size 1000, replications 5000.

Table 4: Ratios of variances to complete sampling, MA(1), A-B sampling

θ	A-B=3-1			A-B=4-1			A-B=5-1		
	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_{T}$	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_{T}$	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_{T}$
-0.90	2.4605	2.5523	2.0129	2.0436	1.9563	1.7110	1.7596	1.7174	1.5049
-0.70	2.2995	2.3675	1.9718	1.9132	1.8810	1.6549	1.6897	1.6143	1.4802
-0.50	2.3466	2.3822	2.1149	1.9060	1.8594	1.7227	1.6784	1.6266	1.5408
-0.30	2.1173	2.1206	2.0249	1.7385	1.7296	1.6729	1.5593	1.5363	1.5033
-0.10	1.9996	2.0016	1.9889	1.6772	1.6797	1.6668	1.5001	1.4976	1.4940
0.10	2.0683	2.0751	2.0554	1.7274	1.7183	1.7151	1.5116	1.5108	1.5059
0.30	2.1301	2.1338	2.0546	1.7792	1.7439	1.7052	1.5742	1.5673	1.5199
0.50	2.3515	2.3933	2.1147	2.0039	1.9390	1.8011	1.7324	1.6753	1.5755
0.70	2.3256	2.3942	1.9748	1.9799	1.9326	1.6815	1.7362	1.6691	1.5191
0.90	2.4088	2.4772	1.9902	1.9587	1.8900	1.6573	1.7500	1.7189	1.5041

sample size 1000, replications 5000.

Table 5: Fractional ARIMA(0,d,0), from $d = 0.1$ to 0.49

N	d	Bernoulli trials with $p = 0.90$			A-B sampling with A-B=9-1		
		$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$
1000							
	0.10	1.21463	1.21140	1.20862	1.22314	1.21811	1.21924
	0.20	1.17815	1.17117	1.16459	1.13955	1.13372	1.12391
	0.25	1.13042	1.11982	1.11247	1.11397	1.09757	1.09371
	0.30	1.08216	1.06593	1.05692	1.06435	1.05582	1.04992
	0.40	1.02905	1.02243	1.01550	1.02160	1.01904	1.01110
	0.49	1.01744	1.00769	1.00694	1.00627	1.00403	1.00259
5000							
	0.10	1.21830	1.21470	1.21403	1.24943	1.24360	1.24388
	0.20	1.16696	1.15429	1.14856	1.15289	1.14342	1.14084
	0.25	1.11716	1.09653	1.09529	1.07462	1.06828	1.06451
	0.30	1.05818	1.04880	1.04159	1.05192	1.04638	1.04074
	0.40	1.01341	1.00784	1.00537	1.00800	1.00496	1.00391
	0.49	1.00814	1.00259	1.00215	1.00048	1.00277	1.00012

replications 5000.

Table 6: Mean Unknown, Fractional ARIMA(0,d,0), from $d = 0.1$ to 0.49

N	d	Bernoulli trials with $p = 0.90$			A-B sampling with A-B=9-1		
		$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$
1000							
	0.10	1.21953	1.21221	1.21144	1.22640	1.22755	1.22267
	0.20	1.20482	1.19041	1.18801	1.15375	1.14267	1.13846
	0.25	1.16935	1.15422	1.14755	1.14750	1.13695	1.12516
	0.30	1.14322	1.11043	1.10391	1.12362	1.11054	1.09668
	0.40	1.09654	1.06089	1.05189	1.05556	1.03837	1.02924
	0.49	1.09394	1.05055	1.03628	1.05218	1.04061	1.02060
5000							
	0.10	1.21844	1.21564	1.21354	1.25129	1.24967	1.24596
	0.20	1.18341	1.16905	1.16437	1.17327	1.17120	1.16026
	0.25	1.15807	1.14207	1.13044	1.10025	1.09954	1.08742
	0.30	1.10731	1.07947	1.07709	1.10218	1.09227	1.08284
	0.40	1.05217	1.02995	1.02255	1.03862	1.03428	1.02195
	0.49	1.04564	1.01852	1.01152	1.01066	1.01393	1.00334

replications 5000.

Table 7: Variances, Fractional ARIMA(0,d,0), from d = 0.1 to 0.49 (upper line, Mean known; lower line, Mean unknown)

N	d	complete sampling	Bernoulli trials with $p = 0.90$			A-B sampling with A-B=9-1		
			$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_{T}$	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_{T}$
1000								
	0.10	1.13180	1.37471	1.37107	1.36792	1.38435	1.37866	1.37993
		1.11171	1.35576	1.34763	1.34677	1.36340	1.36468	1.35926
	0.20	1.56162	1.83983	1.82893	1.81865	1.77955	1.77044	1.75513
		1.38224	1.66534	1.64543	1.64211	1.59475	1.57944	1.57362
	0.25	0.29615	0.33478	0.33164	0.32946	0.32990	0.32505	0.32390
		0.22492	0.26301	0.25961	0.25811	0.25810	0.25573	0.25308
	0.30	0.82339	0.89103	0.87767	0.87025	0.87637	0.86935	0.86449
		0.48391	0.55322	0.53735	0.53419	0.54373	0.53740	0.53070
	0.40	0.11966	0.12314	0.12235	0.12152	0.12225	0.12194	0.12099
		0.04222	0.04629	0.04479	0.04441	0.04456	0.04384	0.04345
	0.49	0.01176	0.01197	0.01186	0.01185	0.01184	0.01181	0.01180
		0.00357	0.00391	0.00375	0.00370	0.00376	0.00372	0.00365
5000								
	0.10	1.12800	1.37425	1.37019	1.36943	1.40936	1.40278	1.40309
		1.12330	1.36867	1.36553	1.36317	1.40558	1.40375	1.39959
	0.20	1.55142	1.81045	1.79079	1.78190	1.78862	1.77393	1.76993
		1.41238	1.67142	1.65113	1.64453	1.65710	1.65417	1.63873
	0.25	0.28954	0.32347	0.31749	0.31714	0.31115	0.30932	0.30822
		0.21649	0.25071	0.24725	0.24473	0.23819	0.23804	0.23541
	0.30	0.80881	0.85586	0.84828	0.84245	0.85080	0.84632	0.84176
		0.44952	0.49776	0.48524	0.48417	0.49545	0.49099	0.48675
	0.40	0.12101	0.12264	0.12196	0.12166	0.12198	0.12161	0.12149
		0.02838	0.02986	0.02923	0.02902	0.02947	0.02935	0.02900
	0.49	0.00862	0.00869	0.00864	0.00864	0.00862	0.00864	0.00862
		0.00172	0.00179	0.00175	0.00174	0.00173	0.00174	0.00172

replications 5000.

Table 8: Variances, Fractional ARIMA(0,d,0), from $d = 0.1$ to 0.49, Uniform distribution and t-distribution (upper line, Mean known; lower line, Mean unknown)

distribution	d	complete sampling	Bernoulli trials with $p = 0.90$			A-B sampling with A-B=9-1		
			$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$	$\hat{\rho}_{PDR}$	$\hat{\rho}_{SST}$	$\hat{\rho}_T$
uniform								
	0.10	1.14757	1.38978	1.38909	1.38664	1.39278	1.38981	1.38813
		1.12806	1.37449	1.36852	1.36922	1.37585	1.37234	1.37185
	0.20	1.52957	1.79709	1.78018	1.77304	1.82769	1.81468	1.80384
		1.37593	1.65272	1.62700	1.62745	1.67898	1.65533	1.65728
	0.25	0.29978	0.33663	0.33259	0.33040	0.33950	0.33713	0.33429
		0.23093	0.26761	0.26167	0.26165	0.26910	0.26429	0.26470
	0.30	0.76048	0.81956	0.80975	0.80260	0.81151	0.80945	0.79992
		0.46179	0.52360	0.50739	0.50919	0.50813	0.49960	0.49962
	0.40	0.12345	0.12761	0.12652	0.12540	0.12635	0.12624	0.12522
		0.04091	0.04469	0.04292	0.04289	0.04417	0.04315	0.04299
	0.49	0.01071	0.01091	0.01079	0.01076	0.01081	0.01076	0.01075
		0.00358	0.00396	0.00374	0.00371	0.00373	0.00369	0.00365
t-dist								
	0.10	1.07246	1.44369	1.32376	1.39979	1.46156	1.39282	1.47376
		1.06056	1.43376	1.72229	1.39145	1.44275	1.92157	1.45802
	0.20	1.50438	1.79771	1.75058	1.75372	1.79396	1.75177	1.75469
		1.36032	1.64600	1.65579	1.60499	1.64832	1.67203	1.61284
	0.25	0.30006	0.34754	0.33639	0.33795	0.34710	0.34365	0.34303
		0.22242	0.27127	0.27395	0.26146	0.26850	0.27740	0.26422
	0.30	0.80312	0.88733	0.86918	0.86724	0.85696	0.84914	0.84718
		0.47686	0.56595	0.56703	0.54498	0.54158	0.56298	0.53090
	0.40	0.12452	0.14049	0.13180	0.13733	0.14222	0.13253	0.13993
		0.04049	0.06000	0.08858	0.05545	0.06048	0.10552	0.05866
	0.49	0.01207	0.01278	0.01230	0.01252	0.01268	0.01226	0.01245
		0.00358	0.00492	0.00664	0.00456	0.00467	0.00716	0.00463

sample size 1000, replications 5000.

Next we consider a fractional ARIMA(p,d,q) process. $\{X(n)\}$ is expressed as

$$\phi(B)(1-B)^d X(n) = \theta(B)\varepsilon(n),$$

where B is the backward shift operator and $\{\varepsilon(n)\}$ is a white noise process with mean 0 and variance σ_ε^2. d is a real number with $0 < d < \frac{1}{2}$. Let $\phi(z) = 1 - \sum_{j=1}^{p} \phi_j z^j$ and $\theta(z) = 1 - \sum_{j=1}^{q} \theta_j z^j$, where p and q are positive integers. Suppose $\phi(z)$ and $\theta(z)$ have no zeros on or inside the unit circle and no zeros in common. It has the spectral density

$$f(\lambda; d, \phi, \theta, \sigma_\varepsilon^2) = \frac{\sigma_\varepsilon^2}{2\pi}|1 - e^{i\lambda}|^{-2d}\frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2}.$$

And the spectral density is not square integrable when $\frac{1}{4} \le d < \frac{1}{2}$. In the following experiments, we generate the fractional ARIMA(0,d,0) process by *i.i.d.* normal variables $N(0,1)$ and the Durbin-Levinson algorithm [see e.g., Brockwell and Davis (1991)].

Table 5 gives the empirical ratios of $\hat{\rho}_{PDR}(1)$, $\hat{\rho}_{SST}(1)$, and $\hat{\rho}_T(1)$ to those under complete sampling for d = 0.1, 0.2, 0.25, 0.3, 0.4, and 0.49 respectively. We can see from Table 5 that $\hat{\rho}_T(1)$ is relatively efficient and all the values become nearly equal, as d increases. On the whole, the ratios also converge to 1, as N increases, for $d \ge 0.25$. These facts are consistent with Theorem 4.

Next if $EX(n)$ is unknown, we replace $Y(n) = a(n)X(n)$ by $Y(n) - a(n)\hat{X}(n) = a(n)(X(n) - \hat{X}(n))$ and calculate the empirical variances as in Table 5, where $X(n)$ follow fractional ARIMA(0,d,0) processes and $EX(n) = 0$. The values in Table 5 and the corresponding values in Table 6 are almost similar for $d \le 0.25$. On the other hand, the differences between the values in Table 5 and those in Table 6 are large as d increase. But the ratios mostly converge to 1, as N increases, for $d \ge 0.25$ as well as in Table 5. These facts are consistent with the things described in Section 6.

Next we listed the asymptotic variances for a fractional ARIMA(0,d,0) in Table 7 to compare the case of mean known with the case of mean unknown in further details. Here we calculate the empirical variances of $\sqrt{N}(\hat{\rho}_{PDR}(1) - \rho_{PDR}(1))$, $\sqrt{N}(\hat{\rho}_{SST}(1) - \bar{\rho}_{SST}(1))$, and $\sqrt{N}(\hat{\rho}_T(1) - \bar{\rho}_T(1))$ for $d = 0.10, 0.20$, $\sqrt{N/\log N}(\hat{\rho}_{PDR}(1) - \rho_{PDR}(1))$, $\sqrt{N/\log N}(\hat{\rho}_{SST}(1) - \bar{\rho}_{SST}(1))$, and $\sqrt{N/\log N}(\hat{\rho}_T(1) - \bar{\rho}_T(1))$ for $d = 0.25$, $N^{1-2d}(\hat{\rho}_{PDR}(1) - \bar{\rho}_{PDR}(1))$, $N^{1-2d}(\hat{\rho}_{SST}(1) - \bar{\rho}_{SST}(1))$, and $N^{1-2d}(\hat{\rho}_T(1) - \rho_T(1))$ for $d = 0.30, 0.40, 0.49$, where $\bar{\rho}_{PDR}(1)$, $\bar{\rho}_{SST}(1)$, and $\bar{\rho}_T(1)$ are sample means. For each d, the case of mean known is on upper line, the case of mean unknown on lower line in Table 7. When d is small, the difference between the case of mean known and the case of mean unknown is small. Especially, the difference disappears as N is large. It supports the facts indicated by (47) and (48) in Section 6. On the other hand, when d is large — the spectral density is not square integrable, the difference between the case of mean known and the case of mean unknown is large. The reason is that the asymptotic distributions are different as the relations (49) and (50) showed. But the empirical variances under complete sampling and the other ones are nearly equal. Especially, whether missing observation exist or not, all the values are become the same ones as N increase. The experiments support the fact that the existence of missing observations have no influence on the asymptotic variances when the spectral density is not square integrable and N is sufficiently large. It has also already been shown in Remark 2 (i) and at the last paragraph in Section 6.

Finally we assume that $\varepsilon(n)$ are non-Gaussian random variables. First $\{\varepsilon(n)\}$ is a sequence of *i.i.d.* random variables which follow a uniform distribution; the probability density function is assumed to be

$$p(x) = \frac{1}{2\sqrt{3}} I(-\sqrt{3}, \sqrt{3}),$$

where $I(\cdot)$ is an indicator function. The mean is 0, the variance is 1, and the fourth-order cumulant κ is -1.2. Secondly we also assume that $\{\varepsilon(n)\}$ is a sequence of *i.i.d.* random variables which follow a t-distribution. Let t_ν be a random variable of t-distribution with freedom ν and $\varepsilon(n) \sqrt{\frac{\nu-2}{\nu}} t_\nu$. The 4th-order moment of t_ν exist if $\nu > 4$ holds. We therefore suppose $\nu = 5$. The mean of $\varepsilon(n)$ is 0, the variance is 1, and κ is $\frac{6}{\nu-4} = 6$. In Table 8 as well as in Table 7, we give empirical variances of the case of non-Gaussian. As d is large, the values become similar to those in Table 7. In other words it is assured that non-Gaussianity does not have a serious influence on the limiting distributions.

Acknowledgments

The authors thank Professor Peter M. Robinson for reading the draft and giving them helpful comments.

References

[1] Anderson, T. W. and Walker, A. M. (1964) On the asymptotic distribution of the autocorrelations of a sample from a liner stochastic process. *Annals of Mathematical Statistics*, **35**, 1296-1303.

[2] Avram, F. and Taqqu, M. S. (1987) Noncentral limit theorems and appell polynomials. *Annals of Probability*, **15**, 767-775.

[3] Bartlett, M. S. (1946) On the theoretical specification and sampling properties of autocorrelated time-series. *Journal of Royal Statistical Society Supplement*, **8**, 27-41, 85-97.

[4] Blockwell, P. J. and Davis, R. A. (1991) *Time Series:Theory and Methods*, 2nd ed., Springer, New York.

[5] Breuer, P. and Major, P. (1983) Central limit theorems for non-linear functionals of gaussian fields. *Journal of Multivariate Analysis*, **13**, 425-441.

[6] Cavazos-Cadena, R. (1994) The asymptotic distribution of sample autocorrelations for a class of linear filters. *Journal of Multivariate Analysis*, **48**, 249-274.

[7] Dunsmuir, W. and Robinson, P. (1981) Asymptotic theory for time series containing missing and amplitude modulated observations. *Sankhyā*, **Ser. A, 43**, 260-281.

[8] Granger, C. W. J. and Joyeux, R. (1980) An introduction to a long memory time series models and fractional differencing. *Journal of Time Series Analysis*, **1**, 15-29.

[9] He, S. (1996) A note on the asymptotic normality of sample autocorrelations for a linear stationary sequence. *Journal of Multivariate Analysis*, **58**, 182-188.

[10] Hosking, J. R. M. (1981) Fractional differencing. *Biometrica*, **68**, 165-176.

[11] Hosking, J. R. M. (1996) Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *Journal of Econometrics*, **73**, 261-284.

[12] Mandelbrot, B. B. and Van Ness, J. W. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review*, **10**, 422-437.

[13] Parzen, E. (1963) On spectral analysis with missing observations and amplitude modulation. *Sankhyā*, **Ser.A. 25**, 383-392.

[14] Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1988) *Numerical Recipes in C.* Cambridge Univ, New York.

[15] Qian, W. (1988) Gaussian estimation of first order time series models with bernoulli observations. *Stochastic Processes and their Applications*, **27**, 85-96.

[16] Reinsel, G. C. and Wincek, M.A. (1987) Asymptotic distribution of parameter estimators for non consecutively observed time series. *Biometrika*, **74**, 115-124.

[17] Robinson, P. M. (1985) Testing for serial correlation in regression with missing observations. *Journal of the Royal Statistical Society*, **Soc. B, 47**, 429-437.

[18] Rosenblatt, M. (1979) Some limit theorems for partial sums of quadratic forms in stationary Gaussian variables. *Zeitschrift für Wahrsehsceinlichkeitstheorie und Verwandte Gebiete*, **49**, 125-132.

[19] Shin, D.W. and Sarkar, S. (1995) Estimation for stationary AR(1) models with nonconsecutively observed samples. *Sankhyā*, **Ser. A, 57**, 287-298.

[20] Takeuchi, K. (1995) A comment on "Recent Development of Economic Data Analysis"at the 63th Annual Meeting of Japan Statistical Society.