On a Class of Estimation and Test for Long Memory

Fu, Hui

Jinan University

6 December 2012

On a Class of Estimation and Test for Long Memory

Hui Fu*
Jinan University

Abstract: This paper advances a new analysis technology path of estimation and test for long memory time series. I propose the definitions of time scale series, strong variance scale exponent and weak variance scale exponent, and prove the strict mathematical equations that strong and weak variance scale exponent can accurately identify the time series of white noise, short memory and long memory, especially derive the equation relationships between weak variance scale exponent and long memory parameters. I also construct two statistics which *SLmemory* statistic tests for long memory properties. The paper further displays Monte Carlo performance for MSE of weak variance scale exponent estimator and the empirical size and power of *SLmemory* statistic, giving practical recommendations of finite-sample, and also provides brief empirical examples of logarithmic return rate series data for Sino-US stock markets.

Keywords: Long Memory, Weak Variance Scale Exponent, *SLmemory* Statistic, Time Scale Series.
JEL Classification: C22, C13, C12.

1 Introduction and Setup

The first research of estimation for long memory dated back to Hurst (1951) who proposed classical rescaled range (R/S) method to analyze long-range dependence and subsequently refined by excellent scholars (Mandelbrot and Wallis, 1969; Mandelbrot, 1972, 1975; Mandelbrot and Taqqu, 1979; Peters, 1999, 2002). Following, fractional Gaussian noises process (Mandelbrot and Van Ness, 1968) was the first long memory model, with the long memory parameter H (Hurst exponent, also called self-similarity parameter) satisfying $0 < H < 1$. And Lo (1991), Kwiatkowski et al. (1992) and Giraitis et al. (2003) further promote and develop the study of tests for long memory properties following the R/S-type method.

Fractional differencing noises process (Granger and Joyeux, 1980; Hosking, 1981) was the second long memory model and formed the second technology path of estimation and test for long memory properties, with the long memory parameter (fractional differencing parameter d) satisfying $-0.5 < d < 0.5$. Corresponding work include Geweke and Porter-Hudak (1983), Haslett and Raftery (1989), Beran (1994), Robinson (1995, 2005), among many others.

Lo (1991), Baillie (1996), Giraitis et al. (2003), Robinson (2003), Palma (2007) and Boutahar et al. (2007) provided relevant literature review in detail.

It is no doubt that the model parameter satisfying $0.5 < H < 1$, or $0 < d < 0.5$ is thought as long memory, long-range dependence, or persistence, and as anti-persistence for the situation of $0 < H < 0.5$ or $-0.5 < d < 0$. But there is argument that whether anti-persistence is long memory (Hosking, 1981;

*Hui Fu is from School of Economics, Jinan University, Guangzhou, China. E-mail: fuhuicn@gmail.com or hui_fu@hotmail.com

Lo, 1991; Palma, 2007). The present paper suggest that the situations of $0.5 < H < 1$, $0 < d < 0.5$ regard as persistent long memory, and as anti-persistent long memory for $0 < H < 0.5$ or $-0.5 < d < 0$. Thus, persistent long memory means long-run positive correlation, and anti-persistent long memory implies long-run negative correlation.

The main theoretical contribution of this paper is to build a new systematic technology path for theory analysis of stationary time series, especially for the estimation and test of long memory properties. All of those the definitions of time scale series, strong variance scale exponent and weak variance scale exponent, the equation relationships between the short, long memory time series and variance scale exponents, and the two new statistics for the hypothesis tests of white noise, short memory and long memory time series, jointly form the basic components of the new technology path.

The second contribution of this paper is to show that the asymptotic properties for estimation and test can distinguish anti-persistent long memory from persistent long memory. Further it is the main reason that I advance the two types of (persistence and anti-persistence) long memory in the present paper.

The final contribution is to analyze time series properties based on the both time and frequency domain perspectives, and corresponding long memory models are confined to time and frequency domain analysis perspectives in the paper, which differs the situations that the two traditional technology paths.

The plan of the paper is as follow. In sections 2 and 3, I propose the definitions of time scale series, strong variance scale exponent and weak variance scale exponent, and prove the equations between variance scale exponents and the situations that white noise, short memory and long memory time series. I also construct $Wnoise$ and $SLmemory$ statistics to test for white noise, short memory and long memory time series. Section 4 investigates by Monte Carlo simulations the finite-sample performance of weak variance scale exponent estimator and $SLmemory$ statistic, giving practical recommendations for the choice of maximum time scale n. Section 5 concludes. The derivations are given in the Appendix.

I now detail the setting for the paper. Let $\{x(t), t = 1, 2, \cdots, N, N \to \infty\}$ be a stationary time series with unknown mean μ and lag-i autocovariance γ_i. Let the spectral density of $x(t)$ be denoted by $f(\lambda)$ and defined over $|\lambda| \leq \pi$. d denotes long memory parameter of fractional differencing noises process (Granger and Joyeux, 1980; Hosking, 1981). H denotes long memory parameter of fractional Gaussian noises process (Mandelbrot and Van Ness, 1968). N denotes sample sizes.

2 Definitions and Properties of Strong and Weak Variance Scale Exponent

2.1 Definitions

Definition 2.1.1 Time Scale Series

For the time series $\{x(t)\}$, the new time series $\{x_2^1 = x(2t-1) + x(2t), t = 1, 2, \cdots, M, M = \lfloor N/2 \rfloor\}$ and $\{x_2^2 = x(2t) + x(2t+1), t = 1, 2, \cdots, M, M = [(N-1)/2]\}$ can be called time scale-2 series; And, the new time series $\{x_n^{i+1}, t = 1, 2, \cdots, M, M = \lfloor (N-i)/n \rfloor\}$ are called time scale-n series, where $x_n^{i+1} = x(nt+i) + x(nt-1+i) + \cdots + x(nt-(n-1)+i)$, $i = 0, 1, 2, \cdots, (n-1)$,$n/N \to \infty$, and $\lfloor \cdot \rfloor$ means round toward zero, resulting integers.

I now list four propositions arising from definition 2.1.1:

Proposition 2.1.1 Denoted $D[x_n^{i+1}](i = 0, 1, 2, \cdots, n-1)$ are the n variances of the time scale-n series respectively, then

$$D[x_n^1(t)] = D[x_n^2(t)] = \cdots = D[x_n^n(t)]. \tag{2.1.1}$$

In the case, I denote the n variances of the time scale-n series as $D[x_n(t)]$.

Proposition 2.1.2 Variance $D[x_n(t)]$ of time scale-n series can be expressed in terms of

$$D[x_n(t)] = nD[x(t)] + 2\sum_{i=1}^{n}(n-i)\gamma_i, n = 1, 2, \cdots . \tag{2.1.2}$$

Proposition 2.1.3 Lag-n autocovariance γ_n of the time series $\{x(t)\}$ can be expressed in terms of

$$\gamma_n = \frac{1}{2}\{D[x_{n+1}(t)] - 2D[x_n(t)] + D[x_{n-1}(t)]\}, \tag{2.1.3}$$

where $D[x_0(t)] = 0, D[x_1(t)] = D[x(t)]$, and $n = 1, 2, \cdots$.

Proposition 2.1.4 The sum of lag-i autocovariance γ_i can be expressed in terms of

$$\sum_{i=-n}^{n}\gamma_i = n\{D[x_{n+1}(t)] - D[x_n(t)]\}, n = 1, 2, \cdots . \tag{2.1.4}$$

Definition 2.1.2 Strong Variance Scale Exponent

For the time series $\{x(t)\}$, if the variance of time scale-n series satisfies

$$D[x_n(t)] = n^{2F_n}D[x(t)], n = 1, 2, \cdots , \tag{2.1.5}$$

and if F_n equal to a constant $\dot{F}$. Thus $\dot{F}$ can be called Strong Variance Scale Exponent.

Definition 2.1.3 Weak Variance Scale Exponent

For the time series $\{x(t)\}$, if the variance of time scale-n series satisfies

$$D[x_n(t)] = fn^{2F_n}D[x(t)], \tag{2.1.6}$$

and F_n get the convergence to a constant F, as $n \to \infty$. Thus F and f are called Weak Variance Scale Exponent and adjusted proportion coefficient respectively.

2.2 Time Domain Analysis

Base on four propositions arising from definition 2.1.1, I obtain following results.

Theorem 2.2.1 If the time series $\{x(t)\}$ is white noise series, then

$$D[x_n(t)]/D[x(t)] = n, n = 1, 2, \cdots . \tag{2.2.1}$$

I now list a implication arising from Theorem 2.1.1:

Proposition 2.2.1 If the time series $\{x(t)\}$ is white noise series, then strong variance scale exponent $\dot{F}$, weak variance scale exponent F and adjusted proportion coefficient f satisfy:

$$\dot{F} = 0.5, F = 0.5, f = 1 \tag{2.2.2}$$

respectively.

Theorem 2.2.2 If the time series $\{x(t)\}$ is short memory ARMA(p,q) time series, then

$$D[x_n(t)]/D[x(t)] \sim cn, \tag{2.2.3}$$

where $n \to \infty$, and c is a constant.

Throughout this paper, I take $x_n \sim y_n$ to mean that $x_n/y_n = 1$, as $n \to \infty$.

Proposition 2.2.2 If the time series $\{x(t)\}$ is short memory ARMA(p,q) series, then weak variance scale exponent F and adjusted proportion coefficient f satisfy:

$$F = 0.5, f = 1 \tag{2.2.4}$$

respectively.

Theorem 2.2.3 If the time series $\{x(t)\}$ is fractional Gaussian noises process (Mandelbrot and Van Ness, 1968), then

$$D[x_n(t)]/D[x(t)] \sim n^{2H}, as\ n \to \infty, \tag{2.2.5}$$

where H is the long memory parameter of fractional Gaussian noises process.

Proposition 2.2.3 If the time series $\{x(t)\}$ is fractional Gaussian noises process, then weak variance scale exponent F and adjusted proportion coefficient f satisfy:

$$F = H, f = 1 \tag{2.2.6}$$

respectively.

Theorem 2.2.4 If the time series $\{x(t)\}$ is fractional differencing noises process (Granger and Joyeux, 1980; Hosking, 1981, ARFIMA(0,d,0) time series), then

$$D[x_n(t)]/D[x(t)] \sim \frac{\Gamma(1-d)}{(1+2d)\Gamma(1+d)} n^{1+2d}, as\ n \to \infty, \tag{2.2.7}$$

where d is the long memory parameter of fractional differencing noises process and $\Gamma(\cdot)$ denotes Gamma function.

Proposition 2.2.4 If the time series $\{x(t)\}$ is fractional differencing noises process, then weak variance scale exponent F and adjusted proportion coefficient f satisfy:

$$F = d + 0.5, f = \frac{\Gamma(1-d)}{(1+2d)\Gamma(1+d)} \tag{2.2.8}$$

respectively.

2.3 Frequency Domain Analysis

For the stationary time series $\{x(t)\}$, if the autocovariance γ_i is absolutely summable, then the spectral density $f(\lambda)$ satisfy:

$$f(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-ik\lambda} = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k \cos(k\lambda). \tag{2.3.1}$$

set $\lambda = 0$, then

$$f(\lambda = 0) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k. \tag{2.3.2}$$

Based on Proposition 2.1.2, I obtain following results.

Theorem 2.3.1 If the time series $\{x(t)\}$ is white noise series, then

$$D[x_n(t)] = 2n\pi f(0), n = 1, 2, \cdots. \tag{2.3.3}$$

Theorem 2.3.2 If the time series $\{x(t)\}$ is short memory ARMA(p,q) series, then

$$D[x_n(t)] \sim 2n\pi f(0), as\ n \to \infty. \tag{2.3.4}$$

Proposition 2.2.1 and Proposition 2.2.2 can also arising from Theorem 2.3.1 and Theorem 2.3.2 respectively.

It need to be pointed out that the autocovariance of stationary long memory time series is not absolutely summable, but the generalized relationship theorem for a stationary time series between spectral density $f(\lambda)$ and variance $D[x_n(t)]$ of time scale-n series is display as following:

Theorem 2.3.3 For the stationary time series $\{x(t)\}$, then

$$D[x_n(t)] = 2\int_{-\pi}^{\pi} f(\lambda)\frac{1-cos(n\lambda)}{[2sin(\lambda/2)]^2}d\lambda, n = 1, 2, \cdots . \tag{2.3.5}$$

Obviously, Theorem 2.3.3 is the generalization of Theorem 2.3.1 and Theorem 2.3.2. Arising from Theorem 2.3.3, I educe Theorem 2.3.4:

Theorem 2.3.4 If the time series $\{x(t)\}$ is generalized ARFIMA(p,d,q) series, then

$$D[x_n(t); p, d, q]/D[x(t)] \sim c(\phi, \theta)\frac{\Gamma(1-d)}{(1+2d)\Gamma(1+d)}n^{1+2d}, as\ n \to \infty. \tag{2.3.6}$$

In which, $c(\phi, \theta)$ is a constant. And if $p = q \equiv 0$, then $c(\phi, \theta) = 1$; otherwise, $c(\phi, \theta) \in [c_1, 1) \cup (1, c_2]$, $c_1 = \frac{\min\{|\theta(1)|^2,|\theta(-1)|^2\}}{\max\{|\phi(1)|^2,|\phi(-1)|^2\}}$ and $c_2 = \frac{\max\{|\theta(1)|^2,|\theta(-1)|^2\}}{\min\{|\phi(1)|^2,|\phi(-1)|^2\}}$.

I now list a implication arising from Theorem 2.3.4:

Proposition 2.3.1 If the time series $\{x(t)\}$ is generalized ARFIMA(p,d,q) time series, then weak variance scale exponent F and adjusted proportion coefficient f satisfy:

$$F = d + 0.5, f = c(\phi, \theta)\frac{\Gamma(1-d)}{(1+2d)\Gamma(1+d)} \tag{2.3.7}$$

respectively.

3 Hypothesis Tests

3.1 *Wnoise* Statistic

I construct two hypothesis test statistics to analyze the memory properties considering white noise, short memory and long memory time series.

If the stationary time series $\{x(t), t = 1, 2, \dots\}$ is independent white noise series, as $n \to \infty$, apply Lindbergh Central Limit Theorem, then

$$\sqrt{n}\left(\sum\nolimits_{t=1}^{n} x(t)/n\ -\mu\right) \xrightarrow{d} N(0, D[x(t)]), \tag{3.1.1}$$

and

$$\frac{x_n(t) - n\mu}{\sqrt{nD[x(t)]}} \xrightarrow{d} N(0, 1). \tag{3.1.2}$$

where $\xrightarrow{d}$ denotes convergence in distribution. Denotes $\hat{D}[x(t)]$ as sample variance, based on Slutsky Theorem, then

$$\frac{x_n(t) - n\sum_{t=1}^{N} x(t)/N}{\sqrt{n\hat{D}[x(t)]}} \xrightarrow{d} N(0, 1), \tag{3.1.3}$$

and

$$(N(n)-1)\,\frac{\hat{D}[x_n(t)]}{n\hat{D}[x(t)]} \xrightarrow{d} \chi^2\left(N(n)-1\right), \tag{3.1.4}$$

where $N(n)$ denotes the time series $\{x_n(t)\}$ sample sizes.

Denote

$$Wnoise(n) = (N(n)-1)\,\frac{\hat{D}[x_n(t)]}{n\hat{D}[x(t)]}. \tag{3.1.5}$$

$Wnoise$ statistic tests under independent white noise hypotheses and non-independent stochastic process alternatives.

3.2 *SLmemory* Statistic

As a generalization situation, if the stationary time series $\{x(t), t = 1, 2, \cdots\}$ is short memory series and apply central limit theorem(Anderson, 1971, Theorem 7.7.8; Brockwell and Davis, 1991, Theorem 7.1.2), then

$$\sqrt{n}\left(\sum\nolimits_{t=1}^{n} x(t)/n\ -\mu\right) \xrightarrow{d} N(0, \sum\nolimits_{i=-\infty}^{\infty}\gamma_i), as\ n\to\infty. \tag{3.2.1}$$

Based on Slutsky Theorem and Proposition 2.1.4, I educe the following conclusion:

$$\frac{x_n(t) - n\sum_{t=1}^{N} x(t)/N}{\sqrt{n(\hat{D}[x_{n+1}(t)]-\hat{D}[x_n(t)])}} \xrightarrow{d} N(0,1). \tag{3.2.2}$$

Then

$$\frac{(N(n)-1)\,\hat{D}[x_n(t)]}{n\left(\hat{D}[x_{n+1}(t)]-\hat{D}[x_n(t)]\right)} \xrightarrow{d} \chi^2\left(N(n)-1\right). \tag{3.2.3}$$

Denote

$$SLmemory(n) = \frac{(N(n)-1)\,\hat{D}[x_n(t)]}{n\left(\hat{D}[x_{n+1}(t)]-\hat{D}[x_n(t)]\right)}. \tag{3.2.4}$$

$SLmemory$ statistic tests under white noise, short memory null hypotheses and long memory alternatives. Obviously, $SLmemory$ statistic is a monotonic decreasing function of long memory parameter d.

Further, if the alternatives are limited in the situation of anti-persistent long memory, the two classical long memory parameters satisfies $-0.5 < d < 0$, $0 < H < 0.5$ respectively, the reject region of hypothesis test of $SLmemory$ statistic is:

$$\left\{SLmemory(n) \geq \chi^2_{1-\alpha}\left(N(n)-1\right)\right\}, \tag{3.2.5}$$

where α is significance size.

And if the alternatives are limited in the situation of persistent long memory, two classical long memory parameters satisfies $0 < d < 0.5$, $0.5 < H < 1$ respectively, the reject region of hypothesis test of $SLmemory$ statistic is:

$$\left\{SLmemory(n) \leq \chi^2_{\alpha}\left(N(n)-1\right)\right\}. \tag{3.2.6}$$

4 Monte Carlo Performance

The objective of this section is to illustrate the asymptotic properties for weak variance scale exponent estimator and *SLmemory* statistic, to examine their finite-sample performance, and to give advice on how to choose the time scale n in practical applications. I focus on the MSE for weak variance scale exponent estimator and the empirical size and power for *SLmemory* statistic. Throughout the simulation exercise, the number of replications is 1000.

4.1 Monte Carlo Study for Weak Variance Scale Exponent Estimator

Based on the definition of weak variance scale exponent, it can be transferred to:

$$y(i) = f\,[z(i)]^F\,, \tag{4.1.1}$$

where $y(i) = D[x_i(t)]/D[x(t)]$,$z(i) = i^2$,$i = 1, 2, \cdots, n$. n is the maximum time scale of finite-sample series. Estimate the parameter F in non-linear regression by sample series $\{y(i)\}$ and $\{z(i)\}$, then gain the weak variance scale exponent estimator $\hat{F}$.

Let stationary time series $\{x(t)\}$ be a linear generalized Gaussian ARFIMA (1,d,1) process $(1-\phi B)(1-B)^d X_t = (1-\theta B)\varepsilon_t$, with unit standard deviation, for different values of ϕ, θ and d. I consider four sample sizes N=250, 500, 1000, 2000, and estimate the parameter d using weak variance scale exponent with maximum time scale $n = \lfloor N^m \rfloor$, where m are the 13 numbers which range from 0.2 to 0.8 and step length by 0.05.

Table 1 contains the MSE of the weak variance scale exponent estimator calculated for different values of maximum time scale $n = \lfloor N^m \rfloor$. When $\phi \neq 0$, or $\theta \neq 0$, the MSE show approximate U-type curve shape, and when $d = \phi = \theta$=0, the MSE increase slowly. Generally, m=0.5 and $n = \lfloor N^{0.5} \rfloor$ could be optimal choice of weak variance scale exponent estimator in practice.

4.2 Monte Carlo Study for *SLmemory* Statistic

The Monte Carlo study for *SLmemory* statistic investigate the percentage of replications in which the rejection of a short memory (or white noise) null hypothesis was observed. Thus if a data generating process belongs to the null hypothesis, the empirical test sizes will be calculated. And if it is a long memory process, the empirical power of the tests will be provided.

Let stationary time series $\{x(t)\}$ be a linear ARMA(1,1) model $(1-\phi B)X_t = (1-\theta B)\varepsilon_t$, with unit standard deviation, for different values of ϕ and θ. I consider two sample sizes N=500, 1000, and choose time scale $n = \lfloor N^m \rfloor$, m are the 13 numbers which are from 0.2 to 0.8 and step length is 0.05. Table 2 display empirical size of ARMA(1,1) model for the situations of different time scale $n = \lfloor N^{0.4} \rfloor$. Table 2 shows that, as m increase, empirical size show approximate U-type curve shape.

Table 3 compares the power of the tests under long memory alternatives. I consider the long memory ARFIMA (1,d,1) model $(1-\phi B)(1-B)^d X_t = (1-\theta B)\varepsilon_t$, with unit standard deviation, for different values of ϕ, θ and d, and for two sample sizes N=500, 1000. In table 3, while m increase, empirical power show approximate inverse U-type curve shape, and obviously contrast to the situation of Table 2.

Considering the significant characteristics of the U-type and inverse U-type curve displayed in table 2 and 3 respectively, I advice that m=0.4 and $n = \lfloor N^{0.4} \rfloor$ can be optimal choice of *SLmemory* statistic in practice. The Monte Carlo study results of empirical size and power for *SLmemory* statistic show the significant advantages, comparing with the literature of long memory tests (Lo, 1991; Giraitis et al., 2003, 2005) which investigate bigger values of long memory parameter, d=1/3, $-1/3$ and d=0.2, 0.3, 0.4 respectively.

Table 1
MSE (in %) of weak variance scale exponent estimator. Let stationary time series {x(t)} be a linear generalized Gaussian ARFIMA(1, d, 1) process $(1-\varphi B)(1-B)^d X_t=(1-\theta B)\varepsilon_t$, with unit standard deviation, for different values of φ, θ and d, and N denotes sample sizes.

d	N=250				N=500				N=1000				N=2000			
	φ=0.5 θ=0	φ=0 θ=−0.5	φ=0.5 θ=−0.5	φ=0 θ=0	φ=0.5 θ=0	φ=0 θ=−0.5	φ=0.5 θ=−0.5	φ=0 θ=0	φ=0.5 θ=0	φ=0 θ=−0.5	φ=0.5 θ=−0.5	φ=0 θ=0	φ=0.5 θ=0	φ=0 θ=−0.5	φ=0.5 θ=−0.5	φ=0 θ=0
	$n=\lfloor N^{0.2}\rfloor=3$				$n=\lfloor N^{0.2}\rfloor=3$				$n=\lfloor N^{0.2}\rfloor=3$				$n=\lfloor N^{0.2}\rfloor=4$			
−0.1	9.645	4.319	15.201	0.287	9.511	4.296	15.222	0.167	9.524	4.232	15.162	0.117	7.740	2.564	11.493	0.082
0	6.817	2.648	10.685	0.216	6.832	2.621	10.693	0.106	6.833	2.598	10.690	0.048	5.488	1.490	8.050	0.027
0.1	4.484	1.431	6.847	0.274	4.496	1.460	6.936	0.164	4.555	1.447	6.957	0.096	3.612	0.784	5.200	0.056
	$n=\lfloor N^{0.25}\rfloor=3$				$n=\lfloor N^{0.25}\rfloor=4$				$n=\lfloor N^{0.25}\rfloor=5$				$n=\lfloor N^{0.25}\rfloor=6$			
−0.1	9.645	4.319	15.201	0.287	7.772	2.655	11.556	0.165	6.354	1.769	9.010	0.101	5.247	1.298	7.260	0.068
0	6.817	2.648	10.685	0.216	5.475	1.532	8.038	0.123	4.410	0.990	6.230	0.063	3.604	0.691	4.962	0.036
0.1	4.484	1.431	6.847	0.274	3.538	0.809	5.140	0.171	2.832	0.487	3.942	0.092	2.288	0.332	3.115	0.055
	$n=\lfloor N^{0.3}\rfloor=5$				$n=\lfloor N^{0.3}\rfloor=6$				$n=\lfloor N^{0.3}\rfloor=7$				$n=\lfloor N^{0.3}\rfloor=9$			
−0.1	6.536	1.936	9.112	0.317	5.337	1.414	7.357	0.184	4.415	1.036	5.972	0.100	3.215	0.688	4.241	0.065
0	4.443	1.139	6.244	0.277	3.635	0.784	4.982	0.164	2.973	0.552	4.015	0.086	2.107	0.344	2.776	0.052
0.1	2.792	0.576	3.836	0.357	2.245	0.401	3.071	0.212	1.828	0.265	2.455	0.110	1.265	0.161	1.657	0.065
	$n=\lfloor N^{0.35}\rfloor=6$				$n=\lfloor N^{0.35}\rfloor=8$				$n=\lfloor N^{0.35}\rfloor=11$				$n=\lfloor N^{0.35}\rfloor=14$			
−0.1	5.473	1.525	7.395	0.343	3.873	0.974	5.141	0.212	2.490	0.565	3.213	0.126	1.740	0.381	2.214	0.080
0	3.656	0.911	4.984	0.323	2.553	0.544	3.366	0.207	1.583	0.306	2.036	0.128	1.066	0.193	1.353	0.077
0.1	2.243	0.481	2.992	0.407	1.515	0.309	2.004	0.258	0.906	0.174	1.165	0.155	0.593	0.112	0.748	0.093
	$n=\lfloor N^{0.4}\rfloor=9$				$n=\lfloor N^{0.4}\rfloor=12$				$n=\lfloor N^{0.4}\rfloor=15$				$n=\lfloor N^{0.4}\rfloor=20$			
−0.1	3.520	1.010	4.461	0.431	2.390	0.679	3.015	0.280	1.643	0.419	2.064	0.153	1.048	0.269	1.295	0.101
0	2.266	0.689	2.903	0.453	1.502	0.429	1.877	0.298	1.001	0.256	1.246	0.167	0.611	0.157	0.751	0.106
0.1	1.329	0.462	1.631	0.557	0.857	0.320	1.060	0.356	0.555	0.180	0.686	0.194	0.333	0.121	0.397	0.125
	$n=\lfloor N^{0.45}\rfloor=11$				$n=\lfloor N^{0.45}\rfloor=16$				$n=\lfloor N^{0.45}\rfloor=22$				$n=\lfloor N^{0.45}\rfloor=30$			
−0.1	2.823	0.883	3.447	0.488	1.690	0.592	2.072	0.343	1.016	0.348	1.242	0.198	0.645	0.228	0.758	0.138
0	1.823	0.679	2.251	0.536	1.074	0.451	1.282	0.391	0.625	0.270	0.736	0.234	0.382	0.180	0.442	0.160
0.1	1.067	0.519	1.228	0.652	0.627	0.392	0.723	0.457	0.372	0.241	0.420	0.271	0.235	0.169	0.255	0.178
	$n=\lfloor N^{0.5}\rfloor=15$				$n=\lfloor N^{0.5}\rfloor=22$				$n=\lfloor N^{0.5}\rfloor=31$				$n=\lfloor N^{0.5}\rfloor=44$			
−0.1	2.098	0.826	2.406	0.619	1.223	0.570	1.439	0.426	0.756	0.367	0.882	0.268	0.478	0.243	0.528	0.192
0	1.438	0.772	1.660	0.714	0.851	0.551	0.958	0.528	0.509	0.336	0.562	0.321	0.327	0.241	0.352	0.237
0.1	0.924	0.716	0.948	0.875	0.582	0.530	0.606	0.614	0.385	0.348	0.388	0.389	0.254	0.249	0.260	0.258
	$n=\lfloor N^{0.55}\rfloor=20$				$n=\lfloor N^{0.55}\rfloor=30$				$n=\lfloor N^{0.55}\rfloor=44$				$n=\lfloor N^{0.55}\rfloor=65$			
−0.1	1.760	0.926	1.918	0.805	1.040	0.643	1.170	0.560	0.672	0.448	0.749	0.382	0.435	0.311	0.463	0.278
0	1.367	0.978	1.485	0.959	0.839	0.724	0.894	0.714	0.538	0.472	0.562	0.468	0.377	0.350	0.387	0.353
0.1	1.084	1.041	1.018	1.232	0.744	0.760	0.719	0.865	0.493	0.513	0.492	0.547	0.370	0.386	0.366	0.400
	$n=\lfloor N^{0.6}\rfloor=27$				$n=\lfloor N^{0.6}\rfloor=41$				$n=\lfloor N^{0.6}\rfloor=63$				$n=\lfloor N^{0.6}\rfloor=95$			
−0.1	1.739	1.179	1.787	1.130	1.115	0.836	1.155	0.800	0.703	0.599	0.778	0.535	0.520	0.429	0.517	0.423
0	1.540	1.335	1.599	1.331	1.023	0.982	1.051	0.979	0.696	0.701	0.706	0.703	0.514	0.508	0.521	0.514
0.1	1.503	1.541	1.379	1.758	1.091	1.140	1.043	1.239	0.723	0.777	0.731	0.795	0.594	0.592	0.568	0.624

Table 1 (*continued*)
MSE (in %) of weak variance scale exponent estimator.

d	N=250				N=500				N=1000				N=2000			
	φ=0.5 θ=0	φ=0 θ=−0.5	φ=0.5 θ=−0.5	φ=0 θ=0	φ=0.5 θ=0	φ=0 θ=−0.5	φ=0.5 θ=−0.5	φ=0 θ=0	φ=0.5 θ=0	φ=0 θ=−0.5	φ=0.5 θ=−0.5	φ=0 θ=0	φ=0.5 θ=0	φ=0 θ=−0.5	φ=0.5 θ=−0.5	φ=0 θ=0
	$n=\lfloor N^{0.65}\rfloor=36$				$n=\lfloor N^{0.65}\rfloor=56$				$n=\lfloor N^{0.65}\rfloor=89$				$n=\lfloor N^{0.65}\rfloor=139$			
−0.1	1.969	1.566	1.959	1.544	1.343	1.162	1.361	1.121	0.899	0.829	0.941	0.794	0.695	0.603	0.654	0.629
0	1.888	1.818	1.917	1.817	1.378	1.402	1.393	1.399	0.986	1.035	0.991	1.037	0.757	0.778	0.757	0.783
0.1	2.145	2.207	1.996	2.415	1.580	1.697	1.583	1.724	1.164	1.198	1.146	1.236	0.922	0.909	0.893	0.952
	$n=\lfloor N^{0.7}\rfloor=47$				$n=\lfloor N^{0.7}\rfloor=77$				$n=\lfloor N^{0.7}\rfloor=125$				$n=\lfloor N^{0.7}\rfloor=204$			
−0.1	2.394	2.091	2.383	2.062	1.781	1.639	1.791	1.604	1.348	1.219	1.300	1.260	0.995	0.962	1.000	0.948
0	2.480	2.425	2.498	2.426	1.967	2.021	1.974	2.019	1.476	1.586	1.477	1.589	1.130	1.166	1.129	1.171
0.1	2.971	3.116	2.920	3.226	2.335	2.376	2.274	2.464	1.909	1.853	1.805	1.971	1.452	1.491	1.476	1.479
	$n=\lfloor N^{0.75}\rfloor=62$				$n=\lfloor N^{0.75}\rfloor=105$				$n=\lfloor N^{0.75}\rfloor=177$				$n=\lfloor N^{0.75}\rfloor=299$			
−0.1	3.158	2.923	3.193	2.852	2.489	2.321	2.449	2.324	1.952	1.885	1.959	1.870	1.545	1.551	1.571	1.501
0	3.535	3.427	3.547	3.419	2.786	2.872	2.787	2.871	2.305	2.411	2.300	2.416	1.774	1.753	1.770	1.764
0.1	4.257	4.406	4.251	4.457	3.471	3.327	3.237	3.583	2.806	2.866	2.835	2.862	2.375	2.389	2.358	2.365
	$n=\lfloor N^{0.8}\rfloor=82$				$n=\lfloor N^{0.8}\rfloor=144$				$n=\lfloor N^{0.8}\rfloor=251$				$n=\lfloor N^{0.8}\rfloor=437$			
−0.1	4.722	4.286	4.560	4.338	3.600	3.398	3.523	3.434	2.985	2.841	2.917	2.891	2.491	2.502	2.511	2.454
0	5.501	4.992	5.539	4.971	4.227	4.438	4.225	4.428	3.453	3.592	3.449	3.599	2.723	2.716	2.755	2.775
0.1	6.447	6.159	6.025	6.567	5.274	4.931	4.851	5.373	4.344	4.182	4.130	4.367	3.625	3.776	3.822	3.699

Table 2
Empirical test sizes (in %) of *SLmemory* statistic of time series under the null hypothesis of ARMA(1, 1) model, $(1-\varphi B)X_t=(1-\theta B)\varepsilon_t$, with standard normal innovations.

N	n	φ=0 θ=0			φ=0.5 θ=0			φ=−0.5 θ=0			φ=0.8 θ=0		
		α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
500	$n=\lfloor N^{0.2}\rfloor=3$	0.223	0.123	0.040	0.999	0.995	0.978	0.994	0.986	0.962	1.000	1.000	1.000
	$n=\lfloor N^{0.25}\rfloor=4$	0.204	0.108	0.037	0.973	0.922	0.564	0.520	0.414	0.285	1.000	1.000	1.000
	$n=\lfloor N^{0.3}\rfloor=6$	0.187	0.101	0.037	0.578	0.336	0.028	0.501	0.405	0.257	1.000	1.000	0.981
	$n=\lfloor N^{0.35}\rfloor=8$	0.167	0.106	0.041	0.261	0.076	0.003	0.431	0.338	0.226	0.989	0.947	0.457
	$n=\lfloor N^{0.4}\rfloor=12$	0.169	0.115	0.048	0.100	0.028	0.009	0.357	0.272	0.164	0.655	0.275	0.000
	$n=\lfloor N^{0.45}\rfloor=16$	0.180	0.120	0.062	0.103	0.053	0.027	0.292	0.220	0.122	0.224	0.024	0.001
	$n=\lfloor N^{0.5}\rfloor=22$	0.176	0.122	0.067	0.106	0.071	0.037	0.248	0.199	0.137	0.063	0.023	0.006
	$n=\lfloor N^{0.55}\rfloor=30$	0.179	0.143	0.093	0.144	0.103	0.074	0.239	0.187	0.124	0.082	0.056	0.026
	$n=\lfloor N^{0.6}\rfloor=41$	0.199	0.168	0.115	0.169	0.139	0.091	0.239	0.184	0.138	0.119	0.089	0.059
	$n=\lfloor N^{0.65}\rfloor=56$	0.231	0.186	0.150	0.248	0.197	0.148	0.264	0.218	0.179	0.194	0.170	0.132
	$n=\lfloor N^{0.7}\rfloor=77$	0.251	0.217	0.178	0.265	0.225	0.189	0.286	0.254	0.205	0.248	0.213	0.176
	$n=\lfloor N^{0.75}\rfloor=105$	0.329	0.300	0.277	0.317	0.288	0.266	0.336	0.309	0.279	0.322	0.292	0.254
	$n=\lfloor N^{0.8}\rfloor=144$	0.363	0.345	0.318	0.363	0.338	0.312	0.363	0.348	0.321	0.373	0.356	0.339
1000	$n=\lfloor N^{0.2}\rfloor=3$	0.210	0.102	0.022	1.000	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000
	$n=\lfloor N^{0.25}\rfloor=5$	0.179	0.094	0.028	0.982	0.941	0.667	0.862	0.816	0.700	1.000	1.000	1.000
	$n=\lfloor N^{0.3}\rfloor=7$	0.150	0.082	0.024	0.679	0.476	0.094	0.613	0.516	0.358	1.000	1.000	1.000
	$n=\lfloor N^{0.35}\rfloor=11$	0.144	0.075	0.028	0.221	0.068	0.003	0.395	0.298	0.172	0.989	0.944	0.505
	$n=\lfloor N^{0.4}\rfloor=15$	0.140	0.076	0.026	0.104	0.026	0.004	0.324	0.253	0.147	0.750	0.440	0.013
	$n=\lfloor N^{0.45}\rfloor=22$	0.124	0.069	0.031	0.087	0.038	0.014	0.256	0.183	0.101	0.228	0.042	0.000

Table 2 (*continued*)
Empirical test sizes (in %) of *SLmemory* statistic of time series under the null hypothesis of ARMA(1, 1) model, $(1-\varphi B)\, X_t=(1-\theta B)\varepsilon_t$, with standard normal innovations.

N	*n*	φ=0 θ=0			φ=0.5 θ=0			φ=−0.5 θ=0			φ=0.8 θ=0		
		α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
1000	$n=\lfloor N^{0.5}\rfloor=31$	0.157	0.091	0.046	0.106	0.064	0.026	0.240	0.166	0.090	0.082	0.025	0.012
	$n=\lfloor N^{0.55}\rfloor=44$	0.172	0.128	0.060	0.142	0.098	0.044	0.213	0.153	0.096	0.085	0.059	0.032
	$n=\lfloor N^{0.6}\rfloor=63$	0.196	0.143	0.089	0.156	0.119	0.070	0.204	0.154	0.093	0.138	0.100	0.063
	$n=\lfloor N^{0.65}\rfloor=89$	0.234	0.193	0.132	0.217	0.169	0.129	0.251	0.206	0.143	0.170	0.143	0.110
	$n=\lfloor N^{0.7}\rfloor=125$	0.267	0.239	0.191	0.285	0.244	0.188	0.296	0.248	0.196	0.231	0.193	0.150
	$n=\lfloor N^{0.75}\rfloor=177$	0.322	0.285	0.243	0.304	0.275	0.240	0.317	0.288	0.251	0.284	0.255	0.217
	$n=\lfloor N^{0.8}\rfloor=251$	0.383	0.355	0.322	0.376	0.349	0.329	0.390	0.362	0.332	0.373	0.346	0.322

Table 3
Empirical power of the tests (in %) based on *SLmemory* statistic, the alternatives considered are ARFIMA(1, *d*, 1) model, $(1-\varphi B)(1-B)^d X_t=(1-\theta B)\varepsilon_t$, with standard normal innovations.

N	*n*	d=0.1						d=−0.1					
		φ=0.5 θ=0			φ=0 θ=0			φ=0.5 θ=0			φ=0 θ=0		
		α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01	α=0.1	α=0.05	α=0.01
500	$n=\lfloor N^{0.2}\rfloor=3$	0.000	0.000	0.000	0.362	0.530	0.827	0.038	0.092	0.381	0.337	0.417	0.607
	$n=\lfloor N^{0.25}\rfloor=4$	0.001	0.003	0.044	0.465	0.652	0.932	0.349	0.536	0.908	0.37	0.487	0.675
	$n=\lfloor N^{0.3}\rfloor=6$	0.072	0.206	0.702	0.645	0.836	0.991	0.825	0.943	0.994	0.437	0.542	0.708
	$n=\lfloor N^{0.35}\rfloor=8$	0.331	0.598	0.977	0.732	0.894	0.996	0.860	0.939	0.981	0.523	0.637	0.769
	$n=\lfloor N^{0.4}\rfloor=12$	0.730	0.936	0.997	0.874	0.965	0.991	0.804	0.870	0.942	0.580	0.679	0.789
	$n=\lfloor N^{0.45}\rfloor=16$	0.880	0.986	0.999	0.895	0.968	0.992	0.760	0.819	0.891	0.616	0.698	0.805
	$n=\lfloor N^{0.5}\rfloor=22$	0.927	0.976	0.987	0.921	0.963	0.983	0.719	0.790	0.870	0.617	0.686	0.799
	$n=\lfloor N^{0.55}\rfloor=30$	0.925	0.949	0.972	0.906	0.942	0.961	0.685	0.749	0.837	0.642	0.704	0.800
	$n=\lfloor N^{0.6}\rfloor=41$	0.884	0.907	0.933	0.879	0.901	0.931	0.700	0.758	0.813	0.675	0.709	0.784
	$n=\lfloor N^{0.65}\rfloor=56$	0.844	0.869	0.903	0.837	0.870	0.900	0.670	0.74	0.804	0.658	0.714	0.783
	$n=\lfloor N^{0.7}\rfloor=77$	0.780	0.808	0.840	0.773	0.803	0.841	0.648	0.685	0.728	0.628	0.672	0.730
	$n=\lfloor N^{0.75}\rfloor=105$	0.689	0.711	0.746	0.690	0.715	0.748	0.620	0.655	0.677	0.593	0.633	0.676
	$n=\lfloor N^{0.8}\rfloor=144$	0.633	0.658	0.669	0.627	0.647	0.671	0.572	0.59	0.613	0.558	0.584	0.608
1000	$n=\lfloor N^{0.2}\rfloor=3$	0.000	0.000	0.000	0.102	0.204	0.504	0.000	0.001	0.016	0.131	0.196	0.378
	$n=\lfloor N^{0.25}\rfloor=5$	0.000	0.000	0.005	0.290	0.468	0.828	0.458	0.663	0.942	0.230	0.335	0.523
	$n=\lfloor N^{0.3}\rfloor=7$	0.012	0.052	0.305	0.434	0.643	0.943	0.848	0.943	0.996	0.342	0.432	0.623
	$n=\lfloor N^{0.35}\rfloor=11$	0.250	0.510	0.949	0.625	0.855	0.992	0.838	0.919	0.973	0.430	0.560	0.750
	$n=\lfloor N^{0.4}\rfloor=15$	0.549	0.807	0.998	0.749	0.937	0.998	0.753	0.844	0.930	0.491	0.612	0.757
	$n=\lfloor N^{0.45}\rfloor=22$	0.787	0.951	0.997	0.859	0.965	0.995	0.699	0.796	0.899	0.560	0.657	0.800
	$n=\lfloor N^{0.5}\rfloor=31$	0.912	0.991	0.996	0.926	0.985	0.995	0.717	0.787	0.872	0.621	0.705	0.811
	$n=\lfloor N^{0.55}\rfloor=44$	0.940	0.972	0.988	0.932	0.967	0.988	0.697	0.759	0.856	0.634	0.709	0.818
	$n=\lfloor N^{0.6}\rfloor=63$	0.907	0.928	0.961	0.896	0.926	0.960	0.674	0.746	0.821	0.647	0.714	0.800
	$n=\lfloor N^{0.65}\rfloor=89$	0.848	0.878	0.916	0.843	0.878	0.918	0.671	0.719	0.784	0.641	0.703	0.753
	$n=\lfloor N^{0.7}\rfloor=125$	0.782	0.814	0.853	0.786	0.817	0.848	0.644	0.694	0.747	0.634	0.684	0.739
	$n=\lfloor N^{0.75}\rfloor=177$	0.748	0.768	0.795	0.748	0.765	0.795	0.634	0.667	0.716	0.631	0.663	0.693
	$n=\lfloor N^{0.8}\rfloor=251$	0.654	0.684	0.710	0.658	0.680	0.712	0.595	0.621	0.648	0.600	0.625	0.650

4.3 Application to Sino-US Stock Index Return Rate Data

I illustrate the theory and simulations by applying the weak variance scale exponent estimator and *SLmemory* statistic to the logarithmic return rate data of Sino-US stock index. To make the results comparable with the simulations, I divide each logarithmic return rate series into three blocks of approximate lengths (about 1000) as shown in table 4.

Table 4
Descriptive Statistics. CN. and US. Block denotes Shanghai Composite and Standard & Poor 500 index data respectively. Block 1 is for a period of four years from January 1, 2000 to December 31, 2003. Block 2 is also for a period of four years from January 1, 2004 to December 31, 2007. Block 3 ranges a period from January 1, 2008 to May 25, 2012.

	Sample Sizes	Mean	Std.Dev	Minimum	Maximum	Skewness	Kurtosis
CN.Block 1	957	0.0095	1.3666	-6.5437	9.4014	0.7929	7.5504
CN.Block 2	968	0.1298	1.6209	-9.2562	7.8903	-0.4333	3.1152
CN.Block 3	1069	-0.0761	1.9062	-8.0437	9.0348	-0.1541	2.8143
US.Block 1	1004	-0.0278	1.3814	-6.0045	5.5744	0.1447	1.2812
US.Block 2	1006	0.0276	0.7612	-3.5343	2.8790	-0.3088	1.7889
US.Block 3	1110	-0.0097	1.7393	-9.4695	10.9572	-0.2300	6.3233

The weak variance exponent estimators, the *SLmemory* statistic values and corresponding P-values are displayed in table 5. The evidence against the null hypothesis in favor of long memory alternative is strong for the second blocks of Sino-US stock index series which the weak variance exponent estimators are 0.6379, 0.4239 respectively. The first and third blocks of Sino-US stock index series show that their weak variance exponent estimators are all close to 0.5 and the corresponding P-values are all beyond 0.3, that it is inclined to believe that the data exhibit some forms of accepting the null hypothesis discussed in the present paper. In a word, the weak variance exponent estimators are consist with statistical tests for Sino-US stock index series and the results show that there is long memory in Sino-US stock index series for the period that from January 1, 2004 to December 31, 2007, and no long memory in Sino-US stock index series for the other two periods.

Table 5
Estimation and Test Results for Sino-US Stock Index Data.

	CN.Block 1	CN.Block 2	CN.Block 3	US.Block 1	US.Block 2	US.Block 3
$\hat{F}$	0.5093	0.6379	0.5306	0.4596	0.4239	0.5045
$\hat{f}$	0.9752	0.5831	0.9059	1.0177	1.0046	0.6811
SLmemory	57.5542	41.5572	59.4155	66.4455	86.64281	63.36272
Freedom	62	63	65	65	66	68
P-value	0.3636	0.0169	0.3279	0.4269	0.0451	0.3633

It is seen again that applying $n = \lfloor N^{0.5} \rfloor$ for weak variance scale exponent estimator and $n = \lfloor N^{0.4} \rfloor$ for *SLmemory* statistic test can be recommended in practice. For the application to Sino-US stock data, I find that there is different long memory properties between Sino-US stock data series in the period of Block 2, the long memory properties of Shanghai Composite index data series is persistent, and anti-persistent for Standard & Poor 500. Persistence means positive correlated long-range effect and negative for anti-persistence. The economics meanings for the periods of Block 2 can be interpret that Shanghai stock market was more vulnerable to the impact of external events and it lasted for a long time, while Standard & Poor stock market can quickly return to stability under the influence of external shocks.

5 Conclusions

The goal of this paper is to introduce a new system standpoint for theory analysis of stationary time series, especially for long memory properties, referring to estimation and test for long memory properties in the present paper. Combining with both time and frequency domain analysis, the properties of strong and weak variance scale exponent of white noise, short memory ARMA(p,q) and long memory ARFIMA(p,d,q) time series have been derived. And particularly conclude the equations between weak variance scale exponent and long memory parameter $d(-0.5 < d < 0.5)$, $H(0 < H < 1)$. Further, under independent white noise null hypotheses and non-independent stochastic process alternatives, and under white noise, short memory null hypothesis and long memory alternatives, I construct $Wnoise$ and $SLmemory$ statistic tests respectively. The estimation and test allow for the possibility of memory properties of persistence or anti-persistence in theory which differ V/S statistic tests (Giraitis et al., 2003, 2005) only for persistence situations.

A Monte Carlo study for weak variance scale exponent and $SLmemory$ statistic both refer to the selection of maximum time scale n of finite-sample. By minimizing the MSE for estimation and optimizing empirical size and power for tests, finally find that $n = \lfloor N^{0.5} \rfloor$ for weak variance scale exponent estimator and $n = \lfloor N^{0.4} \rfloor$ for $SLmemory$ statistic test can be recommended in practice. In fact, time scale n can also be regarded as an analogy of bandwidth (Abadir et al., 2009).

Further study of the present paper, maybe consider that, to analyze the properties of strong and weak variance scale exponent for non-stationary or non-invertibility time series, more complicated long memory model and corresponding statistical tests. Surely, on estimation of autocovariance, proposition 2.1.3 can be considered as a replacement of Durbin-Levinson algorithm.

Acknowledgments

This paper is mainly based on chapter 4 and 5 of my dissertation at Yunnan University of Finance and Economics. And the research was partially supported by the Yunnan University of Finance and Economics Graduate Student Innovation Project grant YFGI(2011)019. I am grateful to my supervisor Professor Guoqing Zhao for leading me interested in the topic of long memory time series and giving me important guidances. I also thank Yunnan University of Finance and Economics Professor Peng Bai for useful helps during my studying on the topic and the constructive comments during my thesis defense.

Appendix. Proofs of The Theorems and An Auxiliary Lemma

Proof of Theorem 2.2.2 Proposition 2.1.2 can rewrite

$$D\left[x_n(t)\right] = \sum_{i=-n}^{n} (n-i)\gamma_i = n\gamma_0 \sum_{i=-n}^{n} \rho_i - \sum_{i=-n}^{n} i\gamma_i. \tag{A.1}$$

For the short memory time series, the cumulant of autocorrelation ρ_i convergence

$$\lim_{n\to\infty} \sum_{-n}^{n} \rho_i \to c, \tag{A.2}$$

where c is a constant. And under the condition of short memory time series, applying Kronecker lemma:

$$\frac{1}{n}\sum\nolimits_{i=1}^{n} i\gamma_i \to 0, as\ n \to \infty. \tag{A.3}$$

to complete the proof of Theorem 2.2.2. □

Proof of Theorem 2.2.3 The Lag-n autocovariance γ_n of fractional Gaussian noises process $\{x(t)\}$ satisfies:

$$\gamma_n = \frac{\gamma_0}{2}\left[(n+1)^{2H} - 2n^{2H} + (n-1)^{2H}\right], as\ n \to \infty. \tag{A.4}$$

Using proposition 2.1.3,

$$\gamma_n = \frac{1}{2}\{D[x_{n+1}(t)] - 2D[x_n(t)] + D[x_{n-1}(t)]\}, n = 1, 2, \cdots. \tag{A.5}$$

I obtain

$$D[x_n(t)]/D[x(t)] \sim n^{2H}, as\ n \to \infty. \tag{A.6}$$

□

Proof of Theorem 2.2.4 Proposition 2.1.2 can rewrite

$$D[x_n(t)] = n\gamma_0 + 2\gamma_0\left[(n+d)\sum\nolimits_{k=1}^{n}\rho_k - \sum\nolimits_{k=1}^{n}(k+d)\rho_k\right]. \tag{A.7}$$

Denote $S_1(n) = \sum_{k=1}^{n}\rho_k$, $S_2(n) = \sum_{k=1}^{n}(k+d)\rho_k$ respectively, that

$$\begin{aligned} S_1(n) &= \sum\nolimits_{k=1}^{n}\rho_k \\ &= \frac{d}{1-d} + \frac{d(1+d)}{(1-d)(2-d)} + \cdots + \frac{d(1+d)\cdots(n-1+d)}{(1-d)(2-d)\cdots(n-d)} \\ &= \frac{1}{2}\left(\frac{1+d}{1-d} - 1\right) + \frac{1}{2}\frac{1+d}{1-d}\left(\frac{2+d}{2-d} - 1\right) + \cdots + \\ &\frac{1}{2}\frac{(1+d)(2+d)\cdots(n-1+d)}{(1-d)(2-d)\cdots(n-1-d)}\left[\frac{(n+d)}{(n-d)} - 1\right] \\ &= -\frac{1}{2} + \frac{1}{2}\frac{\Gamma(1-d)\Gamma(n+1+d)}{\Gamma(1+d)\Gamma(n+1-d)}, \end{aligned} \tag{A.8}$$

and

$$\begin{aligned} S_2(n) &= \sum\nolimits_{k=1}^{n}(k+d)\rho_k \\ &= d\left[\frac{1+d}{1-d} + \frac{(1+d)(2+d)}{(1-d)(2-d)} + \cdots + \frac{(1+d)(2+d)\cdots(n-1+d)(n+d)}{(1-d)(2-d)\cdots(n-1-d)(n-d)}\right] \\ &= d\frac{1+d}{1+2d}\left(\frac{2+d}{1-d} - 1\right) + d\frac{1+d}{1+2d}\left[\frac{(2+d)(3+d)}{(1-d)(2-d)} - \frac{(2+d)}{(1-d)}\right] + \cdots + \\ &d\frac{1+d}{1+2d}\left[\frac{(2+d)(3+d)\cdots(n+d)(n+1+d)}{(1-d)(2-d)\cdots(n-1-d)(n-d)} - \frac{(2+d)(3+d)\cdots(n+d)(n-d)}{(1-d)(2-d)\cdots(n-1-d)(n-d)}\right] \\ &= d\left[-\frac{1+d}{1+2d} + \frac{1}{1+2d}\frac{\Gamma(1-d)\Gamma(n+2+d)}{\Gamma(1+d)\Gamma(n+1-d)}\right]. \end{aligned} \tag{A.9}$$

Then

$$D[x_n(t)] = \gamma_0 \left[\frac{d}{1+2d} + \frac{1}{1+2d} \frac{\Gamma(1-d)\Gamma(n+1+d)}{\Gamma(1+d)\Gamma(n-d)} \right]. \tag{A.10}$$

By Stirling formula: $\Gamma(x) \sim \sqrt{2\pi} e^{1-x}(x-1)^{x-1/2}$, as $x \to \infty$. Therefore

$$D[x_n(t)]/D[x(t)] \sim \frac{1}{1+2d} \frac{\Gamma(1-d)}{\Gamma(1+d)} n^{1+2d}, -1/2 < d < 1/2, n \to \infty. \tag{A.11}$$

□

Proof of Theorem 2.3.3 Proposition 2.1.2 can rewrite

$$\begin{aligned} D[x_n(t)] &= n\gamma_0 + 2\sum\nolimits_{i=1}^{n} (n-i)\gamma_i \\ &= n\int_{-\pi}^{\pi} f(\lambda)d\lambda + 2\int_{-\pi}^{\pi} f(\lambda) \sum\nolimits_{i=1}^{n-1} (n-i)\cos(i\lambda)d\lambda. \end{aligned} \tag{A.12}$$

Applying the summation formula of $\sum_{i=1}^{n-1} \cos(i\lambda)$ and $\sum_{i=1}^{n-1} i\cos(i\lambda)$ (Gradshteyn and Ryzhik, 2007, 37-38), to prove Theorem 2.3.3. □

Proof of Theorem 2.3.4 For the generalized ARFIMA(p,d,q) time series $\{x(t)\}$, spectral density $f(\lambda)$:

$$f(\lambda;p,d,q) = \frac{\sigma^2}{2\pi} \left[2\sin\frac{\lambda}{2}\right]^{-2d} \left| \frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})} \right|^2, -1/2 < d < 1/2 \textit{ and } d \neq 0. \tag{A.13}$$

(1) if $p = q = 0$, spectral density $f(\lambda)$:

$$f(\lambda;0,d,0) = \frac{\sigma^2}{2\pi} \left[2\sin\frac{\lambda}{2}\right]^{-2d}. \tag{A.14}$$

By Theorem 2.3.3,

$$\begin{aligned} D[x_n(t);0,d,0] &= 2\int_{-\pi}^{\pi} f(\lambda;0,d,0) \frac{1-\cos(n\lambda)}{[2\sin(\lambda/2\ )]^2} d\lambda \\ &= 2\int_{-\pi}^{\pi} \frac{\sigma^2}{2\pi} \left[2\sin\frac{\lambda}{2}\right]^{-2d-2} [1-\cos(n\lambda)]d\lambda. \end{aligned} \tag{A.15}$$

Applying the integral formula for $\int_0^{\frac{\pi}{2}} \cos^{\nu-1}x\cos axdx = \frac{\pi}{2^\nu \nu}/B\ (\frac{\nu+a+1}{2}, \frac{\nu-a+1}{2})$ and $\int_0^{\frac{\pi}{2}} \sin^{\mu-1}xdx = 2^{\mu-2}B(\frac{\mu}{2}, \frac{\mu}{2})$ (Gradshteyn and Ryzhik, 2007, 37-38), to prove Theorem 2.3.4.
(2) if $p \neq 0$, or $q \neq 0$, by Theorem 2.3.3,

$$D[x_n(t);p,d,q] = 2\int_{-\pi}^{\pi} \frac{\sigma^2}{2\pi} \left[2\sin\frac{\lambda}{2}\right]^{-2d-2} \left| \frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})} \right|^2 (1-\cos n\lambda)\, d\lambda. \tag{A.16}$$

Applying Lemma A, then

$$c_1 D[x_n(t);0,d,0] \leq D[x_n(t);p,d,q] \leq c_2 D[x_n(t);0,d,0], \tag{A.17}$$

where $c_1 = \frac{\min\{|\theta(1)|^2, |\theta(-1)|^2\}}{\max\{|\phi(1)|^2, |\phi(-1)|^2\}}$ and $c_2 = \frac{\max\{|\theta(1)|^2, |\theta(-1)|^2\}}{\min\{|\phi(1)|^2, |\phi(-1)|^2\}}$.
Therefore,

$$D[x_n(t);p,d,q] = c(\phi,\theta) D[x_n(t);0,d,0], \tag{A.18}$$

where $c(\phi,\theta)$ is a constant, and satisfies $c(\phi,\theta) \in [c_1, c_2]$.
These complete the proof of Theorem 2.3.4. $\square$

Auxiliary Lemma
Lemma A. Let $\phi(e^{-i\lambda}) = 1 - \sum_{k=1}^{p} a_k e^{-i\lambda}$ and $\theta(e^{-i\lambda}) = 1 - \sum_{k=1}^{q} b_k e^{-i\lambda}$ are autoregressive and moving-average multinomial of a stationary and invertible ARMA(p,q) time series respectively, and have no common zeros. then

$$\frac{\min\left\{|\theta(1)|^2, |\theta(-1)|^2\right\}}{\max\left\{|\phi(1)|^2, |\phi(-1)|^2\right\}} \le \left|\frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})}\right|^2 \le \frac{\max\left\{|\theta(1)|^2, |\theta(-1)|^2\right\}}{\min\left\{|\phi(1)|^2, |\phi(-1)|^2\right\}} \; and \; \left|\frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})}\right|^2 \ne 1. \tag{A.19}$$

Proof of Lemma A
For $\phi(e^{-i\lambda})$ and $\theta(e^{-i\lambda})$, have no common zeros ,then

$$\left|\frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})}\right|^2 \ne 1. \tag{A.20}$$

By autoregressive and moving-average multinomial, I obtain

$$\begin{aligned} \left|\phi(e^{-i\lambda})\right|^2 &= \left|1 - \sum\nolimits_{k=1}^{p} a_k e^{-i\lambda}\right| = \left|1 + \left(\sum\nolimits_{k=1}^{p} a_k\right)^2 - 2\cos\lambda \sum\nolimits_{k=1}^{p} a_k\right| \\ \left|\theta(e^{-i\lambda})\right|^2 &= \left|1 - \sum\nolimits_{k=1}^{q} b_k e^{-i\lambda}\right| = \left|1 + \left(\sum\nolimits_{k=1}^{q} b_k\right)^2 - 2\cos\lambda \sum\nolimits_{k=1}^{q} b_k\right|. \end{aligned} \tag{A.21}$$

Under the conditions of stationarity and invertibility, I conclude that $\left|\phi(e^{-i\lambda})\right|^2$ and $\left|\theta(e^{-i\lambda})\right|^2$ are both monotone functions of λ, and satisfy:

$$\begin{aligned} &\min\left\{|\phi(1)|^2, |\phi(-1)|^2\right\} \le \left|\phi(e^{-i\lambda})\right|^2 \le \max\left\{|\phi(1)|^2, |\phi(-1)|^2\right\} \; and \; \left|\phi(e^{-i\lambda})\right|^2 \ne 1 \\ &\min\left\{|\theta(1)|^2, |\theta(-1)|^2\right\} \le \left|\theta(e^{-i\lambda})\right|^2 \le \max\left\{|\theta(1)|^2, |\theta(-1)|^2\right\} \; and \; \left|\theta(e^{-i\lambda})\right|^2 \ne 1. \end{aligned} \tag{A.22}$$

This completes the proof. $\square$

References

[1] Abadir, K.M., Distaso, W., Giraitis, L., 2009. Two estimator of the long-run variance: Beyond short memory. Journal of Econometrics 150, 56-70.

[2] Anderson, T.W., 1971. The Statistical Analysis of Time Series, New York: Wiley.

[3] Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5-59.

[4] Beran, J., 1994. On a class of M-estimators for Gaussian long-memory models. Biometrika 81(4), 755-766.

[5] Boutahar, M., Marimoutou, V., Nouira, L., 2007. Estimation methods of the long memory parameter: Monte Carlo analysis and application. Journal of Applied Statistics 34, 261-301.

[6] Brockwell, P., Davis, R., 1991. Time Series: Theory and Methods(2nd ed.), Berlin: Springer- Verlag.

[7] Fox, R., Taqqu, M.S., 1986. Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series. Annals of Statistics 14(2), 517-532.

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

[9] Giraitis, L., Kokoszka, P., Leipus, R., Teyssiere, G., 2003. Rescaled variance and related tests for long memory in volatility and levels. Journal of Econometrics 112(2), 265-294.

[10] Giraitis, L., Kokoszka, P., Leipus, R., Teyssiere, G., 2005. Corrigendum to Rescaled variance and related tests for long memory in volatility and levels. Journal of Econometrics 126(2), 571-572.

[11] Granger, Clive W.J., Joyeux, R., 1980. An Introduction to Long-memory Time Series Models and Fractional Differencing. Journal of Time Series Analysis 1(1), 15-29.

[12] Gradshteyn, I.S., Ryzhik, I.M., 2007. Table of Integrals, Series, and Products(Seventh Edition). Academic Press, Elsevier Inc.

[13] Haslett, J., Raftery, A.E., 1989. Space-Time Modelling with Long-Memory Dependence: Assessing Irelands Wind Power Resource. Journal of the Royal Statistical Society. Series C (Applied Statistics) 38(1), 1-50.

[14] Hosking, J.R.M., 1981. Fractional differencing. Biometrika 68(1), 165-176.

[15] Hurst, H.E., 1951. Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers 116, 770-779.

[16] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? Journal of Econometrics 54, 159-178.

[17] Lo, A.W., 1991. Long-term memory in stock market prices. Econometrica 59 (5), 1279-1313.

[18] Mandelbrot, B.B., Wallis, J.R., 1968. Noah, Joseph, and Operational Hydrology. Water Resources Research 4(5), 909-918.

[19] Mandelbrot, B.B., Wallis, J.R., 1969. Robustness of the Rescaled Range R/S in the measurement of Noncyclic Long-run Statistical Dependence. Water resources Research 5(5), 967-988.

[20] Mandelbrot, B.B., 1975. Limit theorems of the self-normalized range for weakly and strongly dependent processes. Probability Theory and Related Fields 31(4), 271-285.

[21] Mandelbrot, B.B., Taqqu, M.S., 1979. Robust R/S analysis of long run serial correlation. 42nd Session of the International Statistical Institute, Manila, Book 2, 69-99.

[22] Palma, W., 2007. Long-Memory Time Series (Theory and Methods). New Jersey: John Wiley & Sons, Inc.

[23] Robinson, P.M., 1995. Gaussian Semiparametric Estimation of Long Range Dependence. The Annals of Statistics 23(5), 1630-1661.

[24] Robinson, P.M., 2003. Time Series with Long Memory. New York: Oxford University Press Inc.

[25] Robinson, P.M., 2005. Robust covariance matrix estimation: HAC estimates with long memory/ anti-persistence correction. Econometric Theory 21, 171-180.