International Journal of Business Marketing and Management (IJBMM)
Volume 7 Issue 2 Mar-Apr 2022, P.P. 26-34
ISSN: 2456-4559
www.ijbmm.com

Sustainable arbitrage based on long-term memory via SEV

Shuo Li, Panlu Shi, Peibiao Zhao

School of Science, Nanjing University of Science and Technology, Nanjing, 210094, Jiangsu, P. R. China, pbzhao@njust.edu.cn

Abstract: *The long-term memory means a sustainable arbitrage or no-arbitrage. There are many methods to determine the characteristics of long-term memory of a financial series.*

In the present paper we investigate and propose a new criterion to measure the characteristics of long-term memory of sequences with any correlations via a generalized SEV argument. The superior performance of this new approach here can be illustrated in empirical analysis. This conclusion implies that there is a new sustainable arbitrage (or no arbitrage) principle via a SEV criterion.

Key words and phrases: *Arbitrage; Long-term memory; Generalized measures of correlation; Nonparametric estimation.*

I. Introduction

The efficient market hypothesis is the most important foundation of the modern financial theories. Since Eugene Fama presented it, many theories based on the efficient market hypothesis have been established and developed. However, it has been challenged because of some deficiencies such like too many hypothetical conditions and the difficulty of empirical testing. In the 1990s, in order to relax the assumption of normal distribution and make the theory closer to reality, some scholars put forward the fractal market hypothesis. The long-term memory is one of the main characteristics of fractal market, it means that the autocorrelations of returns at very high lags, decay at a slower, hyperbolic rate, rather than at the exponential rate in a standard ARMA process.

The long-term memory of time series is closely related to the observed data, and this phenomenon is not first noticed in the economic and financial fields. In fact, some scholars found the long-term memory during the data analysis in the fields of physics and hydrology in the 1950s. Hurst[7], Mandelbrot and Wallis[9], Mcleod and Hipel[10] have studied the long-term memory of time series in the fields of hydraulics, meteorology and other natural sciences. A classic example of long-term memory is the continuity of the annual lowest river stage of the Nile from 1007 to 1206, and it is known as the Hurst effect. Along with the deepening of research, we found that many other financial time series have such characteristic. So the long-term memory has been a focus of the financial research in past 20 years.

There are three main methods to estimate the long-term memory parameters of stock market returns. Firstly, the non-parametric estimation methods, including classical R/S method, modified R/S method etc.[8]. The second is parameter estimation method, which uses the unconditional exact likelihood function derived by Sowell[14] to estimate the parameters of a stationary univariate fractionally integrated time series. And the last one, semi-parameter estimation methods mainly including log-periodogram(LP), local Whittle estimation(LW) and exact local whittle(ELW) etc.[12]. So it is not difficult to find that there is still no uniform judgment standard for long-term memory, and many shortcomings need to be solved in studies.

The long-term memory of financial asset returns refers to the remarkable autocorrelation of the return series, even apart of many spacing, which means the historical events would influence future events for a long time. And it is noteworthy that self-similarity is a special kind of autocorrelation, so self-similarity process must be a long-term memory process. Long-term memory can be defined in both time and frequency domain. In time domain, it appears as the autocorrelation coefficient decays at a slow hyperbolic rate. In frequency domain, the autocorrelation coefficient is given from the spectrum.

In this paper, we mainly research the definition in time domain. Considering a time series $x_t, t = 1,2,3, \ldots, T,$ and we define ρ_k as the autocorrelation at lag k. Then a correlogram can represent the autocorrelation function of a time series and describe some linear properties in most cases. The autocorrelation function ρ_k of the stable and reversible ARMA model is finite. However, many time series observed in practice have long range dependence that maybe not obvious but cannot be ignored. So the autocorrelation satisfies

$$\rho_k \sim k^{-\alpha} M(k), k \to \infty \tag{1}$$

where $M(k)$ is any slowly varying function at infinity and is described in detail by Resnick[11].

McLeod and Hippel proposed a usual definition of long-term memory is that

$$\lim_{n\to\infty} \sum_{k=-n}^{n} |\rho_k| = \infty \tag{2}$$

And a wider definition of long memory is to include any process which possesses an autocovariance function for large k, such that

$$\gamma_k \approx \Theta(k) k^{2H-2} \tag{3}$$

Where $\Theta(k)$ is a slowly varying function. Helson and Sarason[6] show that any process with H>0 and autocovariance function given by (3) is long range dependent.

Certainly, there are several different definitions. And we can see that Pearson's correlation coefficient is generally used as a tool of measuring the dependence of time series. But this coefficient can only consider the linear correlation. Therefore, we noticed that ignoring nonlinear correlation may cause error results of measuring long-term memory. Finding a reasonable approach that can solve this problem is what we mainly did in the article.

In fact, an alternative definition ‘extended memory’ is advanced by Granger and Ding[4], which attempts to capture a more general feature of data. Let $g_{h,n}(I_n) = I[x_{n+h}|x_{n-j}, j \geq 0]$ be the optimum least-squares forecast of x_{n+h} based on the information set $I_n: x_{n-j}, j \geq 0$. If for all $n, g_{h,n}(I_n)$ does not tend to a constant as h become large, then x_t is said to have extended memory. The definition has more impact if deterministic processes such as chaotic and series with limit cycles are excluded from consideration. As the extended memory not only considers the linear properties of time series, but also can include optimum nonlinear forecasts, obviously it is a effective supplement to long-term memory. So far, little research has been done for this conception. So in this perspective, we chose the generalized correlation coefficient, SEV_k, which can measure both linear and nonlinear correlation[2], to replace the autocorrelation coefficient ρ_k in judging methods. We hope to solve the problem of ignoring nonlinear correlation when measuring the long range dependence and prove that SEV_k has an improved performance compared with ρ_k by some specific examples.

II. Preliminaries

According to the above, we need a new correlation coefficient for both linear and non-linear correlation. In this paper, we use the SEV (sure explained variable) which is derived from the generalized measures of correlation-GMC.

The concept of GMC was proposed by Zheng et al. in 2012[15][1]. It's definition comes from the formula of variance decomposition:

$$var(Y) = var\big(E(Y|X)\big) + E(var(Y|X)) \tag{4}$$

where $E(X^2 < \infty \; and \; E(Y^2) < \infty$. $E(Y|X)$ is the condition expectation of Y under given X, and $var(E(Y|X))$ measures the variance of the condition expectation of Y due to X. Then $var(E(Y|X)) = var(Y)$ is the explained variance of Y by X, and define

$$GMC(Y|X) = \frac{var\big(E(Y|X)\big)}{var(Y)} = 1 - \frac{E\big(var(Y|X)\big)}{var(Y)}$$

$$= 1 - \frac{E\big(Y - E(Y|X)^2\big)}{var(Y)} \tag{5}$$

GMC always appears in pairs, i.e. $GMC(Y|X)$ and $GMC(X|Y)$, variables X and Y have the same meaning and there is no difference between response variables and exploratory variables, so it can be used research the asymmetry in explained variance. However, in this paper we need to use correlation coefficient to describe the long-term memory of financial time series, which means the influence of previous data on future data. So we suppose Y is always a response variable and X is an exploratory variable, then get

$$SEV\,(Y|X) = 1 - \frac{E\big[Y - E(Y|X)^2\big]}{var(Y)} \tag{6}$$

as the correlation coefficient we want.

Here in the notation, the similarities and differences and relations of Pearson's correlation coefficient and SEV will be clear from the following Proposition 1.

Proposition 1: Suppose Y is a response variable with a finite second moment, X is an exploratory variable, ε is a random error with finite variance. $g(\cdot)$ is a measurable function and $var(g(X))$ exists. X and ε are independent. The following properties hold:

(1) If $Y = g(X)$, almost surely(a.s.), then $SEV(Y|X) = 1$.

(2) If $Y = g(X) + \varepsilon$, then $SEV(Y|X) = var\big(g(X)\big)/(var(g(X)) + var(\varepsilon))$.

(3) If $Y = ag(X) + b + \varepsilon$, where $g(x) = x$, constants $a \neq 0$, b, then $SEV(Y|X) = \rho_{XY}^2$. Here ρ_{XY} is Pearson's correlation coefficient.

(4) If $\rho_{XY} \neq 0$, $SEV(Y|X) \neq 0$, and if $SEV(Y|X) = 0$, $\rho_{XY} = 0$ naturally.

(5) If $\rho_{XY} = -1$ or 1, $SEV(Y|X) = 1$.

(6) $SEV(Y|X) \geq \rho_{XY}^2$.

And the following simple example gives a impression on how SEV is different from Pearson's correlation coefficient. We suppose X is a standard normal random variable. Let $Y = X^2$. Then $SEV(Y|X) = 1$, while $\rho_{XY} = 0$.

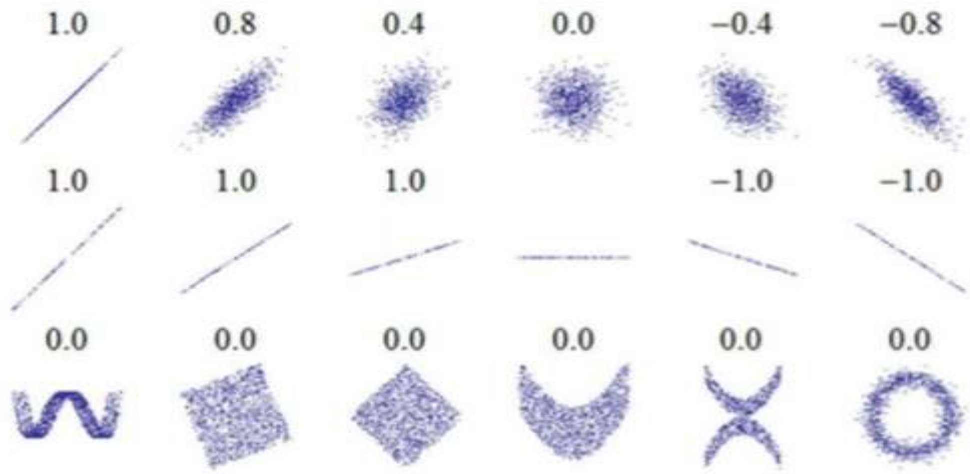

Figure 1: The Pearson's correlation coefficients of some kinds of point distribution

From the above example and figure we can see that for some no-linear dependence cases, Pearson's correlation coefficient is 0. So in terms of measuring the dependence strength, SEV is more powerful than Pearson's correlation coefficient. And the interpretation of SEV is simple and straightforward.

As a result, Pearson's correlation is a special case of SEV. And it is clear that SEV can be applicable in measuring both linear and nonlinear correlation and lead to more meaningful conclusions and improved judgement making.

Then we use SEV to replace the Pearson's correlation coefficient. As we all know, autocorrelation is a mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals. We suppose there is a time series $x_i, i = 1, 2, \dots n$, and the generalized autocorrelation at lag k should be defined as SEV_k,

$$SEV_k = SEV(x_{i+k}|x_i) = \frac{var(E(x_{i+k}|x_i))}{var(x_{i+k})}, k = 1,2, \dots, n-1 \quad (7)$$

Then we can get the generalized autocorrelation function SEV_k and apply it to test long-term memory. Just like ρ_k the decay of correlation is still the most important measure. So we propose a new judging method based on SEV_k. That is if the generalized autocorrelation of a time series satisfies

$$SEV_k \sim k^{-\alpha} G(k), k \to \infty \quad (8)$$

where $G(k)$ is any slowly varying function at infinity , then we think this time series has long-term memory.

III. The nonparametric estimation of SEV

We note that the calculation of SEV requires the variance of conditional expectation and the conditional variance. If the relationship between a response variable and an exploratory variable is known, $SEV(Y|X)$ can be calculated directly. But in practice, it is often un-known, and hence a statistical estimator of $SEV(Y|X)$ has to be provided for numerical evaluations.

Different from the common kernel estimation method, we present a new local smoothing estimation method to improve the accuracy and effectiveness of the nonparametric estimator in this section. Let $(Y_i, X_i), i = 1,2, \dots, n$ be a i.i.d random sample, $E(Y|X = x)$ is the condition expectation of

$$Y_i = f(X_i) + \varepsilon_i, i = 1, 2, \dots, \quad (9)$$

where $f(\cdot)$ is a unkonwn function, ε_i is a set of random error terms with mean of 0 and variance of σ^2. Use

Taylor's expansion, we have

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) \triangleq a + b(x - x_0) \tag{10}$$

where x_0 is a neighbor point of x. So we get the estimator

$$\left(\hat{E}(Y|X = x), \hat{b}\right) = \arg\min_{a,b} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)(Y_i - a - b(x - X_i))^2 \tag{11}$$

In the equation, $K(\cdot)$ is a given Kernel function. Let $K_h(x - X_i) = K((x - X_i)/h)$, then we have the weighted least squares estimation

$$\hat{E}(Y|X = x) = (1,0)(X^T W X)^{-1} X^T W Y \tag{12}$$

While

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, X = \begin{pmatrix} 1 & (x - X_1) \\ \vdots & \vdots \\ 1 & (x - X_n) \end{pmatrix}$$

and $W = diag(K_h(x - X_1, \dots, K_h(x - X_n))$. So, the estimator of $E(E(Y|X = x))^2$ is

$$\hat{E}(E(Y|X = x))^2 = \frac{1}{n} \sum_{i=1}^{n} (\hat{E}(E(Y|X = x)))^2 \triangleq \frac{1}{n} \sum_{i=1}^{n} (\hat{f}(X_i))^2 \tag{13}$$

Set $S_j = \sum_{i=1}^{n} K_h(x - X_i)(x - X_i)^j, j = 0,1,2,$ and $m_l = \sum_{i=1}^{n} K_h(x - X_i)(x - X_j)^l Y_i, l = 0,1.$ Then through simple algebraic calculations, we can obtain a set of numerical solutions of $\hat{f}(x)$, that is

$$\hat{f}(x) = (1,0) \begin{bmatrix} S_0 & S_1 \\ S_1 & S_0 \end{bmatrix} \begin{pmatrix} m_0 \\ m_1 \end{pmatrix} \frac{\sum_{i=1}^{n} K_n(x - X_i)\left(1 - S_1 S_2^{-1}(x - X_i)\right) Y_i}{\sum_{i=1}^{n} K_n(x - X_i)\left(1 - S_1 S_2^{-1}(x - X_i)\right)} \tag{14}$$

Then replace the mean and variance of the population with its sample form, the local smoothing estimator of SEV is

$$\widehat{SEV} = \frac{n^{-1} \sum_{i=1}^{n} \hat{f}(X_i)^2 - \bar{Y}^2}{S_Y^2} \tag{15}$$

With the above nonparametric estimation method, the result of SEV_k of time series can be obtained through R software.

IV. Empirical analysis

In this section, firstly we use two special counter examples to illustrate the limitations of Pearson's correlation coefficient in measuring the long range dependence and prove generalized measures of correlation SEV can solve these problems. So it is necessary and reasonable for us to introduce generalized correlation coefficient.

We consider the following two special time series A and B.

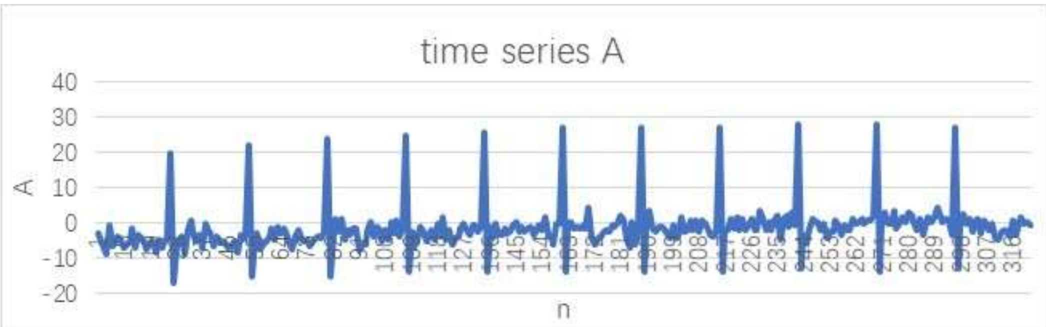

Figure 2: Special time series A

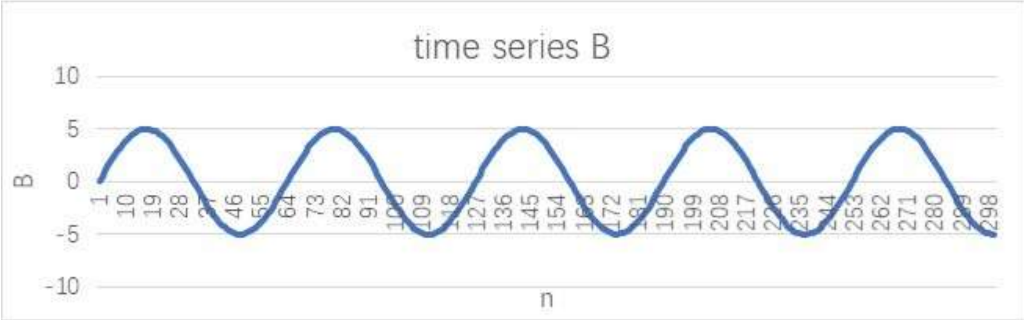

Figure 3: Special time series B

We first observe these two situations from time series images. The time series A is similar to ECG, with almost same regular jumps. Obviously, it is self-similar and shows a general rising trend. According to the existing research on ECG, the Hurst index of normal ECG is always greater than 0.5 and close to 1. And time series B seems like a trigonometric function, it also has self-similarity. It's easy to say that these two special time series are both long range dependent.

Then we apply the R/S analysis method to further prove long-term memory for time series A and B. This method can calculate Hurst index—H, if $H = 0.5$, the time series is an independent process without long-term memory; if $0.5 < H \leq 1$, it is a long range dependent process; and if $0 \leq H < 0.5$, it has anti-persistence.

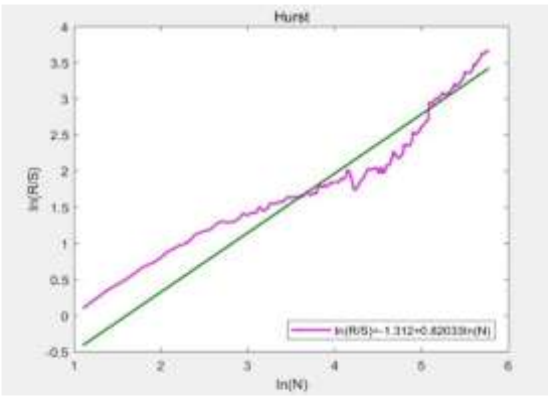

Figure 4: Hurst index of time series A

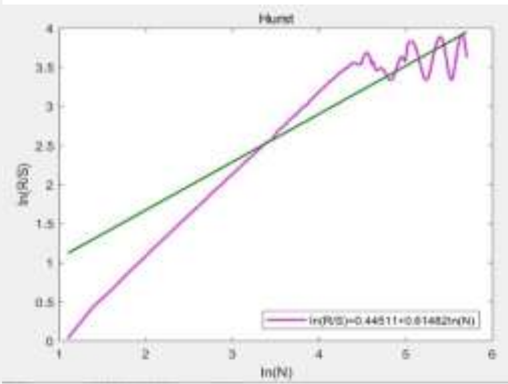

Figure 5: Hurst index of time series B

As a result, $H_a = 0.8208$, $H_b = 0.6148$. It's easy to say they both show long-term memory. But can we get this conclusion from classical autocorrelation function (ACF)?

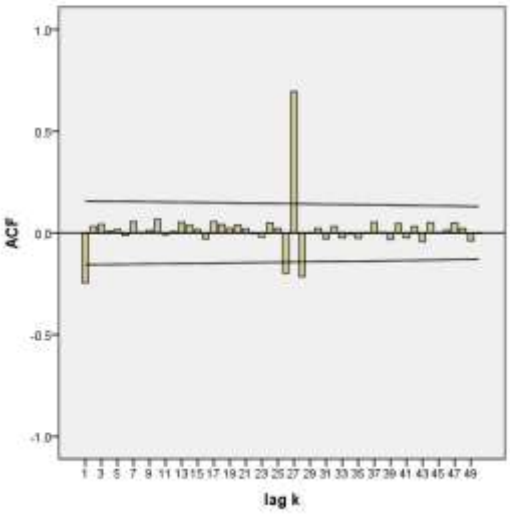

Figure 6: the classical ACF of time series A

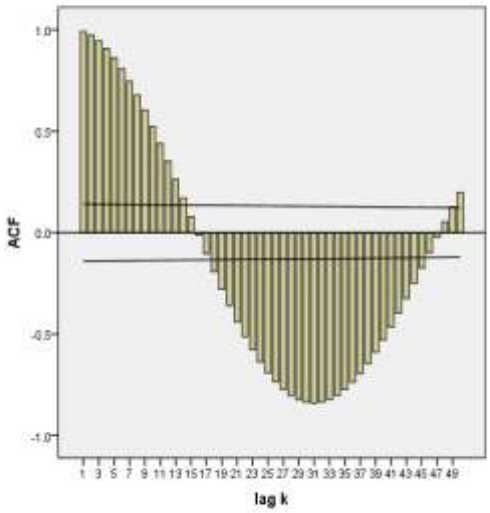

Figure 7: the classical ACF of time series B

According to Figure 6, except at the time points of jumps, the autocorrelation coefficients of time series A are always small | almost 0. So we can not see it decaying at a slow hyperbolic rate. There is no characteristic can tell if the time series A is a long range dependence process.

Then we use SEV_k to replace ρ_k, and obtain the generalized autocorrelation function in Figure 8. The result shows that $SEV_k, k = 1,2,\ldots,50$, vary in an acceptable range and it decrease slowly in general. In conclusion, the time series A has long-term correlation.

Next, we observe these two kinds of ACF of time series B. From Figure 7, there exists long-term memory according to the usual definition Eq(2). However, for some lag times, $\rho_k = 0$. It is not in accordance with the fact that time series B and any lagged version of itself are highly correlated. In Figure 9, the generalized ACF can describe and explain the real characteristics of time series B better, as SEV_k is close to 1 at any k.

Figure 8: the generalized ACF of time series A

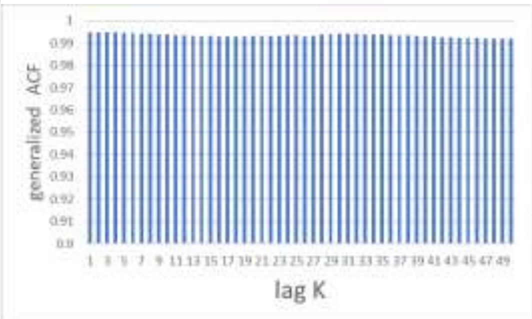

Figure 9: the generalized ACF of time series B

Next, we observe these two kinds of ACF of time series B. From Figure 7, there exists long-term memory according to the usual definition Eq(2). However, for some lag times, $\rho_k = 0$. It is not in accordance with the fact that time series B and any lagged version of itself are highly correlated. In Figure 9, the generalized ACF can describe and explain the real characteristics of time series B better, as SEV_k is close to 1 at any k.

To summarize, the generalized ACF which considering nonlinear correlation can solve the problem that the classical one can not determine long-term memory in some cases. Similarly, it can also help us obtain the result about time series autocorrelation that is more closely fit the actual state.

Testing long-term memory is an important subject in finance. So we apply the test statistics to the practical financial data, the daily returns of B share index of Shenzhen Stock Market from 2nd January, 2018 to 28th December, 2018.

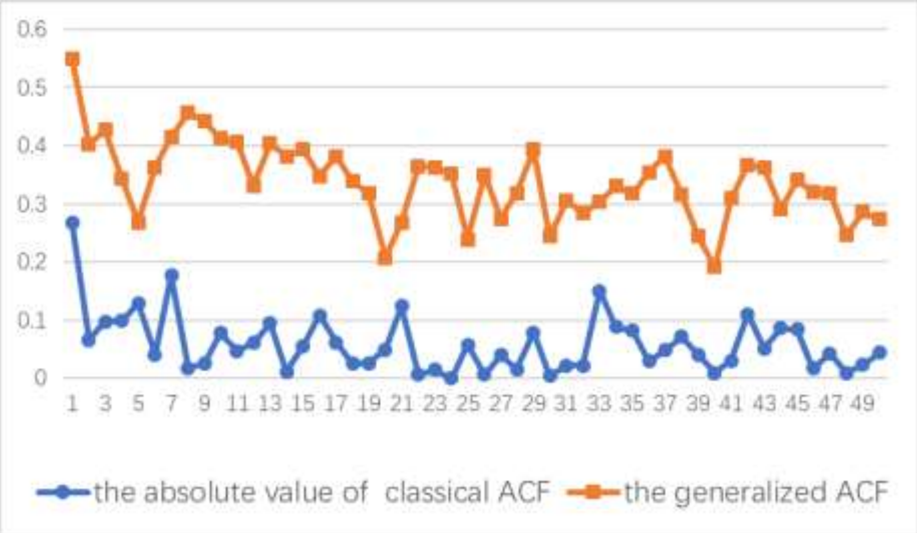

Figure 10: the two kinds of ACFs of real financial data

We use two different methods to analyze the real data and get the corresponding ACF in Figure 10. The presentation states that SEV_k is always larger than ρ_k at the same lag time and this two kinds of ACFs are similar in fluctuation. We can propose that this time series is not a long range dependence process because of the autocorrelation coefficient decaying quickly.

The real example helps us to examine the validity and the generality of the new method for testing long-term memory. In general, the result of generalized ACF is similar to the classical one so it won't affect our judgment or lead to erroneous conclusions.

V. Conclusions

Long-term memory (long range dependence) has been defined in time domain in terms of autocorrelation power law decay. But the nonlinear correlation is not considered in autocorrelation function ρ_k. This paper focuses on this problem and proposes a new method for testing the existence of long-term memory to improve decision making and to give a new criterion to arbitrage to no arbitrage for a financial market. We find a generalized measure of correlation-SEV and introduce a new local smoothing estimation method to improve the accuracy and effectiveness of the nonparametric estimator for it. And the results of empirical analysis show the generalized autocorrelation function SEV_k has superior performance in judgment making and characteristic presenting. In other words, one can induce the related the arbitrage principle via a new model about long-term memory based on the SEV structure frameworks.

References

[1] D. E. Allen and V. Hooper, Generalized Correlation Measures of Causality and Fore-casts of the VIX using Non-linear Models. *Sustainability* 2018, *10(8:2695)*, 1-15.

[2] M. Chen, Y. Lian, Z. Chen & Z. Zhang, Sure Explained Variability and Independence Screening. *Journal of Nonparametric Statistics* 2017, *29:4*, 849-883.

[3] S. R. Bentes, R. Menezes and D. A. Mendes, Long memory and volatility clustering: Is the empirical evidence consistent across stock markets? *Physica A: Statistical mechanics and its applications* 2008, *387(15)*, 3826-3830.

[4] Z. Ding and C. W. J. Granger, Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics* 1996, *73*, 61-77.

[5] J. Elder and A. Serletis, Long memory in energy futures prices. *Review of Financial Economics* 2008, *17(2)*, 146-155.

[6] J. Helson and Y. Sarason, Past and future. *Mathematica Scandanavia* 1967, *21*, 5-16.

[7] H. E. Hurst, Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineering* 1951, *116*, 770-799.

[8] W. Lo, Long term memory in stock market prices. *Econometrica* 1991, *59*, 1279-1313.

[9] B. B. Mandelbro and J. Wallis, Noah, Joseph, and operational hydrology. *Water Re-sources Research* 1968, *4*, 909-918.

[10] A. I. McLeod and K. W. Hipel, Pereservation of the rescaled adjusted range,1: reassessment of the Hurst phenomenon *Water Resources Research* 1978, *14*, 491-508.

[11] S. I. Resnick, Extreme values, regular variation and point processes. *Springer-Verlag, New York, NY* 1987

[12] K. Shimotsu and P. C. B. Phillips, Exact local Whittle estimation of fractional integration. *The Annals of Statistics: An Official Journal of the Institute of Mathematical Statistics* 2005, *33(4)*, 1890-1933.

[13] P. Sibbertsen, Long memory in volatilities of German stock returns. *Empirical Economics* 2004, *29(3)*, 477-488.

[14] F. B. Sowell, Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics* 1992, *53*, 165-188.

[15] S. Zheng, N. Z. Shi and Z. Zhang, Generalized Measures of Correlation for Asymmetry, Nonlinearity, and Beyond. *Journal of the American Statistical Association* 2012, *107*,1239-1252.