Stock Market Volatility, Recurrence Interval and Memory Effect

Zhang Chi
School of Economics & Management
Southeast University
Nanjing, P.R. China
anhuizhc@outlook.com

***Abstract*—The stock market volatility in 2015 has caused serious damage to investors and the market has also suffered severely and immensely. During the last year, large volatility and extreme events were increasingly frequent than before, which made great theoretical and practical significance to gain a deep understanding of extreme value statistics of the volatility. Extreme events also bring huge risk to financial market, therefore the risk prevention, estimation and prediction are of necessity.The research uses 1-min high-frequency datasets of Shanghai 50 Stock Index Futuresin 2015. The data comes from Tongdaxin Database. The paper is the first one to study Shanghai 50 Stock Index Futures and a relationship between recurrence interval and risk estimation has been constructed. We find the recurrence interval of stock volatility can be fitted with stretched exponential function and the recurrence interval decreases when the threshold decreases. Then we demonstrate the existence of short-term and long-term correlations in recurrence intervals. We further construct a hazard function and define a loss probability to evaluate risk and find a crossover point in the loss probability plot. The study would enable one to improve risk estimation andthere are some shortcomings and need to be perfect in the future.**

***Key words*: extreme events; recurrence interval; probability distribution; memory effect; risk estimation**

I. INTRODUCTION

The volatility pattern of financial market has been the focus of government and investors. Large volatility in 2015 Chinese stock market made market and government wonder: will it happen again or not? As a result, when volatility happens in stock market, the time interval between volatility became market and government's concern. Is there memory effect of time interval between volatility, and how will it affect volatility behavior in the short and long term? Can the time interval between volatility now occurred be the evidence to predict the time and extent of next-time volatility? Can we measure the risk by time interval? These questions have implications for understanding the volatility in China stock market. Hence, in this paper, we use Shanghai 50 Stock Index Futures to investigate the volatility pattern of stock market.

Currently, the frequency of extreme events in financial market is higher than normal distribution[1], investors can hardly maintain perfect hedge. And because the income and loss caused by big volatility are higher than small volatility, investors have more interests on extreme events than small volatility. Hence, to measure big volatility is not only useful to understand present market, but also can be a help to predict occurrence probability of extreme events in future. However, extreme events do not happen a lot, which make it hard to describe its profile in statistic. In order to study the behavior of financial market form a brand new perspective, researchers have created a new interdisciplinary research field, Econophysics, which employs theories and methods from physics, like uncertainty, random process, nonlinear kinetics and etc. to solve economic problems[2]. In this paper, we use one method from Econophysics, Recurrence interval analysis (RIA), to study behavior of Shanghai 50 Stock Index Futures.

Recurrence interval is the time interval of two adjacent events beyond or under one certain threshold q. And recurrence interval analysis is a statistical method that uses the events occurred within a period in past to estimate its probability to happen again in future. RIA is more applied to risk analysis, assuming occurrence probability of future events is constant in time and independent to past events. RIA is widely used to analyze extreme events in time series data in many fields, like climate[3], earthquake activities[4], heartbeat monitoring[5] and financial volatility[6]. Therefore, RIA can help predict catastrophic events, such as floods, droughts and stock crisis. By discovering the scale behavior of events of different thresholds, it can also effectively solve the problem of insufficient data[7]. At present, the recurrence interval between volatility in financial market has been widely studied[8].

In this paper, we use 1-minute high frequency data of Shanghai 50 Stock Index Futures to study the recurrence pattern between volatility. Though there have been many detailed work on global currency market, U.S. stock market[9] and China's stock market[10], this is the first paper to investigate recurrence interval between volatility of Shanghai 50 index futures. Then, most studies only focus on the statistical characteristic of recurrence interval between extreme events, this paper brings recurrence interval analysis into risk estimation, establishing the function relationship between average recurrence interval and VaR, and explains why traditional methods would overestimate or underestimate VaR when assessing complex financial system.

The rest of this paper is organized as follows: Section 2 describes data and sample selection, and presents basic statistics. Section 3 presents empirical study, including distribution function, scaling properties, memory effect and risk estimation. Section 4 is the conclusion.

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

II. METHOD AND DATA

Shanghai 50 Stock Index Futures, launched based on Shanghai 50 Stock Index, was created by China Financial Futures Exchange on April 16, 2015. The data here are obtained from Tongdaxin Database. The sample period is from April 16th, 2015 through December 31st, 2015. Excluding weekends and holidays for which trading is closed, we end up with 47,791 price observations. The return of time serious are calculated by taking differences of the logarithms of price as

$$Zha\ n\ gCiZha\ S\ gCiZh\ S\ c\ ha \quad (1)$$

Where p(t) is the closing price of the tth time and Δt=1 because the data is 1-min frequent. The statistics of logarithmic returns of Shanghai 50 Stock Index Futures is summarized in Table 1

Currently, most researches are more interested in using asset price to study the properties and dynamical mechanisms of financial market, e.g., price discovery[11], volatility transmission[12] and risk estimation[13]. From fgure1, it is clear to see there exists clusters in the volatility, i.e., small fluctuations tend to follow small fluctuations, and large fluctuations tend to follow large fluctuations, which indicates long-term memory effect[14]. During volatile time periods when large volatilities cluster, the recurrence intervals are small and dense. In contrast, during calm periods with small volatility, the recurrence interval are large and sparse. While what fascinates us is whether this can be a new method to describe volatility behavior and to estimate the time interval between large fluctuations, which can be a help to improve risk estimation in complex market full of financial derivatives? E.g. when will next large fluctuation occur after a large fluctuation is being observed?

Table I. Statistics of the logarithmic returns

Average	-4.83e-06
Maximum	0.084
Minimum	-0.0794
Standard deviations	0.0015
Skewness	-0.850
Kurtosis	548.331
Nobs	47790

III. EMPIRICAL RESEARCH

Consider o as the recurrence interval when threshold is l , the overwhelming consensus is that the recurrence intervals of financial volatility are distributed as a stretched exponential

$$f_E Zoa\ n\ ns^{\&ZMa^t} \quad (2)$$

(2) means that given a threshold l , the distribution of recurrence interval o is $f_E Zoa$, which is supported by a handful of empirical evidence using daily or high-frequency data[15]. Not only in financial field, stretched exponential function has also been adopted in other natural and social sciences. And the scaling function can be well-approximated by the stretched exponential function $uZUa\ n\ \alpha os^{\&Zr\ gNa^t}$. In Chinese stock market it has also been found similar scaling memory property of the recurrence interval between volatilities, and the correlation exponents j are different for different threshold l .However, different researches receive different results, some stocks possess scaling behavior while others exhibit multiscaling behaviors in a same market.

3.1. Probability Density Function

To examine the memory behaviors and to improve risk estimation of Shanghai 50 Stock Index Futures, we first analyze the probability density function of recurrence interval between volatility [16]. In order to apply the recurrence interval analysis, we need to normalize the price returns of each series by dividing its standard deviation:

$$,\ ZPa\ n\ \frac{.\ ZRa}{z@\ ZRa^k\ \&@^k.\ ZRaI^{TOk}} \quad (3)$$

Where $zD\ Zha^p\ S\ D^p\ ZhaI^{bOp}$ is the standard deviation of Zha , for normalized time series , ZPa , we measure the recurrence internals by introducing a threshold l . The recurrence interval τ is defined as the waiting time between successive volatilities exceeding a given threshold (l d L2or l 0 1). Assume that the volatility at time t is lower than l : , $ZPa\ 0\ 1$. Then the recurrence is

$$oZha\ n\ 5w:\ ?h'\ b\ 3\ h\ -'a\ 0\ 1\ 'h' \quad 3 \quad (4)$$

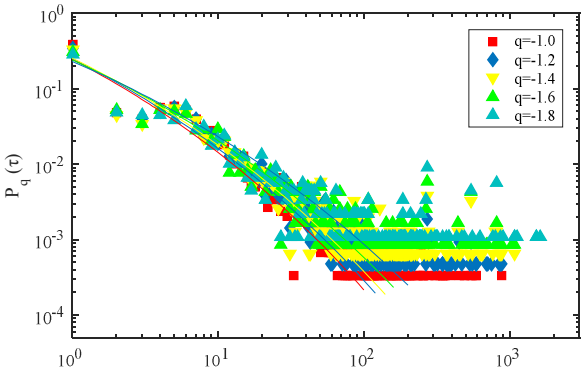


Fig 1. Empirical probability distribution of recurrence intervals between returns with different thresholds

From figure 1, it can be seen that for large fluctuations (the value of -l - is high), $f_E Zoa$ decays slowly, indicating that with -l - rises, recurrence intervals will be longer, in agreement with the fact that large fluctuations have more long intervals and less short intervals than small fluctuations. When threshold l is fixed,$f_E Zoa$ decreases with increasing recurrence intervalo. This implies that when last fluctuation exceeding -l -occurredt units of time ago, if $c\ h_b\ 0\ c\ h_p$, the probability $f_E Zoa$ of next fluctuation exceeding -l - to occur $c\ h_b$ later is larger thanc h_p .

The result suggests that if investors suffer a huge loss in last large fluctuation, then they should be aware that the probability that next large fluctuation occurs within a shorter time interval is larger than that of a longer time interval. On the other hand, when recurrence interval is fixed, $f_E Zoa$ increases with increasing -l -. If an extreme event occurs,the probability that the next extreme event would occur will be higher for a larger threshold -l -.

The PDF of recurrence intervals of Shanghai 50 Stock Index Futures can be fitted with stretched exponential function

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

3! " αovw&Zr ŷNat . The parameters calculated by maximum likelihood estimation are shown in Table 2. In Table 2, we can find that the coefficient of exponent γ decreases by increasing γ. By the observation on Table 2 and Figure 2, we wonder is there scaling behavior between these PDFs, i.e. are these five PDFs can be described using only one function by some kind of transformation?

Table II. Estimates of the coefficients of stretched exponential functions

q			
-1.0	6.402	2813.502	0.168
-1.2	4.066	1983.807	0.166
-1.4	2.649	1452.430	0.164
-1.6	1.523	1109.364	0.160
-1.8	0.890	860.8695	0.156

To examine that, we use Yamasaki et al method,

$$u_E Zo Oosx n f_E Zoaov \quad (5)$$

Where ovis the mean recurrence interval that depends on the threshold l , and $f_E Zoaov$is scaled PDFs, as depicted in Figure 2.

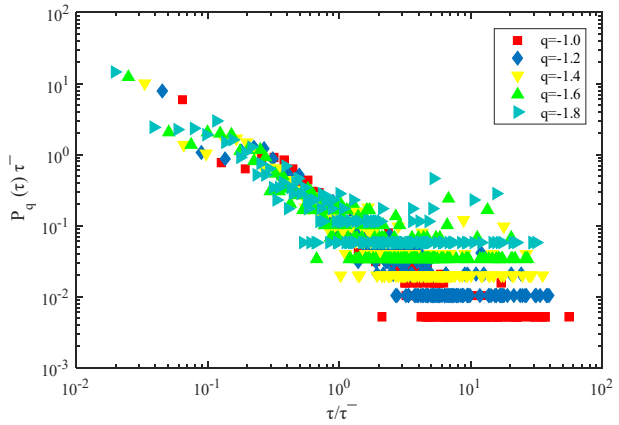


Fig 2. Scaled probability distributions of recurrence intervals for different thresholds

From figure 4, we do not observe that the curves of $u_E Zo Oosx$ have a better collapsing than the original curves of $f_E Zoa$, so it is not clear that if there is a scaling behavior and we will further test it. Theoretically, if $u_E Zo Oosx$ is independent of l , so when $l_]$ ($l_)$, there is $u_{E_3} Zo Oosx n u_{E_4} Zo Oosx$. In this paper, we employ Kolmogorov-Smirnov test to check if $u_{E_3} n u_{E_4}$.

KS test is designed to test the hypothesis that the distribution of data is equal to a particular distribution by comparing their cumulative distribution functions, which also can be applied to test if two samples have the same probability distribution. Assume 6_{E_3} the cumulative probability function of u_{E_3}, we calculate the KS value by comparing the two CDFs in the overlapping region:

$$B7 n 589 ' V6_{E_3} S 6_{E_4} WVl_] (l_) \quad (6)$$

When KS value is less than a critical Value CV, the hypothesis is accepted. The critical value is

$$\Delta = n K_\alpha \tau \overline{Z \quad \gamma \quad CaY'C} \quad (7)$$

Where ′ and C are the numbers of recurrence intervals for threshold $l_]$ and $l_)$, and at significance level %n /‾ the K_α n >“" .The null hypothesis is that $u_{E_3} n u_{E_4}$ and it is rejected if JX d „ 2at the significance level of /‾ . We have performed KS tests for each pair of samples' o_{E_3}'o_{E_4}W The results are presented in Table 3， we can see JX d „ in some cases. It means that there is no scaling in the distribution of recurrence intervals for different thresholdsl .

Table III. KS test and critical value CV at significance level of 5%

q		CV	KS
-1.0	-1.2	0.0474	0.0385
-1.0	-1.4	0.0816	0.0424
-1.0	-1.6	0.0743	0.0466
-1.0	-1.8	0.0902	0.0514
-1.2	-1.4	0.0342	0.0453
-1.2	-1.6	0.0269	0.0492
-1.2	-1.8	0.0428	0.0538
-1.4	-1.6	0.0073	0.0523
-1.4	-1.8	0.0085	0.0567
-1.6	-1.8	0.0159	0.0599

3.2. Memory effect

Previous works have shown that the recurrence intervals of stock and stock index volatility possess both short and long-term memory effect. A question comes out here that for Shanghai 50 Stock Index Futures, is there memory effect of volatility recurrence intervals.

3.2.1.Short-Term Memory

To investigate short-term correlations in the recurrence intervals, we use Yamasaki's method. We begin by calculating and comparing conditional probability density functions $f_E Zo-o a$, which is the probability of finding a recurrence interval o immediately after the recurrence interval o . If there is no short-term correlations, $f_E Zo-o a$ is independent of o . However, due to the size of interval sample is not sufficiently large, it is impossible to calculate $f_E Zo-o a$ for a single value o . Thus we adopt a method based on the idea of coarse graining and analyze for values of o in a certain intervals.

For a given threshold *q*, the set of all the recurrence interval is partitioned into four non-overlapping subsets, n b p , where]) n 'w(. In the partitioning procedure, all recurrence intervals in are sorted in an increasing order and then assigned it turn into the subsets with the same size. Hence, the 1/4 smallest recurrence intervals are put in the first subset b while contains the largest quarter of *T*. We estimate the conditional probability density functions$f_E Zo-]a$ n $f_E Zo-o]a$, if there is no short-term correlations, we should find $f_E Zo-]a$ n $f_E Zo-]a'w($.

Figure 3 shows the results of $f_E Zo-o aov$as a function of oOov for o in the smallest subset b (filled symbols) and largest subset (open symbols), obviously $f_E Zo-]a$ ($f_E Zo-]a$. We also find that, $f_E Zo-o bais$ bigger than $f_E Zo-o a$ for

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

small oOoy while for big oOoy f_EZo-o $_b$a is smaller than f_EZo-o a. The fact that small o tends to follow small o , and big o tends to follow big o indicates that short-term correlation exists in the recurrence intervals. These results mean that if the time interval between two consecutive extreme events is short, then the next extreme event would occur with a larger probability within a smaller time interval, i.e. when there is a sharp decline in market, it may continue to fall in a high probability.

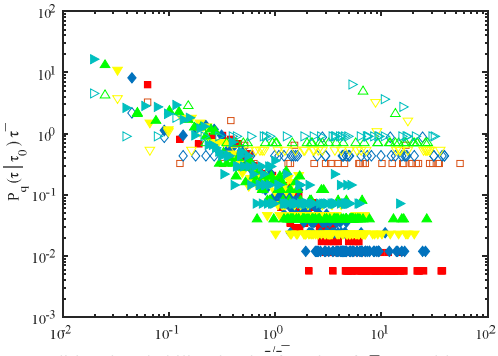


Fig 3. Conditional probability density functionsf_EZo-o awitho $_b$ (filled symbols) and o (open symbols)

3.2.2.Long-Term Memory

The volatility clusters in Figure 1 indicate that there are long-term correlations in the time series. To test it we adopt DFA method and the results are shown in Figure 4.

DFA, known as Detrended Fluctuation Analysis, was invented by Peng [17] (1994) to investigate the long-range dependence in DNA nucleotide sequences. The properties of DFA have been extensively studied and is a powerful method to characterize the long-term correlations in time series. DFA computes the detrended fluctuation 6Z a of a time sequence within a window of time units and determined the exponent *H* from the form 6Z a . If H>0.5, the sequence is long-term correlated while the sequence is uncorrelated for H=0.5.

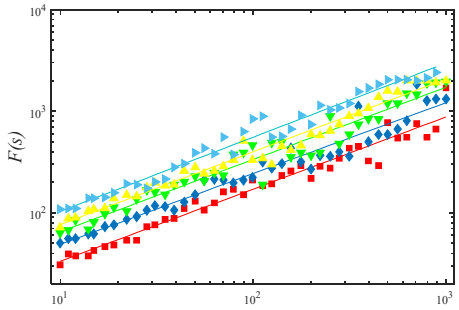


Fig 4. Detrended fluctuation function 6Z a of the recurrence intervals(top to bottom:

l n S>“ ’l n S>“ ’l n S>“ ’l n S>“ ’l n S>“1)

In figure 4, from top to bottom, each fitted line is:

6Z a n > “>" “ (9)

6Z a n > “”> “ $_b$ (10)

6Z a n > “ bpp (11)

6Z a n “ ” “ (12)

6Z a n ""“ " “ bb (13)

It can be seen that the Hurst exponent of each line is greater than 0.5, suggesting that long-term correlations exist in the recurrence intervals. This is in agreement with previous studies on the recurrence intervals in finance market and confirms the fact that long-term memory effect on recurrence intervals exists in Shanghai 50 Stock Index Futures.

IV. Risk Estimation

Consider the fact that t units of time have passed since last large volatility greater than -l -, what fascinates us is the probability that next large volatility greater than -l - will occur within c h units of time. Mathematically, the hazard probability can be expressed as follows:

$$_E\text{Zc P-ha n} \frac{^{c} \text{Zea e}}{\text{Zea e}} \quad (14)$$

Since each distribution f_EZoahas been fitted with a stretched exponential, we can calculate the theoretical value of f_EZoawith the parameters given in Fig 3. Besides, in order to determine the $_E$Zc P-ha empirically, we can rewrite (14) in the following way:

$$_E\text{Zc P-ha n} \frac{\text{RRe R cR}}{\text{e R R}} \quad (15)$$

Where “ KL Ch$_E$ d h ” is the number of recurrence intervals greater than h units of time and “KL Ch 0 o$_E$ hγ c ha” is the number of recurrence intervals greater than h and not greater than hγ c h for a given l .

Color symbols shown in Figure 7 are the hazard function $_E$Zc Pn >-ha obtained numerically by using (15) and the curves are the theoretical values of $_E$Zc Pn >-ha by using (14). We can observe that they are fitted very nicely and the discrepancy between empirical and theoretical curves decreases when h increases. Furthermore, $_E$Zc h n >-ha decreases with increasing h , confirming that recurrence intervals exhibit clustering behaviors and long-term memory between volatilities. We can also see that for small h, when -l $_b$- 0 -l $_p$-, there exists:

$$_{E_T}\text{Zc h-ha d } _{E_k}\text{Zc h-ha} \quad (16)$$

It means when -l - increases, the hazard probability will decrease, andsimultaneouslythere are less recurrence intervals. In general, $_E$Zc P-ha is a monotonously increasingly function of c h.

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

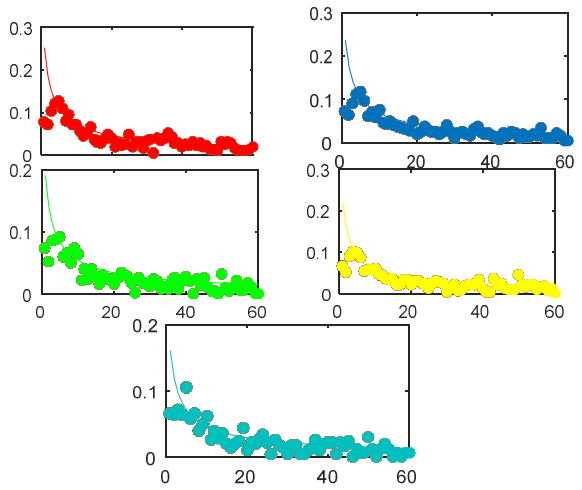


Fig 5. Theoretical (curves) and empirical (color symbols) value of $_E$Zc h n >-ha(x-rays are t, y-rays are values of $_E$Zc Pn >-Pafor l n S>"l 'l n S>" 'l n S>" 'l n S>"" 'l n S>" from left to right)

In the end, we use loss probability density function in recurrence interval analysis to estimate the value at risk (VaR), which defines therisk at loss l :

$$\& \quad f Z, a \quad , n \; f \qquad (17)$$

Where f Z, a is probabilitydensity function of normalized series , Zha , and f is the loss probability. The mean recurrence interval is defined as o$_E$ n $\frac{b}{}$ $\frac{e}{]\ b}$ o$_{E']}$, and $_E$ is the number of intervals which fall below the threshold l , so $\frac{e}{]\ b}$ o$_{E']}$ is approximately equal to the total number of returns and $_E$ γ > is the number of returns below threshold l . Then we can get a relation between mean recurrence interval and VaR:

$$\frac{b}{y} \; n \quad \frac{E}{\&} \quad f Z, a \quad , n \; \frac{2 \; .2 \; ZRa2 \qquad \Sigma E}{R \quad RM1 \qquad 2 \; .2 \; ZRa} \qquad (18)$$

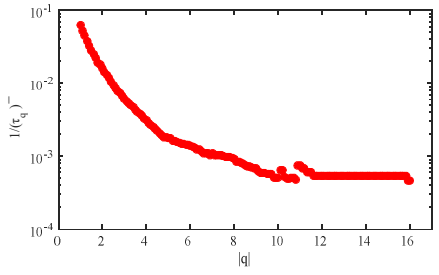


Fig 6. The reciprocal of mean recurrence interval>Oo$_E$ as a function of absolute thresholdl (l n S>"ASS>)

If investors want to know the risk level at 1% probability of loss they can find the l for $\frac{b}{y}$ n > , which is the VaR one is looking for.

V. CONCLUSIONS

The paper uses recurrence interval analysis to investigate the property of recurrence intervals of volatilities for different thresholds using 1-min high frequency data, attempting to understand the behaviors of large volatilities.. We use RIA to analyze the characteristics of volatility in Shanghai 50 Stock Index Futures and to check the short-term and long-term memory and last to estimate the VaR.

First we observe the distribution functions of recurrence and find the results that for different thresholds, the probability density function fit the stretched exponential function. We found that there is no scaling behavior in the distributions for different thresholds after the recurrence intervals are scaled with the mean recurrence interval, which has been verified by KS test. Then we use conditional probability density function and detrended fluctuation analysis to investigate the short-term and long-term memory separately, indicating that the clusters of recurrence intervals of volatilities are not only caused by present condition but also long-term correlations. Last, we apply recurrence interval analysis to evaluate the risk for Shanghai 50 Stock Index Futures, which provide relatively accurate estimates of hazard and forge a link between loss possibility and VaR.

This paper, as the first study to explore the risk of Shanghai 50 Stock Index Futures, still have deficiencies, which can be perfected by subsequent researches to help investors to evaluate risk.

REFERENCES

[1]. Sornette, Didier. "Why stock markets crash : critical events in complex financial systems." Physics Today vol.57,2004, pp. 78-79.

[2]. Mantegna, Rosario N., H. Eugene Stanley, and Neil A. Chriss. "An Introduction to Econophysics: Correlations and Complexity in Finance." Physics Today vol. 53,2007,pp.70.

[3]. Bunde, Armin, et al. "Return intervals of rare events in records with long-term persistence." Physica A-statistical Mechanics and Its Applications vol.342,2004,pp.308-314.

[4]. Livina, V. N., et al. "Recurrence Intervals Between Earthquakes Strongly Depend on History." Physica A-statistical Mechanics and Its Applications ,2005,pp.591-595.

[5]. Bogachev, Mikhail I., et al. "Statistics of return intervals between long heartbeat intervals and their usability for online prediction of disorders." New Journal of Physics vol.11,2009.

[6]. Suo, Yuanyuan, Donghua Wang, and S. X. Li. "Risk estimation of CSI 300 index spot and futures in China from a new perspective." Economic Modelling ,2015,pp.344-353.

[7]. Chattopadhyay, Amit K., and Nigel J. Burroughs. "Close contact fluctuations : the seeding of signalling domains in the immunological synapse." EPL vol.77,2007.

[8]. Ren, Fei, and Weixing Zhou. "Recurrence interval analysis of high-frequency financial returns and its application to risk estimation." New Journal of Physics vol.12 ,2009.

[9]. Wang, Fengzhong, et al. "Scaling and memory of intraday volatility return intervals in stock market." Physical Review E vol.73 ,2005.

[10]. Ren, Fei, and Weixing Zhou. "Multiscaling behavior in the volatility return intervals of Chinese indices." EPL vol. 84 ,2008.

[11]. Tokat, Ekin, and Hakki Arda Tokat. "Shock and Volatility Transmission in the Futures and Spot Markets: Evidence from

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

Turkish Markets." Emerging Markets Finance and Trade vol.46 ,2010,pp.92-104.
[12]. Chulia, Helena, and Hipolit Torro. "The Economic Value of Volatility Transmission Between the Stock and Bond Markets." Journal of Futures Markets vol.28 ,2007,pp. 1066-1094.
[13]. Yang, Jian, Yinggang Zhou, and Zihui Yang. "Intraday price discovery and volatility transmission in stock index and stock index futures markets: Evidence from China." Journal of Futures Markets vol.32 ,2010,pp. 99-121.
[14]. Tabak, Benjamin M., and Daniel O. Cajueiro. "Are the crude oil markets becoming weakly efficient over time? A test for time-varying long-range dependence in prices and volatility." Energy Economics vol.29 ,2007,pp. 28-36.
[15]. Jung, Woosung, et al. "Volatility return intervals analysis of the Japanese market." European Physical Journal vol. 62 ,2007,pp. 113-119.
[16]. Bogachev, Mikhail I., and Armin Bunde. "Improved risk estimation in multifractal records: Application to the value at risk in finance." Physical Review E vol.80 ,2009.
[17]. Peng, Chengbin, et al. "Mosaic organization of DNA nucleotides." Physical Review E vol.49 ,1994,pp. 1685-1689.
[18]. Bogachev, Mikhail I., Jan F. Eichner, and Armin Bunde. "Effect of Nonlinear Correlations on the Statistics of Return Intervals in Multifractal Data Sets." Physical Review Le

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)