RISK IDENTIFICATION ANALYSIS OF STATISTIC DATA FOR BUILDING THE INVESTMENT FORECAST WITH THE HELP OF BROWNIAN MOTION MODEL

Kyzemin Oleksandr, Irina Gurina

Abstract: *Shocks and jumps of different funds' functions bring extreme negative values of great consequences to investors. There is always a factor of uncertainty in any economic situation, and in order to make the right investment decisions, or to choose the right business strategy, we require some form of workable hypothesis (that takes into account uncertainty and randomness) to base our decisions upon. Although both risk and upward potential are related to uncertainty of future events, risks usually play a more dominant role in investment decisions since investors are risk averse. At most cases detection of risk is one of the most important parts of a financial analysis.*

This article will examine ways to measure and manage risk in making investment decisions. Here we determine Brownian motion as the estimation method of the investment risks and show the specific features of the given method during the statistic data analyses.

Keywords*: Brownian motion, risk, function, financial analysis, investment market, non-recurrent cycles, deviations, noises, inaccuracies, strategy, knowledge, the Hurst analyses, correlation, entropy.*

ACM Classification Keywords: *G.3 Probability and statistics - Stochastic processes*

Introduction

Every risk strategy of investment business has the probability of heavy material losses. Misinformation, a lack of knowledge about the investment, or market pressures are influencing investors thinking in some way and leading them to wrong decisions. Successful implementation of the risk strategy involves overcoming these factors and rationally examining the proposed investment. The main strategy for reducing investment risk is diversification [Granot, Soši´, 2003]. By spreading investments among several asset classes such as stocks, bonds and cash, the firm effectively reduces its overall investment risk because there is almost always at least one asset class that is doing well [Plambeck, Taylor, 2003]. For example, if stocks are down, then interest rates may be up, which means that bonds could be a good investment. Another way to lower investment risk is by using derivatives such as options and futures. Derivatives allow the firm to hedge its investments and increase its overall return with limited risk and can protect firm's portfolio from possible downside exposure. One of the easiest ways to lower equity investment risk is to simply give them time to grow. Long-term investors are much more likely to see their investments grow over time than market timers or those who invest for five years or less. However, all these metods need qualitative mathematical analysis to choose the right strategy [Salvatore, 2003].

The building of the classical lineal models cannot describe the saturation, resonance jumps and vibrations. The exceptions to this rule are only steady growth and equally steady decrease [Kronover, 2000]. The Hurst analyses showed that the market had the long term memory of the long term investment horizons. We can present the behavior of the market function as the filter of the informational materials, which are shown in its structure of performance [Guasoni, 2006].

According to the theory of chaos errors, deviations, noises, inaccuracies are more fundamental and ontological, rather than the processes described by classical science.

According to the sinergetic point of view, we can talk about the investment market as an open dynamic, dissipative system, the system that has the possibility of self-development, self-sustaining and self-control [Markov, 2009].

The behavior of the sale value of certain goods or funds in the time interval is characterized by uncertainty, which is caused by entropy effects of independent market participants. The investment market has non-lineal development because its structures are in a constant state of struggle with arising fluctuations – deviations from some of the averages indicators. For example, the beginning of new centers of attraction is nothing like a strong fluctuation, which arose under the influence of environment (creation of new businesses, economic crisis, etc.) [Markov, 2008]. This behavior of the function can be described as the Brownian motion of the elements, where the change of the point's parameters is caused by other points' functions. At each time this function experiences unbalanced influence from other functions and as the result the movement of the point during the time interval has the stochastic behavior [Peters, 2000]. The fractal methods such as Brownian motion method are used in case if a real object cannot be presented in the classical view. It means that we work with nonlinear constraints and nondeterministic (random) kind of data. Nonlinearity at the worldview sense means multiversion of routes development, a choice of alternative routes and certain rate of evolution and the irreversibility of evolutionary processes [Peters, 2000].

Brownian motion (named after Robert Brown, who first observed the motion in 1827, when he examined pollen grains in water), or pedesis (from Greek: πήδησις "leaping") is the assumably random movement of particles suspended in a fluid (i.e., a liquid such as water or a gas such as air) or the mathematical model used to describe such random movements, often called a particle theory.

Brownian motion has desirable mathematical characteristics, where statistics can be estimated with great precision, probabilities can be calculated, and hence scientists and analysts often turn to such an independent process when faced with the analysis of a multidimensional process of unknown origin (i.e. the stock market). The mathematical model of Brownian motion has several real-world applications. Brownian motion is among the simplest of the continuous-time stochastic (or probabilistic) processes, and it is a limit of both simpler and more complicated stochastic processes. One of the main distinctive features of the Brownian motion is that it has non-recurrent cycles and these cycles are the main subject of analysis [Peters, 2000].

The essence of work

Goals and objectives of our work:

For many years investors, economists, statisticians, and stock market players have been interested in developing and testing models of stock price behavior. One important model that needs attention and more thorough investigation is the Brownian motion method.

The aim is to analyze (like examlpe) the value function of the sold shares of the open joint stock company "State Energy Generating Company "Centrenergo" [Ukrainskaja birzha]. In this module we use collected primary data and processed secondary to calculate various parameters and uncertainty modeling required for running the model.

In this document, we do not present in detail all the parameter calculations that are necessary for project's economic and financial evaluation [Gabidulin, 2007]. Rather, we only present the calculations and modeling of several key parameters, including the function of the values; the correlation integral; the correlation entropy; the auto correlation function; the Hurst value.

Scientific novelty:

Great theoretical and practical interest presents the study, that allows according to the result of complex fluctuations in an example of a separate company, give an answer about the nature of modes of operation of the entire dynamic of the market as a whole, from the standpoint of modern concepts of deterministic chaos, where the role of the traditional linear approaches to the study of these patterns is strongly limited.

When making capacity, inventory, and production decisions, firms are typically uncertain about future market conditions such as demand, prices, and exchange rates. Within the framework of the investment market the Brownian motion is one of the first attempts to take into account the fluctuations of the financial market. The firms always take a certain risk. The use of stochastic methods would help to perform the high accuracy analyses of different data and to study the main driving forces of investment market. All that would help to create a well-established and empirically tested forecast updating model in which the forecast evolution process follows a geometric Brownian motion.

Practical value:

The actuality of research is caused by high practical significance and poor study of the problem of the investment risks at the market, submitted in shocks and jumps. It involves the need of the researchers and investors to create an effective instrument for adequate assessment of investment risks. With the help of Brownian motion method and other fractal methods we can get useful recommendations necessary for more accurate assessment of the market risks. And it can improve the work both of private and governmental establishments.

Applied methods of the investment risks today:

The main methods of the investment risk assessment can be divided into the following groups: building confidence intervals; finding out indexes; expert methods; historical methods; discounting methods; structure methods; defining the level of the project's stability; sensitivity analysis of the project; scenario analysis of the project development.

However, these methods are not steady to shocks and don't show enough impact of the initial conditions and noise. Meta heuristics, fractal methods, and stochastic methods are more flexible and show high accuracy of the calculations.

The most popular and traditional measure of risk is volatility. The main problem with volatility, however, is that it does not care about the direction of an investment's movement: a stock can be volatile because it suddenly jumps higher.

The proposed method of analyses:

We use the time row to assess the value of a statistical data. We located the row at the time sequence, and with the help of that we can identify the level's rate of change, determine the trends, and follow certain patterns of variable to identify the factors influencing the change process in the future [Orlov, 2004].

We analyzed the value function of the sold shares of the open joint stock company "State Energy Generating Company "Centrenergo"", which has 15% of the of all electric power in Ukraine and plays a significant role in supporting and regulating energy balance in the whole country. This analysis could be useful for investors and researches who work in the sphere of the energy (for example in the field of the renewable energy). For the sample data, we took an interval of one minute and 8521 values of the sold shares (UAH) [Anupindi, Bassok, Zemel, 2001]. The following figure (1) illustrates the model of the sold shares; the values of the function have the high wide variation, from which we can conclude that this function is stochastic [Peters, 2000].

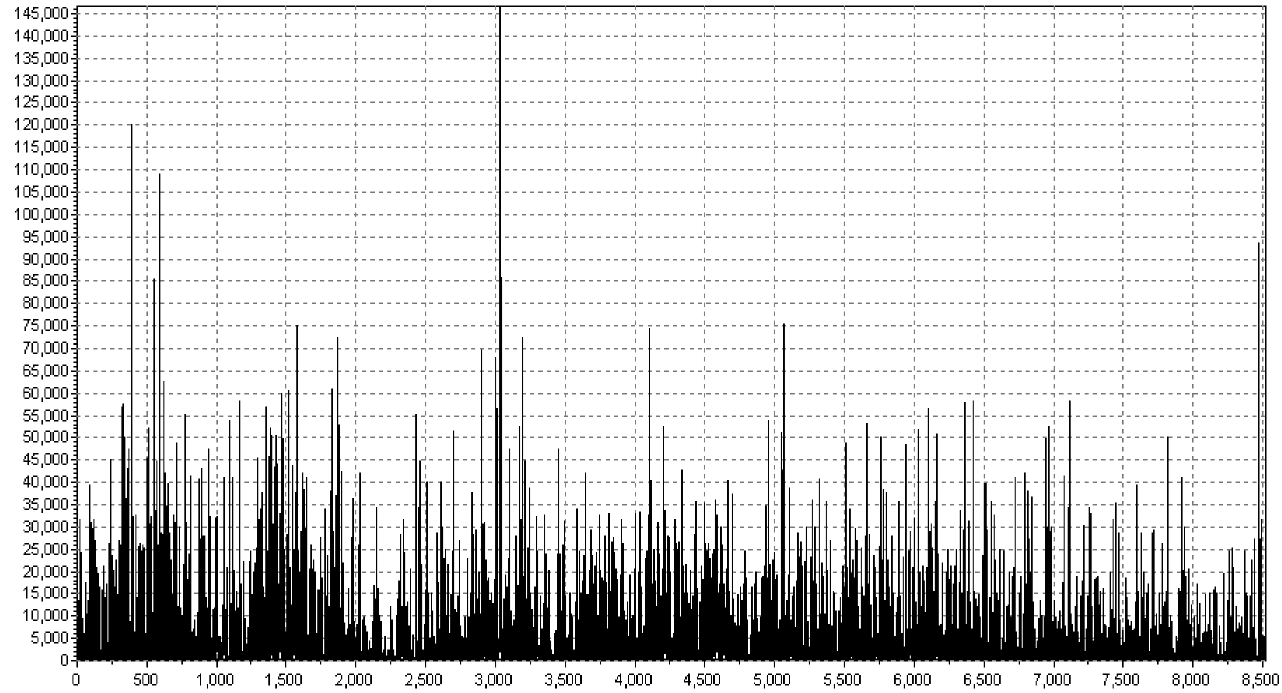


Figure 1 - The model of sold shares

Correlation dimension:

Correlation coefficient for short is a measure of the degree of linear relationship between two variables. We calculated the correlation size of this time sequence, which includes the minimum number of space dimensions that includes the trajectories[Markov, 2009].

With increasing embedding dimension, correlation dimension increases. However, for a deterministic system, no matter how chaotic it seems the correlation dimension sequence stops growing after reaching a certain dimension. If we observe the saturation at the sertain level, we can take this dimension as the estimate of the correlation dimension of attractor. For random data such saturation is not observed and the correlation dimension increases monotone [Kantz, Schreiber]. Correlation dimension $C(\varepsilon)$ is calculated using the following formula (1):

$$C(\varepsilon) = Lim_{n \longrightarrow \infty} \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \theta(\varepsilon - \left\| x(i) - x(j) \right\|), x(i) \in R^m \quad (1)$$

where ε is the threshold distance, $\|x(i) - x(j)\|$ is a norm (Euclid norm), θ is the function of Heaviside, R^m is Euclid multiplicity, N is the number of the observations, n is set of points in an m-dimensional space. The correlation integral helps to calculate the correlation dimension. In chaos theory, the correlation integral means the probability that two variables at different time intervals would have similar values (2):

$$C(\varepsilon) = Lim \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \theta(\varepsilon - \left\| x(i) - x(j) \right\|), x(i) \in R^m \quad (2)$$

In this sample it has the value 4.19. According to the figure (2), at the last iteration the correlation size starts to decrease and it means that the behavior of the function isn’t occasional and is subordinated by certain parameters’ principles.

It can show the change of the sold activity during the day, month or season. Depending on the size of the resulting values of the correlation dimension, all graphics can be divided into two types, reflecting the degree of presence of deterministic chaos in the term structure: the cycles of the chaotic dynamics of the average dimension (C ≤ 6); the cycles of the chaotic dynamics of high dimension (P> 6).

For the sales of “Centrenergo” correlation size is in the middle dimension [Kuusela, Jartti, Tahvanainen].

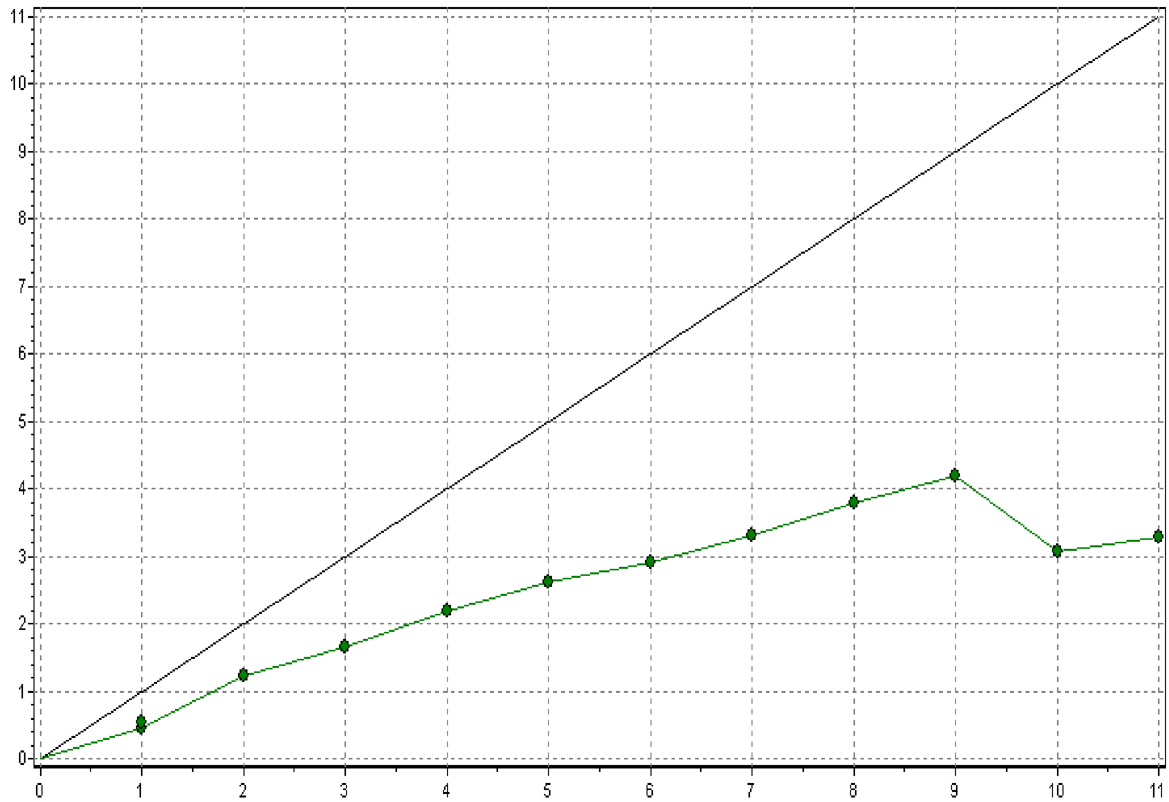


Figure 2 – The correlation size behavior

Correlation dimension can be used to assign a sample to a specific group depending on the presence of random processes, or to evaluate possible changes in the structure of sales of the company as it develops.

Correlation entropy:

We found out the correlation entropy that showed the degree of the recession of the near phase trajectories and allows measuring the amount of information necessary to forecast the behavior of this function in the future [Markov, 2009]. The correlation entropy $H(x)$ can be calculated by the following formula (3):

$$H(x) = -\sum_{k=1}^{m} w(x_k)\Delta x Log\{w(x_k)\Delta x\} \tag{3}$$

where $w(x_k)$ is probability density of the quantized continuous quantity, $\Delta(x)$ is chosen interval. According to the following figure (3) we found out the minimal value of the line, thus, for this research the correlation entropy is 3.8 and it means that the optimal time of function's stability is 3.8 minutes:

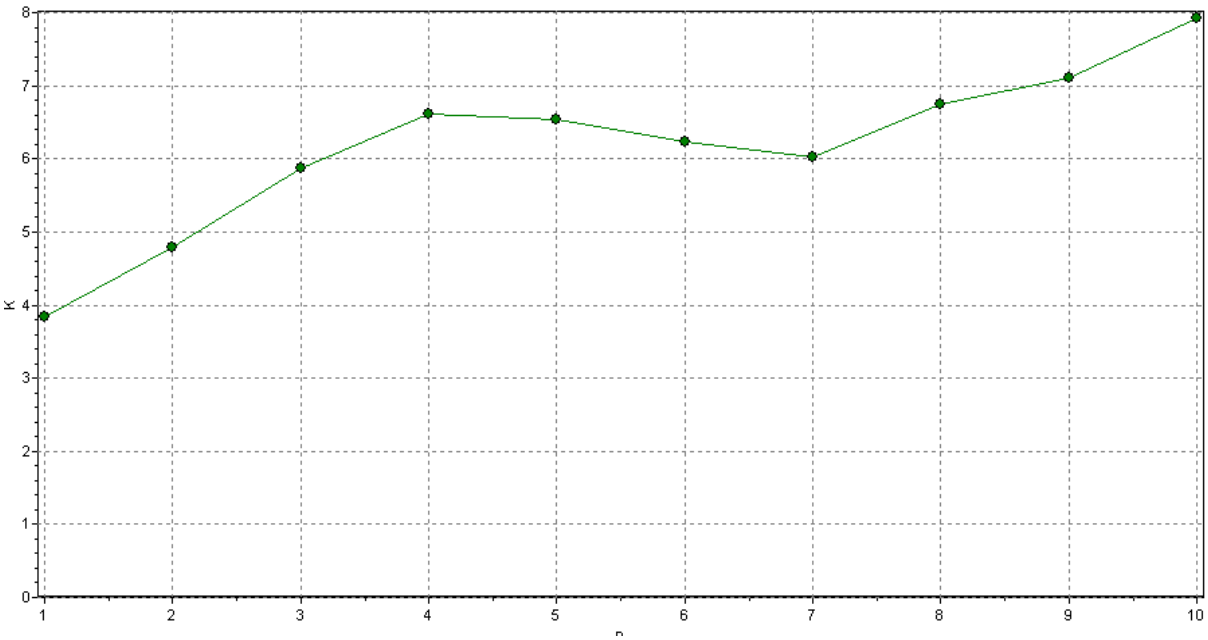


Figure 3 - Correlation entropy

The auto correlation function:

In statistics the auto correlation function shows the correlation dependence between the volumes of sales of the series with different time lag [Markov, 2008].

The autocorrelation function $R(t,s)$ can be calculated by the following formula (4):

$$R(t,s) = \frac{E[(x_t - \mu_t)(x_s - \mu_s)]}{\sigma_t \sigma_s} \tag{4}$$

where E is the expected value operator; X is the value produced by a given run of this function at time t or S; μ is correlation and σ is dispersion [Markov, 2008]. According to the received result, the largest point of the auto correlation function is achieved at the three-four values, and it tells that the inertness of the current sales time is saved for three-four minutes.

The Hurst value:

It shows the trendiness or persistence of row numeric incomes. Basically, this method is used to identify when a value is persistence i.e. the tendency of the value to continue its current direction and also antipersistence i.e. the tendency of the value to reverse itself rather than to continue its current direction. Or it is random and unpredictable. The statistic graphic can be used to estimate the cycle length. With the help of these calculations we can find out how many cycles has the value for a definite period. The Hurst calculations were made according to the formula from the work of Albert Einstein about Brownian motion of particles. According to these calculations

Brownian particle moves a distance equal to the square root of time spent on this promotion. If H=0.5 the system passes the same distance as the Brownian particle, during the time T. If T has the value more than 0.5, the system passes over a greater distance during the same period T as a Brownian particle. The calculation of the Hurst value H could be done according to the next formula (5):

$$H = \frac{Log(R/S)}{Log(aN)} \quad (5)$$

where H - the Hurst exponent; S - standard deviation of a set of observations; R - amplitude of the accumulated deviation; N - number of observation periods; a - a given constant, positive number (Hurst calculated this exponent constant empirically for a relatively short time series of natural phenomena as 0.5).

The following picture shows time dependence of the normalized value in a double logarithmic scale and its linear approximation. The slope coefficient of the approximating line at this picture is the Hurst value [Markov, 2008].

The Hurst value of this interval is 0.71 and it lies at the (0.5, 1] interval. It means that this row is trendiness and it means that the next value would have the same sign as the previous (figure 5).

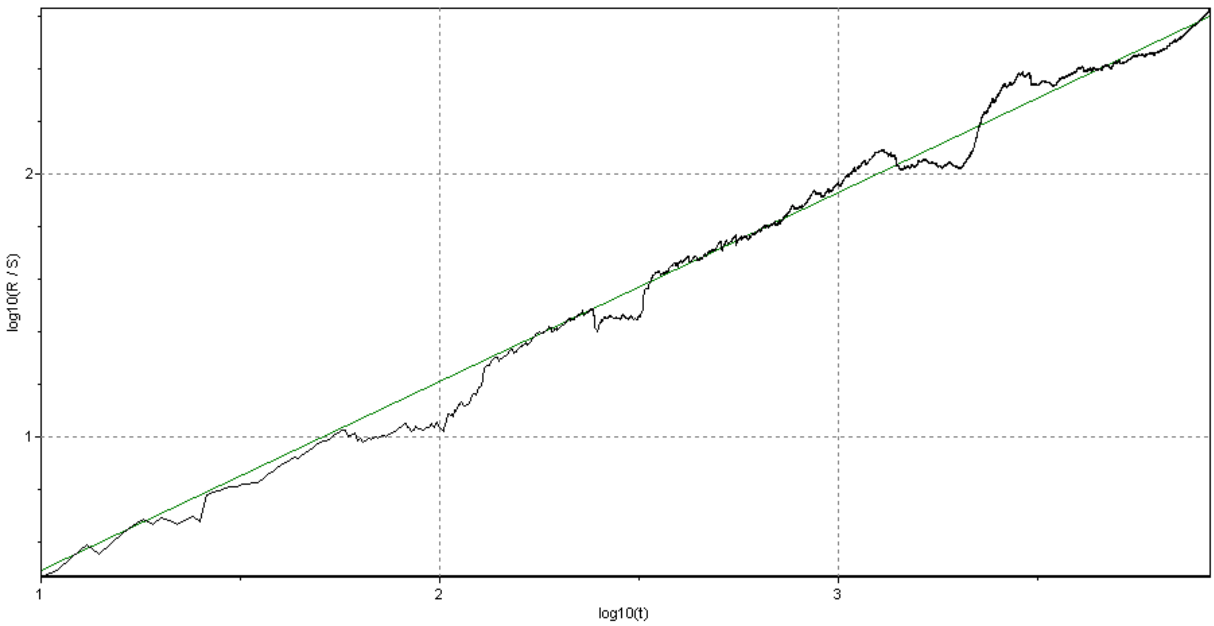


Figure 5 - The Hurst value

Investment planning starts at the top of the organization with strategic planning, which includes the analyses of the different goods and stocks. In order to assess the situation, we can look at past histories of the similar operations or projects and various factors that influence the future situation. The analyses of the data of company "Centrenergo" helped to identify useful information.

We found out the correlation dimension and with the help of that we found the probability that two variables at different time intervals would have similar values. This function of sold shares has the cycles of the chaotic dynamics of the average dimension (C ≤ 6). Thus, we found out the information that would help to build the investment forecast with the help of cycles. The correlation entropy of our function is 3.8 and it means that the optimal time of function's stability is 3.8 minutes. It would help to consider the risks, during the market analysis that takes into account the function of sales of company"Centrenergo".

We found out the the inertness of the current sales time and it can be used to build some kind of forecast for this function with certain probability. As distancing from the inertness time, the risk of deviation from the real function increases. With the help of the auto correlation function we could compare many pairs of cycles and models and search the interdependence. Comparison of models, finding out the interdependence and identification of the similar or reverse regions on the graphics of models constitute some important aspects of model validation that deserve serious study.

The Hurst value is the global characteristic, which shows the connection between the next increment and the previous increments, consisting in the fact that if the value of the function increased, the density distribution of the next increment is biased towards growth. These values also help to consider the risk.

With the help of all these characteristics we can make the recognition of the graphic view of the model and carry them for the sample comparison. It can be done by scaling of the image and splitting the image into pieces with different coefficients. But during the modification process of the of the sample model, its inner part should remain changeless.

With the help of Brownian motion method we can create a unique trade system, which would handle the data from different data bases and recourses and it would help to minimize risks and make reasonable decisions.

Conclusion

Every investment involves some degree of risk, but a solid understanding of risk in its different forms can help investors to understand better the opportunities, trade-offs and costs involved with different investment approaches. Described in this work the analysis method of statistic data for building the investment forecast with the help of Brownian motion model can be used to measure the risk during the planning of the investment activities, at the market alone or with complex of other methods, and can be faster and more accurate than the traditional existing ones. With the features described above, the model allows one to determine risk and quantify and optimize investment opportunities in different manners and categories. The cognitive power of these calculations is that they help to order and divide the analysed investment proccesses by their chaotic values and complexity, and, thus, classify and systematize them. It would help to determine, more accurately, the investment horizons, to develop the action strategy with the risk minimization and identify some features of the market as a whole.

Bibliography

[Ukrainskaja birzha] http://forum.ux.ua/. Dannye dlja tehnicheskogo analiza (in Russian)

[Salvatore, 2003] Salvatore D. Microeconomics: Theory and Applications. 4th ed.Oxford University Press, New York, 2003.

[Kronover, 2000] Kronover R.M. Fraktaly i haos v dinamicheskih sistemah. M.: Postmarket, 2000. (in Russian)

[Guasoni, 2006] Guasoni P. No arbitrage under transaction costs, with fractional Brownian motion and beyond 11 Mathematical Finance, Volume 16, Issue 3. -2006-pp. 569-582.

[Markov, 2009] Markov A.A. Ocenka granic stoimosti opcionov na fraktal'nom rynke // Vestnik jekonomicheskoj integracii. - M., 2009. - №8 (18), (in Russian)

[Markov, 2008] Markov A.A. Fraktal'naja razmernost' rossijskih i nekotoryh zarubezhnyh fondovyh indeksov // Matematicheskie metody analiza finansovyh vremennyh rjadov: Sbornik nauchnyh statej / Pod red. V.B. Gisina i A.B. Shapovala. - M.: Finakademija, 2008. - s. 49-61. (in Russian)

[Peters, 2000] Peters Je. Haos i porjadok na rynkah kapitala. Novyj analiticheskij vzgljad na cikly, ceny i izmenchivost' rynka. M.: Mir, 2000. (in Russian)

[Orlov, 2004] Orlov A.I. Prikladnaja statistika M.: Izdatel'stvo «Jekzamen», 2004. (in Russian)

[Gabidulin, 2007] Gabidulin, Je. M., Pilipchuk, N. I. Lekcii po teorii informacii. M.: MFTI, 2007. — 214 s. (in Russian)

[Anupindi, Bassok, Zemel, 2001] Anupindi, R., Y. Bassok, E. Zemel.. A general framework for the study of decentralized distribution systems. Manufacturing Service Oper. Management 3(4) 2001, 349–368.

[Granot, Soši´, 2003] Granot, D., G. Soši´.. A three-stage model of a decentralizedc distribution system of retailers. Oper. Res. 51(5) 2003, 771–784.

[Plambeck, Taylor, 2003] Plambeck, E. L., T. A. Taylor.. Sell the plant? The impact ofcontract manufacturing on innovation, capacity and profitability. Management Sci. 2003, 51(1) 133–150.

[Kuusela, Jartti, Tahvanainen, 2002] Kuusela T.A., Jartti T.T., Tahvanainen K.U.O., Kaila T.J. Nonlinear methods of biosignal analysis in assessing terbutaline-induced heart rate and blood pressure changes // Am. J. Physiol. Heart Circ. Physiol. 282: H773-H781, 2002.

[Kantz, Schreiber, 1997] Kantz H., Schreiber T. Nonlinear time series analysis. Cambridge: Cambridge University Press, 1997: 168-171.

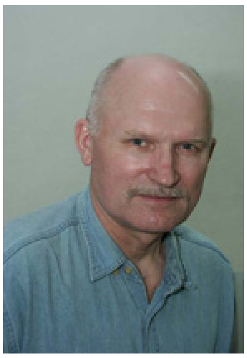
Authors' Information

Gurina Irina *– postgraduate student; Kharkiv National University of Radioelectronics; Kharkiv, Ukraine;*

e-mail: _charmel_@mail.ru

tel.: +380 93 25 97 630, +380 66 93 93 785

Major Fields of Scientific Research: General theoretical information research, Knowledge Discovery and Engineering, Business Informatics.



Oleksandr Kuzomin

Chief of Innovation Marketing Department

Professor of Information Science

14, Lenin Ave., 61166, Kharkiv, UKRAINE

Tel/fax: +38(057)7021515

*mailto:*kuzy@kture.kharkov.ua