VYTAUTAS MAGNUS UNIVERSITY

INSTITUTE OF MATHEMATICS AND INFORMATICS

Audrius KABAŠINSKAS

STATISTICAL ANALYSIS OF FINANCIAL MARKETS AND METHODS OF STATISTICAL MODELING

This dissertation was accomplished in the since 2003 to 2007 at the Institute of Mathematics and Informatics, Lithuania.

The right for the doctoral studies in informatics was granted to the Institute of Mathematics and Informatics together with Vytautas Magnus University by Government of the Republic of Lithuania, Decree No. 1285, issued on the 13th of December 2004.

Scientific Advisor:

Prof. Dr. Habil. Leonidas SAKALAUSKAS (Institute of Mathematics and Informatics, Physical sciences, Informatics, 09P)

Council of Scientific field:

Chairman

Prof. Dr. Habil. Vytautas Kaminskas (Vytautas Magnus University, Physical sciences, Informatics, 09P)

Members:

Prof. Dr. Habil. Romualdas Baušys (Vilnius Gediminas Technical University, technological sciences, Informatics engineering, 07 T)

Asoc. Prof. Dr. Habil. Vitalijus Denisovas (Klaipėda University, Physical sciences, Informatics, 09P)

Asoc. Prof. Dr. Antanas Lipeika (Institute of Mathematics and Informatics, Physical sciences, Informatics, 09P),

Prof. Dr. Habil. Aleksandras Vytautas Rutkauskas (Vilnius Gediminas Technical University, Social Sciences, Economics, 04 S)

Opponents:

Prof. Dr. Habil. Remigijus Leipus (Vilnius University, Physical sciences, mathematics, 01 P)

Prof. Dr. Habil. Laimutis Telksnys (Institute of Mathematics and Informatics, Physical sciences, Informatics, 09P)

The dissertation will be defended at the public meeting of the Scientific Council in the field of Informatics in the Conference and Seminars Center of the Institute of Mathematics and Informatics at 11 a.m. on January 10, 2008.
Address: Goštauto str. 12, LT-01108, Vilnius, Lithuania.

The summary of the doctoral dissertation was distributed on 10 December, 2007.

A copy of the doctoral dissertation is available for review at the M. Mažvydas National Library of Lithuania, the Library of the Institute of Mathematics and Informatics and the Library of the Vytautas Magnus University.

VYTAUTO DIDŽIOJO UNIVERSITETAS

MATEMATIKOS IR INFORMATIKOS INSTITUTAS

Audrius KABAŠINSKAS

FINANSINIŲ RINKŲ STATISTINĖ ANALIZĖ IR STATISTINIO MODELIAVIMO METODAI

Disertacija rengta 2003-2007 metais Matematikos ir Informatikos Institute.

Doktorantūros teisė suteikta kartu su Vytauto Didžiojo universitetu 2004 m. gruodžio mėn. 13 d. Lietuvos Respublikos Vyriausybės nutarimu Nr. 1285.

Mokslinis vadovas:

Prof. habil. dr. Leonidas SAKALAUSKAS (Matematikos ir Informatikos Institutas, fiziniai mokslai, informatika, 09P)

Disertacija ginama Vytauto Didžiojo universiteto Informatikos mokslo krypties taryboje:

Pirmininkas:

prof. habil. dr. Vytautas Kaminskas (Vytauto Didžiojo universitetas, fiziniai mokslai, informatika, 09 P),

Nariai:

prof. habil. dr. Romualdas Baušys (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija, 07 T),

doc. dr. Vitalijus Denisovas (Klaipėdos universitetas, fiziniai mokslai, informatika, 09 P);

doc. dr. Antanas Lipeika (Matematikos ir Informatikos institutas, fiziniai mokslai, informatika, 09 P),

prof. habil. dr. Aleksandras Vytautas Rutkauskas (Vilniaus Gedimino technikos universitetas, socialiniai mokslai, ekonomika, 04 S).

Oponentai:

prof. habil. dr. Remigijus Leipus (Vilniaus universitetas, Fiziniai mokslai, matematika, 01 P);

prof. habil. dr. Laimutis Telksnys (Matematikos ir Informatikos institutas, fiziniai mokslai, informatika, 09 P).

Disertacija ginama viešame Fizinių mokslų, Informatikos mokslo krypties tarybos posėdyje 2008 m. sausio mėn. 10 d. 11 val. Matematikos ir informatikos instituto konferencijų ir seminarų centre. Adresas: Goštauto g. 12, LT-01108, Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2007 m. gruodžio mėn. 10 d.

Disertaciją galima peržiūrėti M. Mažvydo nacionalinėje bibliotekoje, Matematikos ir informatikos instituto ir Vytauto Didžiojo universiteto bibliotekose.

1 General Characteristics of the Dissertation

1.1 The topicality of the scientific problem

Modeling and analysis of financial processes is an important and fast developing branch of computer science, applied mathematics, statistics, and economy. Probabilistic-statistical models are widely applied in the analysis of investment strategies. Adequate distributional fitting of empirical financial series has a great influence on forecast and investment decisions. Real financial data are often characterized by skewness, kurtosis, and heavy tails. Stable models are proposed (in scientific literature) to model such behavior.

Statistical models of financial data series and algorithms of forecasting and investment are the topic of this research.

1.2 The exploration level of scientific problems

Since the middle of the last century, financial engineering has become very popular among mathematicians and analysts. Stochastic methods were widely applied in financial engineering. Gaussian models were the first to be applied, but it has been noticed that they inadequately describe the behavior of financial series. Since the classical Gaussian models were taken with more and more criticism and eventually have lost their positions, new models were proposed. Stable models attracted special attention; however their adequacy in the real market should be verified. Nowadays, they have become an extremely powerful and versatile tool in financial modeling. For a long time processes in economics and finance have been described by Gaussian distribution (Brownian motion). Real data are often characterized by skewness, kurtosis and heavy tails and because of that reasons they are odds with requirements of the classical models. There are two essential reasons why the models with a stable paradigm (max-stable, geometric stable, α-stable, symmetric stable and other) are applied to model financial processes. The first one is that stable random variables (r.vs) justify the generalized central limit theorem (CLT), which states that stable distributions are the only asymptotic distributions for adequately scaled and centered sums of independent identically distributed random variables (i.i.d.r.vs). The second one is that they are leptokurtotic and asymmetric. This property is illustrated in Figure 1, where (a) and (c) are graphs of stable probability density functions (with additional parameters) and (b) is the graph of the Gaussian probability density function, which is also a special case of stable law.
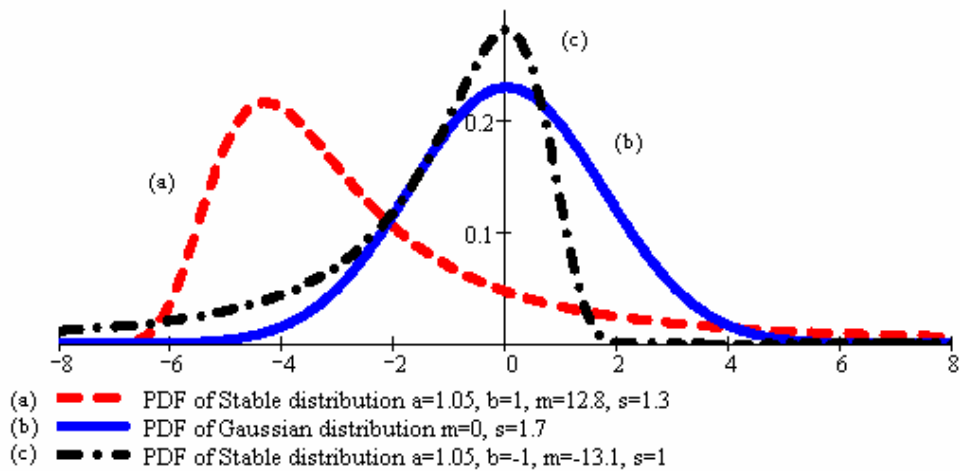

(a)	- - -	PDF of Stable distribution a=1.05, b=1, m=12.8, s=1.3
(b)	——	PDF of Gaussian distribution m=0, s=1.7
(c)	▪·▪·▪	PDF of Stable distribution a=1.05, b=-1, m=-13.1, s=1

Figure 1. Stable distributions are leptokurtotic and asymmetric (here *a* is a stability parameter, *b* - asymmetry parameter, *m* – location parameter and *s* is a scale parameter)

Following to S.Z. Rachev, "the α-stable distribution offers a reasonable improvement if not the best choice among the alternative distributions that have been proposed in the literature over the past four decades".

Each stable distribution $S_\alpha(\sigma,\beta,\mu)$ has the stability index α that can be treated as the main parameter, when we make an investment decision, β is the parameter of asymmetry, σ is that of

scale, and μ is the parameter of position. In models that use financial data, it is generally assumed that $\alpha \in (1;2]$. Stable distributions only in few special cases have analytical distribution and density functions. That is why they are often described by characteristic functions (CF). Several statistical and robust procedures are examined in creating the system for stock portfolio simulation and optimization. The problem of estimating the parameters of stable distribution is usually severely hampered by the lack of known closed form density functions for almost all stable distributions. Most of the methods in mathematical statistics cannot be used in this case, since these methods depend on an explicit form of the PDF. However, there are numerical methods that have been found useful in practice and are described below in this dissertation.

Since fat tails and asymmetry are typical of stable random variables, they better (than Gaussian) fit the empirical data distribution. Long ago in empirical studies it was noted that returns of stocks (indexes, funds) were badly fitted by the Gaussian law, while stable laws were one of the solutions in creating mathematical models of stock returns. There arises a question, why stable laws, but not any others are chosen in financial models. The answer is: because the sum of n independent stable random variables has a stable and only stable distribution, which is similar to the CLT for distributions with a finite second moment (Gaussian). If we are speaking about hyperbolic distributions, so, in general, the Generalized Hyperbolic distribution does not have this property, whereas the Normal-inverse Gaussian (NIG) has it. In particular, if Y_1 and Y_2 are independent normal inverse Gaussian random variables with common parameters α and β but having different scale and location parameters $\delta_{1,2}$ and $\mu_{1,2}$, respectively, then $Y = Y_1 + Y_2$ is $NIG(\alpha, \beta, \delta_1 + \delta_2, \mu_1 + \mu_2)$. So NIG fails against a stable random variable, because, in the stable case, only the stability parameter α must be fixed and the others may be different, i.e., stable parameters are more flexible for portfolio construction of different asymmetry. Another reason why stable distributions are selected from the list of other laws is that they have heavier tails than the NIG and other distributions from the generalized hyperbolic family (its tail behavior is often classified as "semi-heavy").

Foreign financial markets and their challenges were always of top interest for stock brokers. The new investment opportunities emerged after expansion of the European Union in 2004. Undiscovered markets of the Baltic States and other countries of Central and Eastern Europe became very attractive for investors. Unbelievable growth of the gross domestic product (GDP) 3-8% (the average of the EU is 1.5–1.8%) and high profitability overcame the risk. But a deep analysis has not yet been made in those markets. For a long time it has been known that financial series are the source of self-similar and multifractal phenomena and numerous empirical studies support that. In this dissertation, the analysis of daily stock returns of the Baltic States and some world wide known indexes is made. Financial series in the Baltic States bear two very important features (compared with the markets of the USA and EU):

1. Series are rather short: 10–12 years (not exceeding 2000 data points), but only recent 1000–1500 data points are relevant for the analysis;

2. A stagnation phenomenon is observed in empirical data (1993–2005). Stagnation effects are characterized by an extremely strong passivity: at some time periods stock prices do not change because there are no transactions at all.

To avoid the short series problem, the bootstrap method was used. The bootstrap is a method for estimating the distribution of an estimator or test statistic by treating the data as if they were the population of interest. In a word, the bootstrap method allows us to "make" long enough series required in multifractality and self-similarity analysis, from the short ones.

The second problem, called a "daily zero return" problem, is more serious than it may seem. The Baltic States and other Central and Eastern Europe countries have "young" financial markets and they are still developing (small emerging markets), financial instruments are hardly realizable and therefore they are often non-stationary, and any assumptions or conclusions may be inadequate when speaking about longtime series. Stagnation effects are often observed in young markets. In such a case, the number of daily zero returns can reach 89%. A new kind of model should be developed and analyzed, i.e., we have to include one more additional condition into the model – the

daily stock return is equal to zero with a certain (rather high) probability *p*. Anyway, this problem may be solved by extending a continuous model to the mixed one, where daily returns equal to zero are excluded from the series when estimating the stability parameters. The series of non-zero returns are fitted to the stable distribution. Stable parameters are estimated by the maximal likelihood method. Goodness of fit is verified by the Anderson-Darling distributional adequacy test. The stability is also tested by the homogeneity test, based on the fundamental property of stable laws. Unfortunately, because of strong passivity, continuous distribution fitting tests (Anderson-Darling, Kolmogorov-Smirnov, etc) are hardly applicable. An improvement based on mixed distributions is proposed and its adequacy in the Baltic States market is tested. In this dissertation the Koutrouvelis goodness-of-fit, test based on the empirical characteristic function and modified χ^2 (Romanovski) test, was used.

When constructing a financial portfolio, it is essential to determine relationships between different stock returns. In the classical economics and statistics (the data have finite first and second moments), the relationship between random variables (returns) is characterized by covariance or correlation. However under the assumption of stability (sets of stock returns are modeled by stable laws), covariance and correlation (Pearson correlation coefficient) cannot be applied, since the variance (if the index of stability $\alpha < 2$) and the mean (if the index of stability $\alpha < 1$) do not exist. In this case, we can apply rank correlation coefficients (ex. Spearman or Kendall) or the contingency coefficient. Under the assumption of stability, it is reasonable to apply generalized covariance coefficients – covariation or codifference. Therefore the generalized Markowitz problem is solved taking the generalized relationship measures (covariation, codifference). It has been showed that the implementation of codifference between different stocks greatly simplifies the construction of the portfolio.

Typical characteristics of the passivity phenomenon are constancy periods of stock prices. The dissertation deals with the distributional analysis of constancy period lengths. Empirical study of 69 data series from the Baltic states market and modeling experiments have showed that constancy period lengths are distributed by the Hurwitz zeta distribution instead of geometrical distribution. An improved mixed stable model with dependent states of stock price returns is proposed.

1.3   The aim and objectives of the work

The aim of this work is to create and investigate the methods of financial market analysis and financial portfolio management under the assumption of data series stability.

In order to achieve this aim, the following problems were solved:

1. To determine, analyze and experimentally test the statistical methods of stability analysis:
   - Test the stability by nonparametric fitting and homogeneity of aggregated and original series;
   - Test the self-similarity and multifractality of financial data;
   - Create a special software to the applied methods mentioned;
2. To study the passivity effect-phenomenon in emerging markets:
   - Develop a theoretical background for the analysis of passivity in the financial market;
   - Analyze the possibilities of implementing the mixed-stable model for financial series modeling;
   - Experimentally test the reliability of parameter estimation methods and distribution fitting tests, for laws with non continuous probability distribution and density functions;
3. To analyze, experimentally test and compare to another the stable parameter estimation methods that consider the passivity effect.
4. To develop a software system for the analysis of financial market and planning of financial resources, and to test an interactive demo version of the system.

1.4   The object of research

The objects of research in the dissertation presented are the historical data of financial securities (stock, equity, currency exchange rates, financial indices, etc.), statistical models of stock returns, parameter estimation methods, effects of self-similarity and multifractality, and algorithms of financial portfolio selection.

In this dissertation, data series of the developed and emerging financial markets are used as an example. The studied series represent a wide spectrum of stock market. Information that is typically (finance.yahoo.com , www.omxgroup.com , etc.) included into a financial database is:

- Unique trade session number and date of trade;
- Stock issuer;
- Par value;
- Stock price of last trade;
- Opening price;
- High - low price of trade;
- Average price;
- Closure price;
- Price change %;
- Supply – Demand;
- Number of Central Market (CM) transactions;
- Volume;
- Maximal – Minimal price in 4 weeks;
- Maximal – Minimal price in 52 weeks;
- Other related market information.

We use here only the closure price, because we will not analyze data as a time series and its dependence. We analyzed the following r.vs

$$X_i = \frac{P_{i+1} - P_i}{P_i}$$

where P is a set of stock prices. While calculating such a variable, we transform data (Figure 2) from price to return.
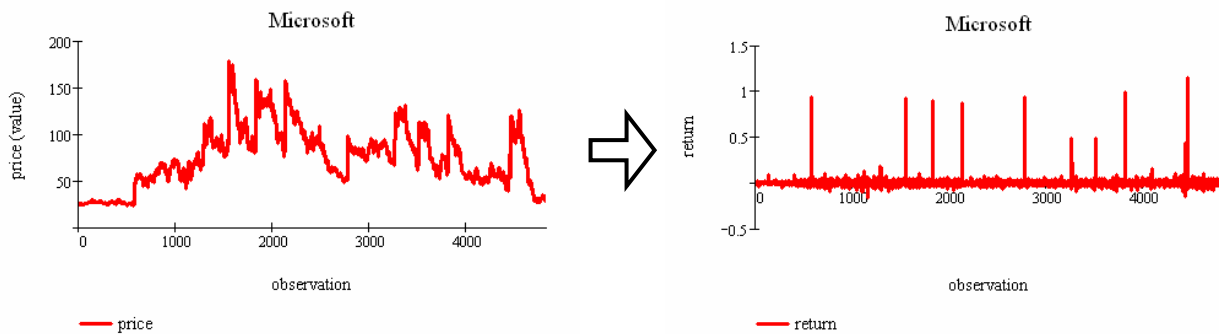

Figure 2. Data transformation

The length of series is very different starting from 1566 (6 years, NASDAQ) to 29296 (107 years, DJTA). Also very different industries are chosen, to represent the whole market. The Baltic States (64 companies) series studied represent a wide spectrum of the stock market (the whole Baltic Main list and Baltic I-list). The length of series is very different, starting from 407 to 1544. The average of data points is 1402. The number of zero daily stock returns differs from 12% to 89%, on the average 52%.

Almost all the data series are strongly asymmetric ($\hat{\gamma}_1$), and the empirical kurtosis ($\hat{\gamma}_2$) shows that density functions of the series are more peaked than that of Gaussian. That is why we make an assumption that Gaussian models are not applicable to these financial series.

1.5   Novelty and theoretical significance of research

The new research results are as follows:

1. Numerical methods (MLE and robust) for parameter estimation of stable models (with asymmetry) have been created and their efficiency were compared;

2. Complex analysis methods of testing stability hypotheses have been created and special software was developed:
   - nonparametric distribution fitting tests were performed and homogeneity of aggregated and original series was tested;
   - theoretical and practical analysis of self-similarity and multifractality was made (in the World wide and the Baltic states markets);
3. To solve the passivity problem in emerging markets the mixed-stable model was introduced. This model generalizes the stable financial series modeling.
4. Analysis of stagnation periods was made. It has been shown that lengths of stagnation periods may be modeled by the Hurwitz zeta law.
5. Methods of statistical relationship measures between shares returns were studied and algorithms of significance were introduced.

1.6   Practical significance of the work

1. A programming system (which includes procedures in C++, JAVA and MathCad) was developed to model and analyze the financial data series. This system allows estimating stable parameters (including passivity effect), simulating new series, and constructing an optimal portfolio.
2. Empirical tests with the data from the Baltic States and other markets have showed that continuous (α-stable, hyperbolic, and even normal) distributions are usually inadequate models of the returns. However, the data almost always fit the mixed-stable law. It should be noted that some properties of this model may be applied to model the failures in computer networks.
3. Portfolio selection methods are proposed in the case of the mixed-stable model.

1.7   Work approbation

The research results were presented and discussed at 10 national and 11 international conferences in Lithuania and abroad.

The main results of this dissertation were published in **9** scientific journals: **6 articles** in reviewed scientific publications and **3 articles** in the proceedings of scientific conferences.

1.8   The scope of the scientific work

The work is written in Lithuanian. It consists of 5 chapters (introduction, analytical part, methodology, results and conclusions), the list of references, and appendices. There are 171 page of the text, 58 tables, 52 figures, 30 algorithms, 3 appendices, and 167 bibliographical sources.

Chapter 1 is introductory and describes the relevance of the problem, the scientific novelty of the results and their practical significance, the objectives and tasks of the work are formulated, some abbreviations and terminology used are also explained.

Chapter 2 explores an overview of scientific and practical researches of the subject and presents a structural analysis of financial modeling and related fields. In this chapter, the following problems are discussed: development of financial engineering, traditional methods of portfolio selection (including models with transaction costs) the modern financial models are reviewed, an overview of stability analysis methods is given, some approaches on self-similarity and multifractality are discussed, an overview of the existing modeling, optimization and planning software is given, financial data sources are analyzed and characterized, and some portfolio optimization problems are discussed.

Chapter 3 is designated to modeling and optimization of financial markets.

The stable distributions and an overview of their properties

We say that a r.v. X is distributed by the stable law and denote

$$X \overset{d}{=} S_\alpha(\sigma, \beta, \mu) \ ,$$

where S_α is the probability density function, if a r.v. has the characteristic function:

$$\phi(t) = \begin{cases} \exp\left\{-\sigma^{\alpha}\cdot|t|^{\alpha}\cdot\left(1-i\beta\,\mathrm{sgn}(t)\tan(\frac{\pi\alpha}{2})\right)+i\mu t\right\}, & \text{if } \alpha \neq 1 \\ \exp\left\{-\sigma\cdot|t|\cdot\left(1+i\beta\,\mathrm{sgn}(t)\frac{2}{\pi}\cdot\log|t|\right)+i\mu t\right\}, & \text{if } \alpha = 1 \end{cases}$$

Each stable distribution is described by 4 parameters: the first one and most important is the stability index $\alpha\in(0;2]$, which is essential when characterizing financial data. The others, respectively are: skewness $\beta\in[-1,1]$, a position $\mu\in\mathbf{R}$, the parameter of scale $\sigma>0$.

The probability density function is

$$p(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\phi(t)\cdot\exp(-ixt)dt .$$

In the general case, this function cannot be expressed as elementary functions. The infinite polynomial expressions of the density function are well known, but it is not very useful for Maximal Likelihood estimation because of infinite summation of its members, for error estimation in the tails, and so on. We use an integral expression of the PDF in standard parameterization

$$p(x,\alpha,\beta,\mu,\sigma) = \frac{1}{\pi\sigma}\int\limits_{0}^{\infty}e^{-t^{\alpha}}\cdot\cos\left(t\cdot\left(\frac{x-\mu}{\sigma}\right)-\beta t^{\alpha}\tan\left(\frac{\pi\alpha}{2}\right)\right)dt .$$
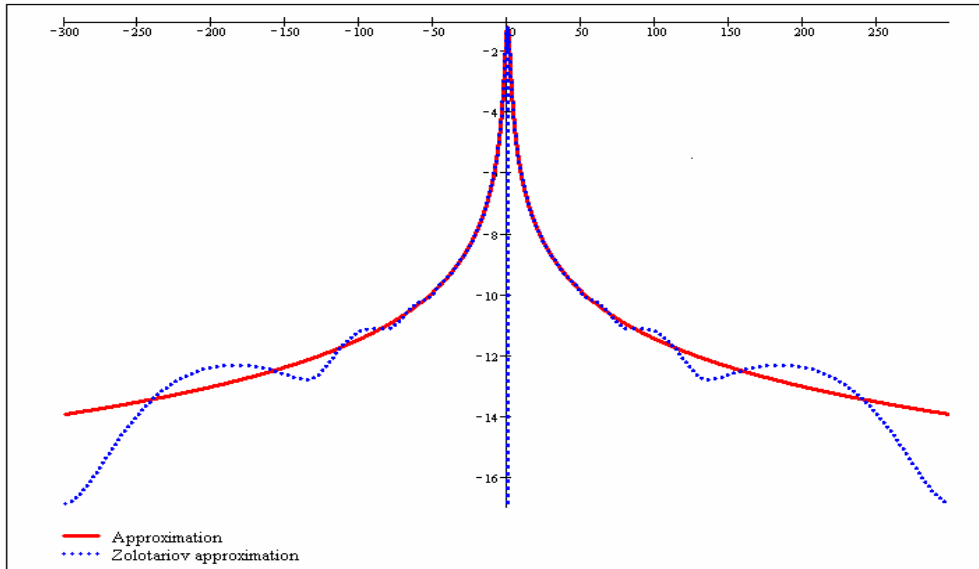

Figure 3. Logarithm of the probability density function S1.5(1,0,0)

It is important to note that Fourier integrals are not always convenient to calculate PDF because the integrated function oscillates. That is why a new Zolotarev formula is proposed which does not have this problem:

$$p(x,\alpha,\beta,\mu,\sigma) = \begin{cases} \dfrac{\alpha\left|\dfrac{x-\mu}{\sigma}\right|^{\frac{1}{\alpha-1}}}{2\sigma\cdot|\alpha-1|}\displaystyle\int\limits_{-\theta}^{1}U_{\alpha}(\varphi,\theta)\exp\left\{-\left|\dfrac{x-\mu}{\sigma}\right|^{\frac{a}{\alpha-1}}U_{\alpha}(\varphi,\theta)\right\}d\varphi, & \text{if } x \neq \mu \\ \dfrac{1}{\pi\sigma}\cdot\Gamma\left(1+\dfrac{1}{\alpha}\right)\cdot\cos\left(\dfrac{1}{\alpha}\arctan\left(\beta\cdot\tan\left(\dfrac{\pi\alpha}{2}\right)\right)\right), & \text{if } x = \mu \end{cases}$$

$$U_{\alpha}(\varphi,\vartheta) = \left(\frac{\sin\left(\frac{\pi}{2}\alpha(\varphi+\vartheta)\right)}{\cos\left(\frac{\pi\varphi}{2}\right)}\right)^{\frac{\alpha}{1-\alpha}}\cdot\left(\frac{\cos\left(\frac{\pi}{2}((\alpha-1)\varphi+\alpha\vartheta)\right)}{\cos\left(\frac{\pi\varphi}{2}\right)}\right),$$

where $\theta = \arctan\left(\beta \tan \dfrac{\pi\alpha}{2}\right)\dfrac{2}{\alpha\pi}\cdot \mathrm{sgn}(x-\mu)$.

If $\mu=0$ and $\sigma=1$, then $p(x,\alpha,\beta)=p(-x,\alpha,-\beta)$.

A stable r.v. has a property, tat may be expressed in two equivalent forms:

- If X_1, $X_2,...,$ X_n are independent r.vs. distributed by $S_\alpha(\sigma,\beta,\mu)$, then $\sum\limits_{i=1}^{n} X_i$ will be distributed by $S_\alpha(\sigma\cdot n^{1/a},\beta,\mu\cdot n)$.

- If X_1, $X_2,...,$ X_n are independent r.vs. distributed by $S_\alpha(\sigma,\beta,\mu)$, then

$$\sum_{i=1}^{n} X_i \overset{d}{=} \begin{cases} n^{1/\alpha}\cdot X_1 + \mu\cdot(n-n^{1/\alpha}),\ \text{if}\ \alpha\neq 1 \\ n\cdot X_1 + \dfrac{2}{\pi}\cdot\sigma\cdot\beta\cdot n\ln n,\ \text{if}\ \alpha = 1 \end{cases}.$$

One of the most fundamental stable law statements is as follows.

Let $X_1, X_2,...,X_n$ be independent identically distributed random variables and

$$\eta_n = \frac{1}{B_n}\sum_{k=1}^{n} X_k + A_n,$$

where $B_n>0$ and A_n are constants of scaling and centering. If $F_n(x)$ is a cumulative distribution function of r.v. η_n, then the asymptotic distribution of functions $F_n(x)$, as $n\to\infty$, may be stable and only stable. And vice versa: for any stable distribution $F(x)$, there exists a series of random variables, such that $F_n(x)$ converges to $F(x)$, as $n\to\infty$.

The pth moment $E|X|^p = \int_0^\infty P(|X|^p > y)dy$ of the random variable X exists and is finite only if $0<p<\alpha$. Otherwise, it does not exist.

Stable processes

A stochastic process $\{X(t), t\in T\}$ is stable if all its finite dimensional distributions are stable.

Let $\{X(t), t\in T\}$ be a stochastic process. $\{X(t), t\in T\}$ is α-stable if and only if all linear combinations $\sum\limits_{k=1}^{d} b_k X(t_k)$ (here $d\geq 1$ $t_1,t_2,...,t_d \in T$, b_1, $b_2,...,b_d$ – real) are α-stable. A stochastic process $\{X(t), t\in T\}$ is called the (standard) α-stable Levy motion if:

(1) $X(0)=0$ (almost surely);
(2) $\{X(t): t\geq 0\}$ has independent increments;
(3) $X(i)-X(s)\sim S_\alpha((t-s)^{1/\alpha}, \beta, 0)$, for any $0\leq s<t<\infty$ and $0<\alpha\leq 2$, $-1\leq\beta\leq 1$.

Note that the α-stable Levy motion has stationary increments. As $\alpha=2$, we have the Brownian motion.

Parameter Estimation Methods

The problem of estimating the parameters of stable distribution is usually severely hampered by the lack of known closed form density functions for almost all stable distributions. Most of the methods in mathematical statistics cannot be used in this case, since these methods depend on an explicit form of the PDF. However, there are numerical methods that have been found useful in practice and are described below. Given a sample $x_1,...,x_n$ from the stable law, we will provide estimates $\hat\alpha$, $\hat\beta$, $\hat\mu$, and $\hat\sigma$ of α, β, μ, and σ. Also, some empirical methods were used:

- Method of Moments (empirical CF);
- Regression method.

Comparison of estimation methods

We simulated a sample of 10 thousand members with the parameters $\alpha=1.75$, $\beta=0.5$, $\mu=0$ and $\sigma=1$. Afterwards we estimated the parameters of a stable random variable with different estimators. All the methods are decent, but the maximal likelihood estimator yields the best results. From the

practical point-of-view, MLM is the worst method, because it is very time-consuming. For large sets (~10.000 and more) we suggest using the regression (or moments) method to estimate α, β and σ, then estimate μ by MLM (optimization only by μ). As a starting point you should choose α, β, σ and sample mean, if $\alpha>1$ and a median, otherwise, for μ. For short sets, use MLM with any starting points (optimization by all 4 parameters).

A mixed stable distribution model

Let $Y \sim B(1,p)$ and $X \sim S_\alpha$. Let a mixed stable r. v. Z take the value 0 with probability 1 if $Y=0$, else $Y=1$ and $Z=X$. Then we can write the distribution function of the mixed stable distribution as

$$P(Z < z) = P(Y = 0) \cdot P(Z < z \mid Y = 0) + P(Y = 1) \cdot P(Z < z \mid Y = 1) = p \cdot \varepsilon(z) + (1 - p) \cdot S_\alpha(z) \quad (1)$$

where $\varepsilon(x) = \begin{cases} 0, & x \le 0 \\ 1, & x > 0 \end{cases}$, is the cumulative distribution function (CDF) of the degenerate distribution. The PDF of the mixed-stable distribution is

$$f(x) = p \cdot \delta(x) + (1 - p) \cdot p_\alpha(x),$$

where $\delta(x)$ is the Dirac delta function.

Cumulative density, probability density and characteristic functions of mixed distribution

For a given set of returns $\{x_1, x_2, \cdots, x_n\}$, let us construct a set of nonzero values $\{\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_{n-k}\}$. The equity ZMP1L (Žemaitijos pienas), from Vilnius stock exchange is given as an example ($p=0,568$). Then the likelihood function is given by

$$L(\bar{x}, \theta, p) \sim (1 - p)^k \, p^{n-k} \prod_{i=1}^{n-k} p_\alpha(\bar{x}_i, \theta) \qquad (2)$$

where θ is the vector of parameters (in the stable case, $\theta = (\alpha, \beta, \mu, \sigma)$). The function $(1 - p)^k \, p^{n-k}$ is easily optimized: $p_{\max} = \dfrac{n - k}{n}$. So we can write the optimal CDF as

$$F(z) = \frac{n - k}{n} S_\alpha(z, \theta_{\max}) + \frac{k}{n} \varepsilon(z), \qquad (3)$$
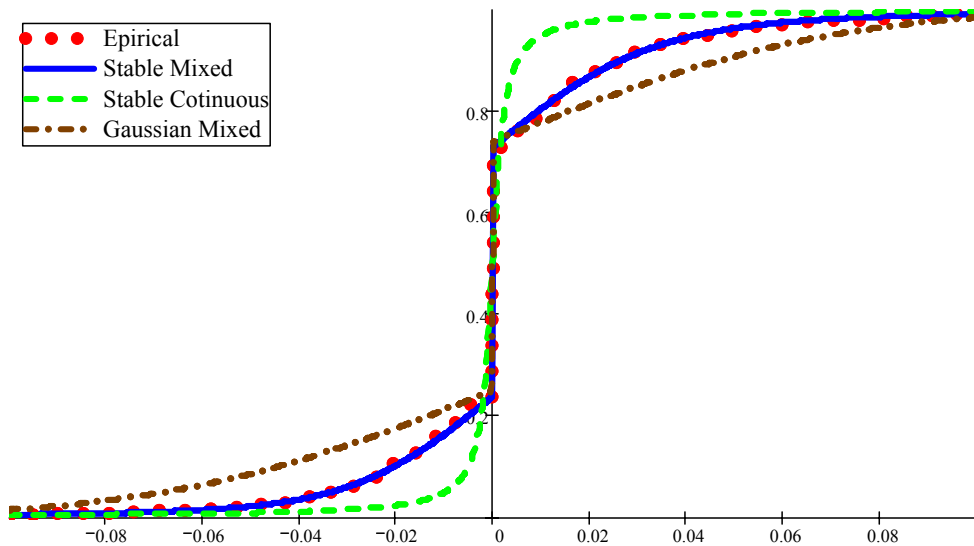

Figure 4, CDF of ZMP1L.

where the vector θ_{max} of parameters is estimated with nonzero returns.

The probability density f unction

$$p(z) = \frac{n - k}{n} p_\alpha(z, \theta_{\max}) + \frac{k}{n} \delta(z). \qquad (4)$$
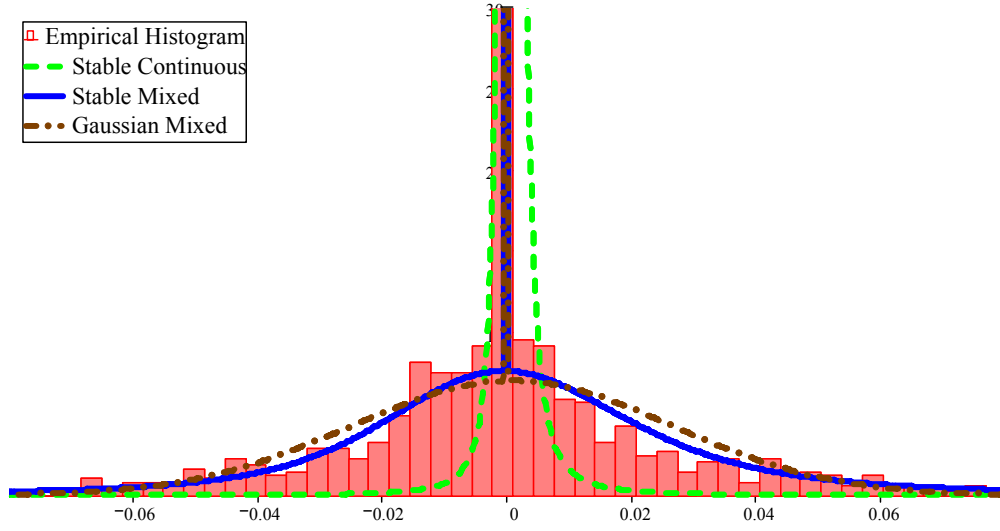
Figure 5. PDF and a histogram of ZMP1L.

Finally we can write down and plot (Figure 6) the characteristic function (CF) of the mixed distribution.

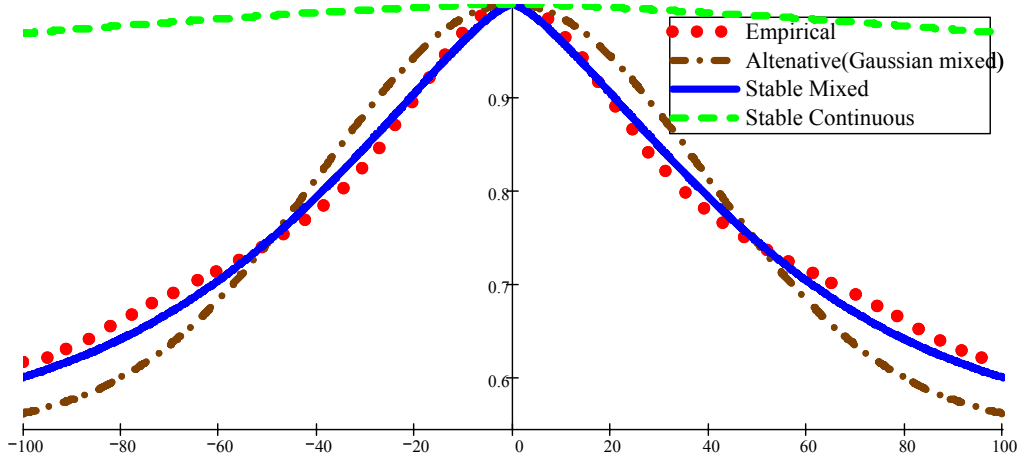$$\phi_{mix}(t) = \frac{n-k}{n} \cdot \phi(t) + \frac{k}{n}$$

Figure 6. Empirical, Gaussian, Stable mixed, and Stable continuous CF of ZMP1L.

The empirical characteristic function $\hat{\phi}(t, X) = \frac{1}{n}\sum_{j=1}^{n} e^{itX_j}$.

Mixed model adequacy

There arises a problem when we are trying to test the adequacy hypothesis for these models. Since we have a discontinuous distribution function, the classic methods (Kolmogorov–Smirnov, Anderson-Darling) do not work for the continuous distribution, and we have to choose a goodness–of–fit test based on the empirical characteristic function, or to trust a modified χ^2 (Romanovski) method. The results (see Table 1) of both methods are similar (match in 48 cases).

Table 1. Results of goodness-of-fit tests (accepted/rejected cases).

Method \ Fit	Gaussian	mixed Gaussian	stable	mixed stable
Modified χ^2	0/64	7/64	0/64	52/64
Empirical CF	0/64	0/64	12/64	52/64
Anderson – Darling	0/64	-	0/64	-

The CF-based test of Brown and Saliu is not so good (89% of all cases were rejected, since they are developed for symmetric distributions). A new stability test for asymmetric (skewed) alpha-stable distribution functions, based on the characteristic function, should be developed, since the existing tests are not reliable. Detailed results of stable-mixed model fitting are given in Table 2.

Table 2. Mixed model fit dependence on the number of zeros in series.

Number of "zeros"	Number of such series	Fits mixed model (χ^2, %)	Fits mixed model (Empirical CF, %)
0,1-0,2	2	100	100
0,2-0,3	2	100	100
0,3-0,4	8	25,00	25,00
0,4-0,5	17	64,71	94,12
0,5-0,6	14	71,43	100
0,6-0,7	15	86,67	100
0,7-0,8	4	100	100
0,8-0,9	2	100	100

One can see that when the number of "zeros" increases, the mixed model fits the empirical data better.

A mixed-stable model of returns distribution was proposed. Our results show that this kind of distribution fits the empirical data better than any other. The implementation of this model is hampered by the lack of goodness-of-fit tests for discontinuous distributions. Since adequacy tests for continuous distribution functions cannot be implemented, the tests, based on the empirical characteristic function and a modified χ^2 test, are used.

Modeling of stagnation intervals in emerging stock markets

We analyzed the following r.vs $X_i = 0$, if $P_{i+1} = P_i$ and $X_i = 1$, if $P_{i+1} \neq P_i$, where $\{P_i\}$ is a set of stock prices and $\{X_i\}$ is set of discrete states, representing behavior of stock price (change=1 or not=0).

Empirical study of lengths distribution of zero state runs

Theoretically if states are independent (Bernoulli scheme), then the series of lengths of zero state runs should be distributed by geometrical law. However, the results of empirical tests do not corroborate this assumption. We will fit the series distribution of lengths of zero state runs by discrete laws (generalized logarithmic, generalized Poisson, Hurwitz zeta, generalized Hurwitz zeta, discrete stable). The probability mass function of Hurwitz zeta law is $P(\xi = k) = v_{s,q}(k+q)^{-s}$

$P(\xi = k) = v_{s,q}(k+q)^{-s}$, where $v_{s,q} = \left(\sum_{i=0}^{\infty} (i+q)^{-s} \right)^{-1}$, $k \in N$, q>0, s>1. The parameters of all discrete distributions are estimated by the maximal likelihood method.

Transformation and distribution fitting

First of all, we will show how financial data from the Baltic States market are transformed to subsets length of zero state series and then we will fit each of the discrete distributions mentioned in above section. Carvalho, Angeja and Navarro have showed that data in network engineering fit the discrete logarithmic distribution better than the geometrical law. So we intend to test whether such a property is valid for financial data from the Baltic States market.

A set of zeros between two units is called a run. The first run is a set of zeros before the first unit and the last one after the last unit. The length of the run is equal to the number of zeros between two units. If there are no zeros between two units, then such an empty set has zero length.

To transform our data (from the state series, e.g., 010011101011100110) the two following steps should be taken: (a) extract the zero state runs (e.g., 0,00,0,0,000,0) from the states series; (b)

calculate the length of each run (1,2,0,0,1,1,0,0,3,0,1,). After the transformation, we estimated the parameters of each discrete distribution mentioned above and tested the nonparametric χ^2 distribution fitting hypothesis.

As mentioned above, theoretically this series should be distributed by geometrical law, however, from Table 2 we can see that other laws fit our data (57 series) much better. It means that zero state series from the Baltic States market are better described by the Hurwitz zeta distribution.

This result allows us to assume that zero-unit states are not purely independent. The Wald–Wolfowitz runs test corroborates this assumption for almost all series from the Baltic States market.

The mixed stable model with dependent states

Since the runs test rejects the randomness hypothesis of the sequence of states, the probability of states (zeros and ones) depends on the position in the sequence. If the lengths of states sequences are distributed by Hurwitz zeta law, then the probabilities of states are

$$P(X_n = 1 | ..., X_{n-k-1} = 1, \underbrace{X_{n-k} = 0, ..., X_{n-1} = 0}_{k}) = \frac{p_k}{1 - \sum_{j=0}^{k-1} p_j}, \; n \in N, k \in Z_0,$$

where p_k are probabilities of Hurwitz zeta law; $P(X_0=1) = p_0$. It should be noted that $P(X_n = 0 | ...) = 1 - P(X_n = 1 | ...), \; n, k \in Z_0$.

With the probabilities of states and distribution of nonzero returns we can generate sequences of stock returns (interchanging in the state sequence units with a stable r.v.). So, the mixed-stable modeling with dependent states is more advanced than that with independent (Bernoulli) states, it requires parameter estimation by both the stable ($\alpha, \beta, \mu, \sigma$) and Hurwitz zeta ($q, s$) law.

Analysis of stability

Examples of stability analysis can be found in the works of Rachev and Weron. In the latter paper, Weron analyzed the DJIA index (from 1985-01-02 to 1992-11-30, 2000 data points in all). The stability analysis was based on the Anderson–Darling criterion and by the weighted Kolmogorov criterion (D'Agostino), the parameters of stable distribution were estimated by the regression method proposed by Koutrouvelis. The author states that DJIA characteristics perfectly correspond to stable distribution.

Almost all data series are strongly asymmetric ($\hat{\gamma}_1$), and the empirical kurtosis ($\hat{\gamma}_2$) shows that density functions of series are more peaked than Gaussian. That is why we make an assumption that Gaussian models are not applicable to these financial series. The distribution (Figure 7) of α and β estimates shows that usually α is over 1.5 and for sure less than 2 (this case 1.8) for financial data.
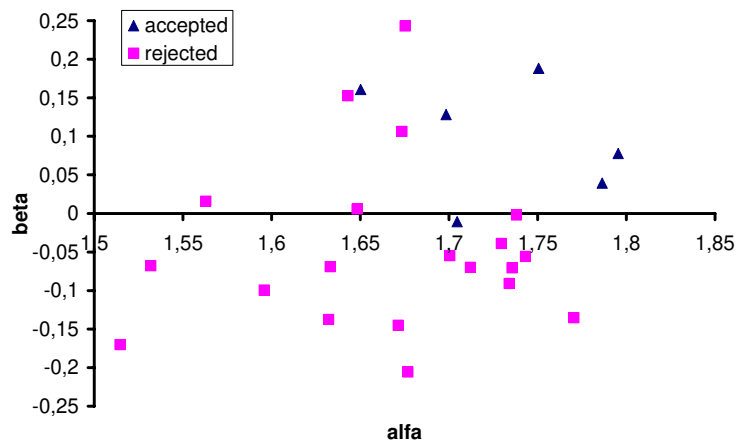

Figure 7. Distribution of α and β (*developed markets*).

Now we will verify two hypotheses: the first one – H_0^1 is our sample (with empirical mean $\hat{\mu}$ and empirical variance $\hat{\sigma}$) distributed by the Gaussian distribution. The second – H_0^2 is our sample (with parameters α, β, μ and σ) distributed by the stable distribution. Both hypotheses are examined by two criteria: the Anderson–Darling (A-D) method and Kolmogorov–Smirnov (K-S) method. The first criterion is more sensitive to the difference between empirical and theoretical distribution functions in far quantiles (tails), in contrast to the K-S criterion that is more sensitive to the difference in the central part of distribution.

The A-D criterion rejects the hypothesis of Gaussianity in all cases with the confidence level of 5%. Hypotheses of stability fitting were rejected only in 15 cases out of 27, but the values of criteria, even in the rejected cases, are better than that of the Gaussian distributions.

To prove the stability hypothesis, other researchers applied the method of infinite variance, because non–Gaussian stable r.vs has infinite variance. The set of empirical variances S_n^2 of the random variable X with infinite variance diverges.

Let $x_1, \ldots x_n$ be a series of i.i.d.r.vs X. Let $n \leq N < \infty$ and $\bar{x}_n$ be the mean of the first n observations, $S_n^2 = \dfrac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2$, $1 \leq n \leq N$. If a distribution has finite variance, then there exists a finite constant $c < \infty$ such that $\dfrac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 \to c$ (almost surely), as $n \to \infty$. And vice versa, if the series is simulated by the non–Gaussian stable law, then the series S_n^2 diverges. Fofack has applied this assumption to a series with finite variance (standard normal, Gamma) and with infinite variance (Cauchy and totally skewed stable). In the first case, the series of variances converged very fast and, in the second case, the series of variances oscillated with a high frequency, as $n \to \infty$. Fofack and Nolan applied this method in the analysis of distribution of Kenyan shilling and Morocco dirham exchange rates in the black market. Their results allow us to affirm that the exchange rates of those currencies in the black market change with infinite variance, and even worse – the authors state that distributions of parallel exchange rates of some other countries do not have the mean ($\alpha < 1$ in the stable case). We present, as an example, a graphical analysis of the variance process of Microsoft corporation stock prices returns (Figure 8).
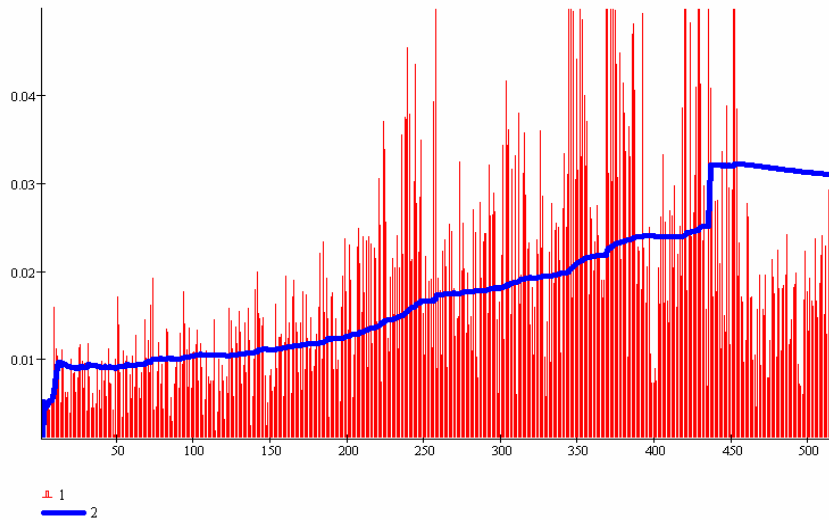
Figure 8. Series of empirical variance of the MICROSOFT company (13-03-86 – 27-05-05)

The columns in this graph show the variance at different time intervals, the solid line shows the series of variances S_n^2. One can see that, as n increases, i.e. $n \to \infty$, the series of empirical variance S_n^2 not only diverges, but also oscillates with a high frequency. The same situation is for mostly all our data sets presented.

16

Stability by homogeneity of the data series and aggregated series

The third method to verify the stability hypothesis is based on the fundamental statement. Suppose we have an original financial series (returns or subtraction of logarithms of stock prices) X_1, $X_2,\ldots,X_n$. Let us calculate the partial sums $Y_1,Y_2,\ldots,Y_{[n/d]}$, where $Y_k = \sum\limits_{i=(k-1)\cdot d+1}^{k\cdot d} X_i$, $k=1\ldots[n/d]$, and d is the number of sum components (freely chosen). The fundamental statement implies that original and derivative series must be homogeneous. Homogeneity of original and derivative (aggregated) sums was tested by the Smirnov and Anderson criteria (ω^2).

The accuracy of both methods was tested with generated sets, that were distributed by the uniform R(-1,1), Gaussian N(0,$1/\sqrt{3}$), Cauchy C(0,1) and stable $S_{1.75}(1,0.25,0)$ distributions. Partial sums were scaled, respectively, by $\sqrt{d}$, $\sqrt{d}$, d, $d^{1/1.75}$. The test was repeated for a 100 times. The results of this modeling show that the Anderson criterion (with confidence levels 0.01, 0.05 and 0.1) is more precise than that of Smirnov with the additional confidence level (for details see Appendix A).

It should be noted that these criteria require large samples (of size no less than 200), that is why the original sample must be large enough. The best choice would be if one could satisfy the condition $n/d>200$.

The same test was performed with real data of that developed and emerging markets, but homogeneity was tested only by the Anderson criterion. Partial series were calculated by summing $d=10$ and 15 elements and scaling with $d^{1/\alpha}$.

One may draw a conclusion from the fundamental statement that for international indexes ISPIX, AMEX, BP, FCHI, COCA, GDAXI, DJC, DJ, DJTA, GE, GM, IBM, LMT, MCD, MER, MSFT, NIKE, PHILE, S&P and SONY the hypothesis on stability is acceptable.

Self–similarity and multifractality

As mentioned before, for a long time it has been known that financial series are not properly described by normal models. Due to that, there arises a hypothesis of fractionallity or self–similarity. The Hurst indicator (or exponent) is used to characterize fractionallity. The process with the Hurst index $H=\frac{1}{2}$ corresponds to the Brownian motion, when variance increases at the rate of $\sqrt{t}$, where t is the amount of time. Indeed, in real data this growth rate (Hurst exponent) is higher. As $0.5<H\leq1$, the Hurst exponent implies a persistent time series characterized by long memory effects, and when $0\leq H<0.5$, it implies an anti-persistent time series that covers less distance than a random process. Such behavior is observed in mean – reverting processes.

There are a number of different, in equivalent definitions of self-similarity. The standard one states that a continuous time process $Y = \{Y(t), t \in T\}$ is self-similar, with the self-similarity parameter H (Hurst index), if it satisfies the condition:

$$Y(t) \overset{d}{=} a^{-H} Y(at), \quad \forall t \in T, \forall a > 0, 0 \leq H < 1, \text{ (5)}$$

where the equality is in the sense of finite-dimensional distributions. The canonical example of such a process is Fractional Brownian Motion ($H=\frac{1}{2}$). Since the process Y satisfying (5) can never be stationary, it is typically assumed to have stationary increments.

Figure 9 shows that stable processes are the product of a class of self-similar processes and that of Levy processes. Suppose a Levy process $X=\{X(t), t\geq0\}$. Then X is self-similar if and only if each $X(t)$ is strictly stable. The index α of stability and the exponent H of self-similarity satisfy $\alpha=1/H$.
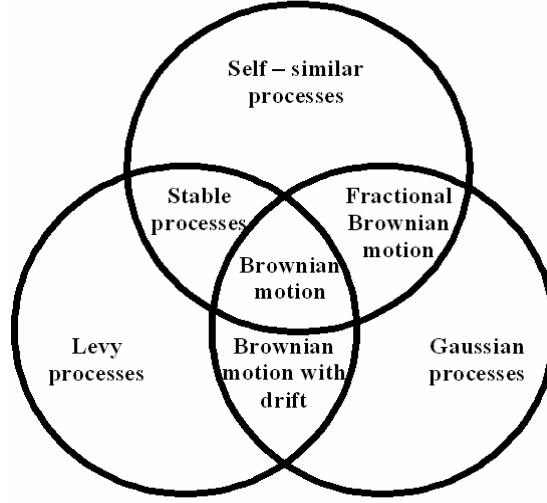
Figure 9. Self-similar processes and their relation to Levy and Gaussian processes

Consider the aggregated series $X^{(m)}$, obtained by dividing a given series of length N into blocks of length m and averaging the series over each block.

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i \text{, here } k=1, 2\ldots[N/m].$$

Self-similarity is often investigated not through the equality of finite-dimensional distributions, but through the behavior of the absolute moments. Thus, consider

$$AM^{(m)}(q) = E\left|\frac{1}{m}\sum_{i=1}^{m} X(i)\right|^q = \frac{1}{m}\sum_{k=1}^{m}\left|X^{(m)}(k) - \overline{X}\right|^q$$

If X is self-similar, then $AM^{(m)}(q)$ is proportional to $m^{\beta(q)}$, which means that $\ln AM^{(m)}(q)$ is linear in $\ln m$ for a fixed q:

$$\ln AM^{(m)}(q) = \beta(q)\ln m + C(q). \qquad (6)$$

In addition, the exponent $\beta(q)$ is linear with respect to q. In fact, since $X^{(m)}(i) \overset{d}{=} m^{1-H} X(i)$, we have

$$\beta(q) = q(H-1) \qquad (7)$$

Thus, the definition of self-similarity is simply that the moments must be proportional as in (6) and that $\beta(q)$ satisfies (7).

This definition of a self-similar process given above can be generalized to that of multifractal processes. A non-negative process $X(t)$ is called multifractal if the logarithms of the absolute moments scale linearly with the logarithm of the aggregation level m. Multifractals are commonly constructed through multiplicative cascades. If a multifractal can take positive and negative values, then it is referred to as a signed multifractal (the term "multiaffine" is sometimes used instead of "signed multifractal"). The key point is that, unlike self-similar processes, the scaling exponent $\beta(q)$ in (6) is not required to be linear in q. Thus, signed multifractal processes are a generalization of self-similar processes. To discover whether a process is (signed) multifractal or self-similar, it is not enough to examine the second moment properties. One must analyze higher moments as well.

However this method is only graphical and linearity is only visual.

Finally, only 9 indices are self-similar: ISPX, AMEX, FCHI, GDAXI, DJC, DJ, DJTA, NIKKEI, S&P.

Hurst exponent estimation. There are many methods to evaluate this index, but in literature the following are usually used:

a. Time-domain estimators,
b. Frequency-domain/wavelet-domain estimators,

The methods: absolute value method (absolute moments), variance method (aggregate variance), R/S method and variance of residuals are known as time domain estimators. Estimators of this type are based on investigating the power law relationship between a specific statistic of the series and the so-called aggregation block of size *m*.

The following three methods and their modifications are usually presented as time-domain estimators:

- Periodogram method;
- Whittle;
- Abry-Veitch (AV).

The methods of this type are based on the frequency properties of wavelets.

All Hurst exponent estimates were calculated using SELFIS software, which is freeware and can be found on the web page http://www.cs.ucr.edu/~tkarag.

Multifractality and self-similarity in the financial markets

In the case of The Baltic States and other Central and Eastern Europe financial markets, the number of daily zero returns can reach 89%. Anyway, this problem may be solved by extending a continuous model to the mixed one, where daily returns equal to zero are excluded from the series when estimating the stability parameters. The series of non-zero returns are fitted to the stable distribution. Stable parameters are estimated by the maximal likelihood method. Goodness-of-fit is verified by the Anderson-Darling distributional adequacy test. The stability is also tested by the homogeneity test, based on the fundamental property of stable laws. The summation scheme is based on the bootstrap method in order to get larger series. Multifractality and self-similarity are investigated through the behavior of the absolute moments. The Hurst analysis has been made by the R-S method.

We have investigated 26 international financial series focusing on the issues of stability, multifractality, and self-similarity. It has been established that the hypothesis of stability was ultimately rejected in 14.81% cases, definitely stable in 22.22%, and the rest are doubtful. It is important to note that, even in the case of rejection, the value of the A-D criterion was much better for stability testing than for the test of Gaussian distribution. No series was found distributed by the Gaussian law.

The stable model parameters were estimated by the maximal likelihood method. The stability indexes of stable series are concentrated between 1.65 and 1.8, which confirms the results of other authors that the stability parameter of financial data is over 1.5. Asymmetry parameters are scattered in the area between -0.017 and 0.2.

The investigation of self-similarity has concluded that only 66.67% of the series are multifractal and the other 33.33% concurrently are self-similar.

The Hurst analysis has showed that the methods of R/S and Variance of Residuals are significant in the stability analysis. Following these two methods, Hurst exponent estimates are in the interval $H \in (0.5; 0.7)$, which means that the stability index $\alpha \in (1.42 ; 2)$. If the Hurst exponent is calculated by the R/S method, $H \in (0.5; 0.6)$, then $\alpha \in (1.666 ; 2)$.

The stable models are suitable for financial engineering; however the analysis has shown that not all (only 22% in our case) the series are stable, so the model adequacy and other stability tests are necessary before model application. The studied series represent a wide spectrum of stock market, however it should be stressed that the research requires a further continuation: to extend the models.

The analysis of stability in the Baltic States market has showed that 49 series of 64 are multifractal and 8 of them are also self-similar. If we removed zero returns from the series, there would be 27 multifractal series, concurrently 9 of them are self-similar.

Relationship measures

In constructing a financial portfolio it is essential to determine relationships between different stock returns. However, under the assumption of stability (sets of stock returns are modeled by

stable laws), the classical relationship measures (covariance, correlation) cannot be applied. Therefore the generalized Markowitz problem is solved by generalized relationship measures (covariation, codifference). We show that implementation of the codifference between different stocks greatly simplifies the construction of the portfolio. We have constructed optimal portfolios of ten Baltic States stocks.

In the classical economic statistics (when the distributional law has two first moments, i.e., mean and variance), relations between two random variables (returns) are described by covariance or correlation. But if we assume that financial data follow the stable law (empirical studies corroborate this assumption), covariance and especially correlation (Pearson) cannot be calculated. In case when the first ($\alpha < 1$) and the second ($\alpha < 2$) moments do not exist, other correlation (rank, e.g., Spearmen, Kendall, etc.) and contingency coefficients are proposed. However, in the portfolio selection problem Samorodnitsky and Taqqu suggest better alternatives, even when mean and variance do not exist. They have proposed alternative relation measures: covariation and codifference.

If X_1 and X_2 are two symmetric i.d. (with $\alpha_1 = \alpha_2 = \alpha$) stable random variables, then the covariation is equal to

$$[X_1, X_2]_\alpha = \int_{S_2} s_1 s_2^{\langle \alpha-1\rangle} \Gamma(ds),$$

where $\alpha > 1$, $y^{\langle \alpha \rangle} = |y|^\alpha \operatorname{sign}(\alpha)$ and Γ is a spectral measure of (X_1, X_2).

In such a parameterization, the scale parameter $\sigma_{X_1}^\alpha$ of symmetric stable r.v. can be calculated from $[X_1, X_1]_\alpha = \sigma_{X_1}^\alpha$. If $\alpha = 2$ (Gaussian distribution), the covariation is equal to half of the covariance $[X_1, X_2]_2 = \frac{1}{2}\operatorname{Cov}(X_1, X_2)$ and $[X_1, X_1]_2 = \sigma_{X_1}^2$ becomes equal to the variance of X_1. However, the covariation norm of $X \in S_\alpha$ ($\alpha > 1$) can be calculated as $\|X\| = ([X, X]_\alpha)^{1/\alpha}$. If $X \sim S_\alpha(\sigma, 0, 0)$ (SαS case), then the norm is equivalent to the scale parameter of the stable distribution $\|X\|_\alpha = \sigma$.

In general case the codifference is defined through characteristic functions
$$\operatorname{cod}_{X,Y} = \ln(E \exp\{i(X - Y)\}) - \ln(E \exp\{iX\}) - \ln(E \exp\{-iY\})$$

$$= \ln\left(\frac{E \exp\{i(X - Y)\}}{E \exp\{iX\} \cdot E \exp\{-iY\}}\right) = \ln\left(\frac{\phi_{X-Y}}{\phi_X \cdot \phi_{-Y}}\right),$$

or empirical characteristic functions

$$\operatorname{cod}_{X,Y} = \ln\left(\frac{n \cdot \sum_{j=1}^{n} e^{i(X_j - Y_j)}}{\sum_{j=1}^{n} e^{iX_j} \cdot \sum_{j=1}^{n} e^{-iY_j}}\right)$$

The codifference of two symmetric (SαS) r.vs X and Y ($0 < \alpha \le 2$) can be expressed through the scale parameters

$$\operatorname{cod}_{X,Y} = \|X\|_\alpha^\alpha + \|Y\|_\alpha^\alpha - \|X - Y\|_\alpha^\alpha$$

If $\alpha = 2$, then $\operatorname{cod}_{X,Y} = \operatorname{Cov}(X, Y)$.

Samorodnitsky and Taqqu have showed that

$$\left(1 - 2^{\alpha-1}\right)\left(\|X\|_\alpha^\alpha + \|Y\|_\alpha^\alpha\right) \le \operatorname{cod}_{X,Y} \le \|X\|_\alpha^\alpha + \|Y\|_\alpha^\alpha,$$

here $1 \le \alpha \le 2$, and, if we normalize (divide by $\|X\|_\alpha^\alpha + \|Y\|_\alpha^\alpha$), we will get a generalized correlation coefficient.

In the general case, the following inequalities

$$\left(1-2^{\alpha-1}\right)\ln\left(\frac{1}{E\exp\{iX\}\cdot E\exp\{-iY\}}\right)\le cod_{X,Y}=\ln\left(\frac{E\exp\{i(X-Y)\}}{E\exp\{iX\}\cdot E\exp\{-iY\}}\right)\le\ln\left(\frac{1}{E\exp\{iX\}\cdot E\exp\{-iY\}}\right)$$

are proper, and if we divide both sides by $\ln(E\exp\{iX\}\cdot E\exp\{-iY\})$, we will get the following system of inequalities for the correlation coefficient

$$\left(1-2^{\alpha-1}\right)\le corr_{X,Y}=\frac{\ln\left(\dfrac{E\exp\{iX\}\cdot E\exp\{-iY\}}{E\exp\{i(X-Y)\}}\right)}{\ln(E\exp\{iX\}\cdot E\exp\{-iY\})}\le1$$

If $0<\alpha\le1$ this correlation coefficient is only non-negative, and if $\alpha=2,\beta=0$, then $-1\le corr_{X,Y}=\rho_{X,Y}\le1$ is equivalent to the Pearson correlation coefficient.

Significance of codifference

The significance of the Pearson correlation coefficient is tested using Fisher statistics, and that of Spearmen and Kendall coefficients, respectively, are tested using Student and Gaussian distributions. But likely that there are no codifference significance tests created. In such a case, we use the bootstrap method (one of Monte-Carlo style methods). The following algorithm to test the codifference significance is proposed:

1) Estimate stable parameters (α, β, σ and μ) and stagnation probability p of all equity returns series;

2) Estimate relation matrix of measure ρ (covariation or codifference) for every pair of equities series;

3) Test the significance of each ρ_{ij} by the bootstrap method:

   i) generate a pair of two ith and jth mixed-stable (with estimated parameters) series, and proceed to the next step;

   ii) calculate the kth relation measure ρ_{ij}^{k} , between the ith and jth series;

   iii) repeat (i) and (ii) steps for $k=1,...,N$ (for example, 10000) times;

   iv) construct ordered series of estimates $\rho_{ij}^{(k)}$;

   v) if $\rho_{ij}^{([N\cdot0.025])}\le\rho_{ij}\le\rho_{ij}^{([N\cdot0.975])}$, then the significance of ρ_{ij} is rejected with the confidence level 0.05, i.e., it is assumed that $\rho_{ij}=0$.

   vi) repeat 3i—3v steps for each pair of equities i and j.

Covariation and codifference are calculated for ten equities with the longest series (MNF1L, LDJ1L, VNF1R, NRM1T, MKO1T, GZE1R, ETLAT, VNG1L, SNG1L, TEO1L). The correlation tables are presented for the series of equalized length 1427.

However, in portfolio the theory covariance (or equivalent measure) is more useful, since in that case, there is no need to know the variance. The generalized covariance tables are calculated for previously mentioned series.

Portfolio construction

Let us construct a portfolio from 10 following equities of the Baltic States market: MNF1L, LDJ1L, VNF1R, NRM1T, MKO1T, GZE1R, ETLAT, VNG1L, SNG1L and TEO1L. To select the optimal portfolio, we can solve the optimization problem presented by Markowitz, but replacing the covariance by codifference.

minimize

$$\lambda\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_i\omega_j\sigma_{ij}-(1-\lambda)\sum_{i}^{n}\varpi_i\mu_i$$

for $\omega_i, i = 1, ..., n,$ and conditions

$$\sum_i^n \varpi_i = 1, \tag{8}$$

$$\omega_i > 0, \quad i = 1, ..., n,$$

where ω_i is the weight of the ith equity of portfolio, μ_i – expected return of the ith equity (in the case of α-stable it is replaced by parameter μ), σ_{ij} is the covariance between the ith and jth equity (codifference), $\lambda \in [0,1]$ is the optimization constant (1/2). The first constituent of this objective function characterizes the risk of portfolio and the second one characterizes expected return.

To solve the portfolio optimization problem, a utility function should be introduced. The most recently used ones are: VAR (value at risk), CVAR (conditional value at risk), MAD – (mean absolute deviation) and Markowitz (mean – variance). Since in the case of the α-stable model, variance (second moment) does not exist, we use such a generalization of the Markowitz utility function. Suppose that returns (investment profit) of the n-stock portfolio are calculated in such a way:

$$r(w, \xi) = \sum_{i=1}^n w_i \cdot \xi_i, \text{ where } \sum_{i=1}^n w_i = 1, \ w_i \geq 0, \ 1 \leq i \leq n,$$

Since in the case of the α-stable model, variance and standard deviation as well as covariance (except $\alpha=2$) do not exist the portfolio selection problem may be formulated by the following objective function:

$$F(w) = c \cdot \left(\frac{1}{N} \sum_{j=1}^N \left| \sum_{i=1}^n w_i \left(X_{j,i} - \mu_i \right) \right|^\gamma \right)^{1/\gamma} - \sum_{i=1}^n w_i \mu_i$$

where $\gamma = \min(\alpha_i)$, $c = 1/\gamma$, μ_i is expected return of the ith equity (or μ), $X_{j,i}$ is return of the ith share at the jth moment. In such a case, the following optimization problem must be solved:

minimize

$F(w)$

for $w_i, i = 1, ..., n,$ and conditions $\tag{9}$

$$\sum_i^n w_i = 1,$$

$$w_i > 0, \quad i = 1, ..., n.$$

If we add the following constraint to systems (8) and (9) $\sum_{i=1}^n w_i \cdot \mu_i = \mu_P$ we will get portfolios with the fixed return μ_P.

For example, given equities, we can construct optimal portfolios with the following weights (see Table 3).

Table 3. Optimal solutions for systems (8) and (9).

Equity	weights of system (8)	weights of system (9)
TEO1L	0.01931127121831664739	0.00000458324159077873
SNG1L	0.14748791669745869859	0.20832600703143941412
VNG1L	0.06520432421240375531	0.07441099228178453540
ETLAT	0.02316874195215026799	0.00877015050036948456
GZE1R	0.06396124707815362131	0.26926385550778603184
MKO1T	0.00856555150906847772	0.12372356319707422667
NRM1T	0.13592436635025012537	0.04210572877402151554
VNF1R	0.28168613842527678859	0.11068516003731768138
LDJ1L	0.11707050740499311270	0.07642465720903288129
MNF1L	0.13761993515192852411	0.08628530221958331803

It is noteworthy that the portfolio gives a higher profit by 29 September 2007.

In the case of a multishare portfolio (*n*>2), together with weight optimization, a diversification problem must be solved. When *n*>2 some weights might shrink to zero and such an investment looses the sense and sometimes even becomes impossible (one has to buy half a share, etc.).

Chapter 4 presents a generalization of the obtained results and conclusions.

1. Parameter estimation methods and software has been developed for models with asymmetric stable distributions. The efficiency of estimation methods was tested by simulating the series. Empirical methods are more effective in time, but the maximal likelihood method (MLM) is more effective (for real data) in the sense of accuracy (Anderson-Darling goodness-of-fit test corroborate that). It should be noted that MLM is more sensitive to changes of the parameters α and σ.

2. Empirical parameters of the Baltic States series and developed market series (respectively 64 and 27 series) have been estimated. Most of the series are very asymmetric ($0.1<|\gamma_1|<30$), and the empirical skewness ($\gamma_2 \neq 0$) suggests that the probability density function of the series is more peaked and exhibits fatter tails than the Gaussian one. The normality hypothesis is rejected by the Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit tests.

3. Distribution of the stability parameter α and asymmetry parameter β in the series of developed markets shows, that usually $1.5<\alpha<2$ and the parameter β is small. Distribution of the stability parameter in the series of the Baltic States market (full series) shows that usually α is lower than 1.5 and close to 1. But if we remove the zero returns from the series, the parameter α is scattered near 1.5, while the parameter β is small usually, but positive.

4. An experimental test of the series homogeneity shows that for the stable series with asymmetry, the Anderson test is more powerful than the Smirnov one. The Anderson test for 27 series from the developed markets shows that 21 series are homogeneous with their aggregated series and only 2 series (64 at all) from the Baltic States market (and only when the zero returns are removed) are homogeneous with their aggregate series (they do not obey the fundamental stable theorem).

5. The analysis of self-similarity and multifractality, by the absolute moments method, indicates that all 27 series (from the developed markets) are multifractal and concurrently 9 of them are self-similar. On the other hand, 49 series (from the Baltic States market) are multifractal and 8 of them are also self-similar, but if we remove the zero returns from the series, then remain only 27 multifractal and 9 self-similar series. This is because the series becomes too short for multifractality analysis.

6. A mixed stable model of returns distribution in emerging markets has been proposed. We introduced the probability density, cumulative density, and the characteristic functions. Empirical results show that this kind of distribution fits the empirical data better than any other. The Baltic States equity lists are given as an example.

7. The implementation of the mixed-stable model is hampered by the lack of goodness-of-fit tests for discontinuous distributions. Since adequacy tests for continuous distribution functions cannot be implemented, the tests based on the empirical characteristic function (Koutrouvelis) as well as modified χ^2, are used. The experimental tests have showed that, if the stability parameter α and the number of zero returns are increasing, then the validity of the tests is also increasing. 99% of the Baltic States series satisfy the mixed stable model proposed (by the Koutrouvelis test).

8. The statistical analysis of the Baltic States equity stagnation intervals has been made. Empirical studies showed that the length series of the state runs of financial data in emerging markets are better described by the Hurwitz zeta distribution, rather than by geometrical. Since series of the lengths of each run are not geometrically distributed, the state series must have some internal dependence (Wald-Wolfowitz runs test corroborates this assumption). A new mixed-stable model with dependent states has been proposed and the formulas for probabilities

of calculating states (zeros and units) have been obtained. Adequacy tests of this model are hampered by inner series dependence.

9. The inner series dependence was tested by the Hoel criterion on the order of the Markov chain. It has been concluded that there are no zero order series or Bernoulli scheme series. 95% of given series are 4th-order Markov chains with φ=0.1% significance level.

10. When constructing an optimal portfolio, it is essential to determine possible relationships between different stock returns. However, under the assumption of stability (stock returns are modeled by mixed-stable laws) traditional relationship measures (covariance, correlation) cannot be applied, since $(1,27 < \alpha < 1,78)$. In such a case, covariation (for asymmetric r.v.) and codifference are offered. The significance of these measures can be tested by the bootstrap method.

11. A wide spectrum of financial portfolio construction methods is known, but in the case of series stability it is suggested to use a generalized Markowitz model. The problem is solved by the generalized relationship measures (covariation, codifference). Portfolio construction strategies with and without the codifference coefficient matrix are presented. It has been shown that the codifference application considerably simplifies the construction of the optimal portfolio. Optimal stock portfolios (with 10 most realizable Baltic States stocks) with and without the codifference coefficient matrix are constructed.

12. The demo version of developed software can be found at *www.kabasinskas.tk*. This application allows us to simulate stable r.vs, estimate the stable parameters and to construct a portfolio (up to 10 selected shares). The experimental test has proved that the parameters estimation time nonlinearly depends on the series lengths.

The list of references and annexes are presented at the end of the dissertation.

2   Published scientific works on the dissertation topic

1. A. Kabašinskas, L. Sakalauskas. A system for financial modelling, *Informacion Scienses* vol. 27, ISSN 1392-0561, Vilnius University, p. 115–120, 2003.

2. A. Kabašinskas, I. Belovas, L. Sakalauskas. Estimation of stable models by maximal likelihood and robust methods, In: S. Aivazian et al. (eds), *Proceedings of the Seventh International Conference „Computer Data Analysis and Modeling, Robustness and Computer Intensive Methods"*, Vol. 1. ISBN 985-445-492-4, Minsk, p. 68–73, 2004.

3. I. Belov, A.Kabašinskas, L.Sakalauskas. A Study of Stable Models of Stock Markets, *Information Technology And Control*, vol. 35, No. 1, ISSN 1392-124X, p. 34–56, 2006.

4. I. Belovas, A. Kabasinskas ir L. Sakalauskas. Multifractality and self-similarity in stock markets, *Matematics ir mathematical modeling* 2, ISSN 1822-2757, p. 6–11, 2006.

5. I. Belovas, A. Kabasinskas ir L. Sakalauskas. Analysis of passivity problems in the Baltic equity market, *Lithuanian Mathematical Journal* 46 (spec. nr.), ISSN 1392-0561, p. 289–294, 2006.

6. I. Belovas, A. Kabasinskas ir L. Sakalauskas. On covariation and codifference in optimal portfolio construction, *Informacion Sciences* vols 42–43. ISSN 1392-0561, Vilnius University, p. 182–188, 2007.

Published conference proceedings on the dissertation topic

7. A. Kabasinskas, L. Sakalauskas. On stock portfolio simulation and optimization, In: H. Pranevicius at al. (eds), *Proceedings of Intern. Conference "Modelling and Simulation of Business Systems"*. ISBN 9955-09-420-6. Kaunas, KTU Press „Technologija", p. 232–234, 2003.

8. I. Belovas, A. Kabasinskas and L. Sakalauskas. Returns modelling problem in the Baltic equity market, In: H. Pranevičius et al. (eds), *Proceedings of the 5th International Conference on Operational Research: Simulation and Optimization in Business and Industry*, ISBN 9955-25-061-5, KTU Press „Technologija", p. 3–8, 2006.

9. I. Belovas, A. Kabasinskas and L. Sakalauskas. On the modelling of stagnation intervals in emerging stock markets. In: S. Aivazian et al. (eds), *Proceedings of the Eighth International*

3   About the author

Audrius Kabašinskas was born on the 29th of December 1977 in Kazlų Rūda, Marijampolė district.

He obtained an Applied Mathematics Bachelor degree at the faculty of Fundamental sciences, Kaunas University of Technology in 2000. A master degree of Applied Mathematics was obtained at the Faculty of Fundamental sciences, Kaunas University of Technology, 2002.

During the period of 2002–2003 A. Kabašinskas was a junior researcher of the Institute of Mathematics and Informatics.

During the period 2002–2007 A. Kabašinskas was an assistant (lecturer) at the Informatics Technology department, Vilnius Gediminas Technical University.

During the period of 2003–2007 A. Kabašinskas was a PhD student of the Institute of Mathematics and Informatics.

Since 2002 he worked as an assistant (in 2007 he became a lecturer) at the department of Mathematical Research in Systems, Kaunas University of Technology.

Email: audrius.kabasinskas@ktu.lt

Reziume

Darbo sritis

Finansinių procesų modeliavimas ir analizė yra svarbi, labai greitai besivystanti mokslo šaka, apjungianti tokias sritis kaip kompiuterių mokslai, taikomoji matematika, statistika ir ekonomika. Tiriant investavimo galimybes ir nagrinėjant finansines laiko sekas, jų elgesiui aprašyti taikomi tikimybiniai–statistiniai modeliai. Adekvatus empirinio duomenų pasiskirstymo nustatymas ženkliai lemia daugelį prognozavimo bei investavimo sprendimų. Daugelio tyrimų rezultatai rodo, kad finansiniai duomenys yra su dideliais nuokrypiais bei pasižymi asimetrija ir leptokurtotiškumu. Tokių duomenų skirstiniams modeliuoti mokslinėje literatūroje pastaruoju metu pasiūlyti stabilieji modeliai.

Šio darbo tyrimų sritis – finansinių rodiklių laiko sekų tikimybinių modelių sudarymo bei prognozavimo ir investavimo algoritmų metodologija.

Darbo aktualumas

Finansinės rinkos ir jose cirkuliuojantys finansiniai srautai yra svarbūs šiuolaikinei ekonomikai. Efektyvių investicinių sprendimų technologijų kūrimas yra aktuali informacinių technologijų sritis. Naujos investavimo galimybės atsivėrė po Europos Sąjungos išsiplėtimo 2004 metais. Iki tol mažai žinomos Baltijos ir kitų Rytų bei Centrinės Europos rinkos tapo patraukliomis investuotojams. Spartus 3–8 % Bendrojo Vidaus Produkto (BVP) augimas (tuo tarpu senosiose ES šalyse tik 1,5–1,8%) ir aukštas pelningumo lygis pritraukė nemažai naujų rinkos dalyvių. Visgi šių rinkų ypatybės kol kas nėra plačiai ištirtos.

Šiame darbe analizuojamos pasaulio ir Baltijos šalių akcijų kainų grąžų kitimas (stebimas vienos dienos intervalu). Baltijos šalių finansinės sekos pasižymi dviem svarbiomis savybėmis, kurios nebūdingos JAV ar senųjų ES narių rinkose cirkuliuojančių vertybinių popierių sekoms:

- sekos yra žymiai trumpesnės: neviršija 2000 stebėjimų (10–12 metų), iš kurių tik 1000–1700 yra tinkami analizei;
- empiriniuose duomenyse stebimas stagnacijos fenomenas (1993–2007 m.). Stagnacija charakterizuojama ypač dideliu pasyvumu, kai tam tikrą laiko tarpą akcijų kainos nesikeičia, nes biržoje šiomis akcijomis neįvyksta nei vieno sandorio.

Realūs duomenys neretai pasižymi asimetrija, ekscesu ir dideliais nuokrypiais. Tai patvirtina nemažai empirinių tyrimų, pavyzdžiui, Janicki ir Weron, Rachev, Samorodnitsky ir Taqqu. Taigi modeliai, paremti duomenų normalumo sąlyga, yra netinkami ir pastaroji sąlyga yra pakeičiama stabilumo sąlyga. Stabilieji dėsniai pasižymi tiek leptokurtotiškumu, tiek asimetrija ir dažnai geriau nei Gauso skirstinys tinka empirinių finansinių duomenų aprašymui. Be to, stabilieji atsitiktiniai dydžiai paklūsta Apibendrintai Centrinei Ribinei Teoremai (ACRT). Ši teorema teigia, kad stabilieji dėsniai yra vieninteliai atitinkamai centruotų ir normuotų bei nepriklausomų ir vienodai pasiskirsčiusių atsitiktinių dydžių sumų asimptotiniai skirstiniai. Kai kurių specialistų nuomone, α-stabilieji dėsniai yra bene geriausia alternatyva Gauso dėsniui iš visų skirstinių, kurie buvo pasiūlyti mokslinėje literatūroje per pastaruosius keturis dešimtmečius.

Ekonominiams ir finansiniams procesams aprašyti dažniausiai taikomi modeliai pagrįsti Gauso skirstiniu (Brauno judesiu). Kadangi pastaruoju metu normaliniai modeliai vertinami kritiškai, iškeliamos šių procesų fraktališkumo arba savastingumo (angl. self-similarity) ir kitokios hipotezės. Savastingumui apibūdinti yra naudojamas Hurst'o rodiklis (angl. Hurst exponent), kurio reikšmė ½ atitinka Brauno judesį. Tačiau tiriant realius duomenis dažnai paaiškėja, kad šis rodiklis yra kitoks.

Baltijos šalių rinkoje "nulinės grąžos problema" yra rimtesnė nei gali pasirodyti iš pirmo žvilgsnio. Baltijos ir kitų Centrinės bei Rytų Europos šalių finansų rinkos yra gana naujos, besivystančios, o kai kurie finansiniai instrumentai yra mažai likvidūs. Neretai besivystančiose rinkose yra stebimas pasyvumo (stagnacijos) efektas. Nulinių grąžų skaičius kartais gali siekti 89% visų stebėjimų, todėl svarbu sukurti ir išanalizuoti tinkamus metodus tokio tipo rinkoms tirti.

Kkompleksinis stabiliosios hipotezės tyrimas yra susijęs su daugeliu efektų, kuriems tirti programinės įrangos nėra sukurta, o standartinės programinės įrangos skirtos statistinei analizei atlikti, ne visada pakanka. Taigi programinės įrangos kūrimas taip pat yra aktuali problema.

Tyrimo objektas

Šio tyrimo objektas yra Baltijos ir pasaulio finansinių rinkų vertybinių popierių istoriniai duomenys ir jų modeliavimo metodai

Darbo tikslas ir uždaviniai

Darbo tikslas – sukurti ir ištirti metodus besivystančių vertybinių popierių rinkos analizei ir finansinio portfelio valdymui, remiantis finansinių sekų stabilumo hipoteze.
Siekiant šio tikslo buvo sprendžiami tokie uždaviniai:
1. Sudaryti finansinių duomenų laiko sekų analizės metodus ir pritaikyti juos Baltijos šalių vertybinių popierių laiko sekų analizei;
2. Aprašyti, išanalizuoti ir eksperimentiškai ištirti metodus stabilumo hipotezei tirti:
   - patikrinti dalinių sekų ir pradinių imčių homogeniškumo bei stabiliųjų dėsnių suderinamumo neparametrines hipotezes;
   - ištirti finansinių duomenų sekų savastingumą ir multifraktališkumą;
   - sukurti specialią programinę įrangą reikalingą išvardintiems statistiniams metodams realizuoti.
3. Ištirti pasyvumo (stagnacijos) efektą besivystančiose rinkose:
   - sukurti teorines prielaidas rinkos pasyvumui aprašyti;
   - ištirti mišriojo stabiliojo modelio pritaikymo galimybes finansinėms sekoms modeliuoti;
   - eksperimentiškai išbandyti parametrų vertinimo metodus ir suderinamumo testų patikimumą, esant trūkiai tikimybinio tankio ir pasiskirstymo funkcijoms.
4. Sukurti programinę sistemą vertybinių popierių rinkos analizei ir finansinių išteklių planavimui bei išbandyti demonstracinę internetinės programinės įrangos versiją.

Mokslinis naujumas

Darbe gauti tokie nauji rezultatai:
1) Stabiliųjų modelių su asimetrija parametrų vertinimui sudaryta programinė įranga, realizuojanti didžiausio tikėtinumo (MTM) ir robastinius metodus bei ištirtas šių metodų efektyvumas.
2) Kompleksiškai ištirti statistiniai metodai ir sukurta juos realizuojanti programinė įranga, reikalinga stabilumo prielaidai pagrįsti:
   a) patikrintos neparametrinės hipotezės apie imties skirstinio funkcijos suderinamumą su stabiliojo a.d. skirstiniu;
   b) patikrintos neparametrinės hipotezės apie dalinių sekų homogeniškumą su visa seka;
   c) atlikti teoriniai ir praktiniai pasaulio ir Baltijos šalių rinkų duomenų sekų savastingumo ir multifraktališkumo tyrimai;
3) Finansinėms sekoms modeliuoti pritaikytas atskiras paslėptųjų Markovo grandinių modelio atvejis – mišrusis stabilusis modelis.
4) Ištirti stagnacijos laikotarpių trukmės pasiskirstymai, kuriems modeliuoti pasiūlyti Hurwitz`o zeta ir apibendrinto Hurwitz`o zeta dėsniai.
5) Ištirti metodai statistiniams ryšiams tarp atskirų finansinių instrumentų nustatyti, kai neegzistuoja antrasis momentas ir parengtos rekomendacijos šiems ryšiams nustatyti bei įvertinti.

Praktinė darbo reikšmė

Darbe gauti tokie praktiniai rezultatai:
1. Sudaryta vertybinių popierių rinkų duomenų analizės ir modeliavimo programinė sistema, naudojant universalių programavimo kalbų Visual Basic, C++ ir JAVA priemones bei programų paketo MathCad. Ši priemonė leidžia įvertinti finansinių sekų stabiliųjų skirstinių parametrus, modeliuoti juos kompiuteriu, bei sudaryti optimalų vertybinių popierių portfelį.

2. Sukurtos sistemos testavimas, pasinaudojant Baltijos šalių ir išsivysčiusių finansinių rinkų duomenimis, parodė, kad α-stabilusis bei kiti alternatyvūs ir klasikiniai dėsniai ne visada tinka akcijų kainų grąžų modeliavimui. Tuo tarpu bendresnis mišrusis–stabilusis modelis tinka beveik visada (99%). Šis pusiau tolydus dėsnis gali būti pritaikytas ir kompiuterinių tinklų informaciniams srautams modeliuoti.
3. Pasiūlyta vertybinių popierių portfelio formavimo metodai naudojant mišrųjį–stabilųjį modelį.

Disertacijos struktūra. Disertaciją sudaro keturi pagrindiniai skyriai (įvadinė dalis, analitinis tyrimas, metodologija ir rezultatai), išvados, literatūros sąrašas ir priedai. Kiekviename skyrelyje pateikiami naudotų, jei tokie buvo, algoritmų aprašymai.

1-asis skyrius yra įvadinis. Šioje dalyje pateikiama glausta informacija apie disertaciją: darbo sritis, darbo aktualumas, tyrimo objektas, darbo tikslas ir uždaviniai, mokslinis naujumas, praktinė darbo reikšmė, darbo rezultatų aprobavimas, terminų ir santrumpų sąvadas.

2-ame skyriuje pateikiama mokslinių bei taikomųjų darbų analitinė apžvalga finansų modeliavimo tema. Skyriuje nagrinėjamos šios temos: finansų inžinerijos vystymasis, klasikiniai vertybinių popierių portfelio sudarymo modeliai (vertybinių popierių portfelio sudarymas esant sandorių kaštams), analizuojami modernūs finansinių rinkų modeliai, apžvelgiami finansinių sekų stabilumo tyrimo metodai, atsižvelgiama į finansinių sekų savastingumą (angl. self-similarity) ir multifraktališkumą, pateikiama esamos programinės įrangos apžvalga, analizuojami ir charakterizuojami finansinių duomenų šaltiniai (duomenų atranka, duomenų patikrinimas, duomenų transformacija, finansinių duomenų kiekis).

3-ame skyriuje pateikiami metodai ir matematiniai modeliai, padedantys pasiekti užsibrėžtus tikslus. Stabilumo prielaidai finansinėse rinkose pagrįsti atliekamas duomenų apdorojimas ir analizė bei tiriama rinkos pasyvumo įtaka duomenims. Detaliai aprašomas stabilusis dėsnis, jo charakteristikos bei numatomi parametrų vertinimo metodai (maksimalaus tikėtinumo metodas, momentų metodas, metodai paremti empirinės charakteringosios funkcijos savybėmis ir kiti). Taip pat aptariamas ir normaliųjų bei stabiliųjų atsitiktinių procesų tinkamumas finansiniams procesams aprašyti. Dar XX amžiaus viduryje buvo pastebėta, kad finansinių sekų elgsenai būdingas savastingumas bei multifraktališkumas, todėl šiame skyriuje pateikiami metodai šioms savybėms tirti (Hurst indeksas, baigtinės dispersijos metodas, pradinės ir agreguotos sekų homogeniškumo nustatymas). Aprašomi metodai reikalingi pasyvumo efekto besivystančiose rinkose tyrimams. Pasyvumo fenomenas mokslinėje literatūroje anksčiau nebuvo nagrinėjamas, todėl aprašoma nulinių grąžų problema (NGP), kuriamas mišrusis stabilusis modelis, aprašomos jo charakteristikos, metodai parametrams vertinti bei suderinamumo testų patikimumas. Mišrusis modelis šiame darbe nagrinėjamas dveim atvejais: kai grąžų būsenos yra priklausomos ir kai jos yra atsitiktinės. Greta jau aptartų vienmačių finansinių sekų modelių, kuriami daugiamačiai modeliai, aprašomas ryšio tarp atskirų akcijų grąžų nustatymas bei jo įtaka vertybinių popierių portfelio formavimui.

4-ame skyriuje aptariami gauti rezultatai ir aprašoma sukurta programinė priemonė (C++ kalboje bei adaptuotos hipertekstinei aplinkai), o taip pat aptariamas ir daugiaprocesorinio superkompiuterio panaudojimas įvairiems daug resursų reikalaujantiems skaičiavimams.

Išvadų skyriuje formuluojamos darbo išvados:
1. Sudarius programinę įrangą, realizuojančią stabiliųjų modelių su asimetrija parametrų vertinimą didžiausio tikėtinumo ir empirinius metodais, bei ištyrus šių metodų efektyvumą laiko sąnaudų ir patikimumo prasme, nustatyta:
   a. laiko sąnaudų prasme efektyvesni yra robastiniai metodai (apie 95%)
   b. didžiausio tikėtinumo metodas yra žymiai efektyvesnis tikslumo prasme, tai patvirtina Andersono–Darlingo suderinamumo testų rezultatai
   c. vertinant stabiliųjų sekų parametrus didžiausio tikėtinumo metodu pastebėta, kad tikslo funkcijos (MTM) reikšmę, o tuo pačiu ir minimumo tašką stipriai įtakoja

parametrų α ir σ pokyčiai, kita vertus funkcija MTM yra mažiau jautri parametrų β ir μ pokyčiams.

1. Nustačius Baltijos šalių ir tarptautinių biržų finansinių sekų (atitinkamai 64 sekos ir 27 sekos) empirinius parametrus, paaiškėjo, kad dauguma sekų yra stipriai asimetriškos ($0.1<|\gamma_1|<30$). Empirinis ekscesas ($\gamma_2 \neq 0$) rodo, kad sekos empirinė tankio funkcija yra smailiaviršūniškenė už Gauss'o tankio funkciją. Vadinasi, galima atmesti hipotezę, kad akcijų kainų grąžų sekos pasiskirstę pagal normalųjį Gauss'o dėsnį. Šią prielaidą pagrindžia Andersono–Darlingo bei Kolmogorovo–Smirnovo suderinamumo testų rezultatai.

2. Tarptautinių rinkų finansinių sekų α ir β parametrų įverčių išsibarstymas rodo, kad dažniausiai $1,5 \leq \alpha \leq 2$. Parametras β šiuo atveju nėra didelis, bet dažniau įgyja neigiamas reikšmes. Baltijos šalių finansinių sekų (nagrinėjant pilną seką) parametrų α ir β įverčių išsibarstymas rodo, kad $\alpha < 1,5$ ir gali įgyti reikšmes artimas 1. Tuo tarpu, Baltijos šalių finansinių sekų (kai iš sekų pašalintos nulinės grąžos) parametras α yra išsibarstęs apie 1,5, bet gali įgyti reikšmes artimas 1 ar 2. Parametras β gali įgyti mažas teigiamas reikšmes, tačiau kai kuriais atvejais gali išaugti iki kritinių reikšmių ($>0,5$).

3. Sekų homogeniškumo testų eksperimentinis patikimumo tyrimas parodė, kad Andersono kriterijus vidutiniškai homogeniškas sekas atpažįsta geriau nei Smirnovo kriterijus. Andersono kriterijus tiriant realių sekų homogeniškumą su jų dalinių sumų sekomis nustatyta, kad 21 išsivysčiusių rinkų seka tenkina pagrindinę stabiliųjų dydžių savybę, o analizuojant Baltijos šalių sekas nustatyta, kad tik dviem atvejais (ETLAT ir LFO1L) ir tik pašalinus nulines grąžas, sekos yra homogeniškos su agreguotomis sekomis.

4. Tarptautinių kompanijų sekų savastingumo tyrimas pagal absoliutinių momentų metodą parodė, kad visos 27 sekos yra multifraktalinės, o 9 iš jų – savastingos. Analogiškai tiriant Baltijos šalių pilnas finansines seks nustatyta, kad 49 sekos yra multifraktalinės, 8 iš jų – savastingos. Pašalinus nulines grąžas gautos 27 multifraktalinės sekos iš kurių 9 – savastingos.

5. Sudarytas mišrusis stabilusis modelis leidžia tirti pasyvumo efektą besivystančiose rinkose. Finansines sekas siūloma modeliuoti mišriuoju stabiliuoju dėsniu. Įvestos šio modelio tikimybinio tankio, pasiskirstymo bei charakteringosios funkcijos. Pasiūlyti metodai leidžia įvertinti mišriojo modelio parametrus esant priklausomoms ir nepriklausomoms grąžų būsenoms. Laukimų trukmę tarp dviejų akcijos kainos pasikeitimų (nulinių grąžų serijų ilgių pasiskirstymą) siūloma modeliuoti ne binominiu, bet Hurwitz dzeta dėsniu.

6. Patikrinus Baltijos šalių rinkų sekų parametrų suderinamumą, pagal Koutrouvelio bei modifikuotą χ^2 (Romanovskio) metodus, nustatyta, kad 99% tenkina mišrųjį-stabilųjį modelį (pagal Koutrouvelis testą). Šių metodų patikimumo tyrimas parodė, kad didėjant tiek stabilumo parametrui α, tiek nulių skaičiui, abiejų testų patikimumas didėja. Atlikus analogišką eksperimentą su tolygiai intervale [0,1] pasiskirsčiusiomis sekomis, nustatyta, kad abu metodai vienodai atmeta suderinamumo hipotezes, o rezultatai yra trivialūs.

7. Nulinių būsenų serijų ilgių pasiskirstymo tyrimas parodė, kad empirinius duomenis geriausiai aprašo Hurwitz'o dzeta dėsnis (tiko 81-95%).Tai leidžia kelti hipotezę, kad nulinių bei vienetinių būsenų pasiskirstymai nėra atsitiktiniai. Atliktas Wald-Wolfowitz ženklų kriterijaus testas parodė, kad beveik visos Baltijos šalių sekos, išskyrus ETLAT, nėra grynai atsitiktinės su pasikliovimu lygmenimis (0.008,..., 0.1).

8. Priklausomybei būsenų sekose tirti pritaikius Hoelio kriterijų, pastebėta, kad nulinės eilės Markovo grandinių arba Bernulio schemą atitinkančių sekų beveik nėra. Esant $\varphi=10\%$ pasikliovimo lygmeniui 4-tos eilės Markovo grandinių yra apie 50%.

9. Aprašytieji ryšio tarp atskirų akcijų grąžų nustatymo metodai leidžia ryšį tarp sekų nustatyti netgi tuo atveju, kai neegzistuoja vieno ar kito atsitiktinio dydžio dispersija. Tam siūloma naudoti kovariantiškumo (simetriniams a.d.) ir kodiferencijos matus, kurių reikšmingumą galima nustatyti savirankos metodu.

10. Vertybinių popierių portfelio sudarymui, kai duomenys yra iš stabiliosios imties, siūloma naudoti modifikuotą Markowitz'o modelį. Šiame modelyje vietoje kovariacijų matricos siūloma naudoti kodiferencijų arba kovariantiškumo matricas.

11. Sukurta programinės įrangos bandomoji versija (*www.kabasinskas.tk*). Ši programinė priemonė leidžia generuoti atsitiktinių dydžių sekas, įvertinti pasirinktų sekų stabiliuosius parametrus bei sudaryti vertybinių popierių portfelį (iki 10 akcijų). Testavimai parodė, kad sekų parametrų vertinimo trukmė netiesiškai priklauso nuo sekos ilgio, tai patvirtino ir bandymai su superkompiuteriu VILKAS.

Disertacijos pabaigoje pateikiamas cituotos literatūros sąrašas. Darbo pabaigoje pateikiami tokie priedai:

Priedas A. Pateikiamos tyrimų rezultatų lentelės.

Priedas B. Pateikiami statistiniai metodai naudoti arba minimi disertacijoje: neparametrinių suderinamumo hipotezių tikrinimas Andersono–Darlingo ir Kolmogorovo–Smirnovo kriterijais, homogeniškumo tikrinimas (Andersono ir Smirnovo kriterijais), Wald–Wolfowitz serijų testas, mišriojo modelio adekvatumas, teorinis suderinamumo testų patikimumo nustatymas, koreliacijos koeficiento lygybė nuliui, Markovo tipo priklausomybės duomenų sekose nustatymas.

Priedas C. Aprašomas portfelio tipo parinkimas ir aptariami vidurkio–rizikos portfelio modeliai bei Markowitz modelis ir uždavinys.

4   Apie autorių

Audrius Kabašinskas gimė 1977 m. gruodžio mėn. 29 d., Kazlų Rūdoje, Marijampolės apskrityje.

Taikomosios matematikos bakalauro laipsnį įgijo 2000 m. Kauno Technologijos Universitete, Fundamentaliųjų mokslų fakultete. Ten pat, 2002 m. įgijo Taikomosios matematikos magistro laipsnį.

2002–2003 m. A. Kabašinskas – jaunesnysis mokslo darbuotojas Matematikos ir Informatikos institute. 2003–2007 m. studijavo doktorantūroje (Informatika).

2002–2007 m. A. Kabašinskas dirbo lektoriumi Vilniaus Gedimino Technikos Universitete, Informacinių Technologijų katedroje.

Nuo 2002 m. dirbo asistentu (2007 metais tapo lektoriumi) Kauno technologijos universitete, Matematinės Sistemotyros katedroje.

Email: audrius.kabasinskas@ktu.lt