Comparative analysis of conclusions uncertainty on connections between stocks of stock markets.

Koldanov Petr

National Research University Higher School of Economics,
Laboratory of Algorithms and Technologiies for Network Analysis (LATNA)
Nijni Novgorod, Russia
Joint work with Koldanov Alexander and Semenov Dmitry
pkoldanov@hse.ru

October 10, 2024

Outline

Introduction

- It is well known that any conclusions by observations contain uncertainty.
- Could all conclusions be separated by significant and admissible ones?
- Example - data of stock returns.
- Question: is uncertainty of different markets differ or no?
- New methodology for comparing uncertainty across different markets is developed
- Practical application of the methodology on stock markets of Russia, United States, and France is demonstrated.
- Our research is conducted within the framework of the concept of random variable network.

Introduction

Random variable network is a general model related with:

1. biological and medical studies
   Batushansky, A., Toubiana, D., Fait, A. Correlation-based network generation, visualization, and analysis as a powerful tool in biological studies: A case study in cancer cell metabolism, BioMed Res. Int. (2016), 19.,
2. gene expession or gene co-expression analysis
   Drton, M., Mathius, M., 2017. Structure learning in graphical modeling. Ann. Rev. Stat. Appl. 4, 365-393.,
3. market network analysis
   Kalyagin, V. A., Koldanov, A. P., Koldanov, P., Pardalos, P. M. Statistical Analysis of Graph Structures in Random Variable Networks. Springer, 2020.,
4. climate network analysis
   Tsonis, A.A., Roebber, P.J. The architecture of the climate network. Physica A, 333 (2004) 497504.
5. and others.

Introduction

- Simple and popular graph structure in random variable network is a threshold similarity graph which is called market graph in market network analysis
  Boginski, V., Butenko, S., Pardalos, P.M. Statistical analysis of financial networks. Computational Statistics & Data Analysis. 48 (2) (2005) 431-443.
- Problem of graph structure identification was considered in
  Kalyagin, V. A., Koldanov, A. P., Koldanov, P., Pardalos, P. M. Statistical Analysis of Graph Structures in Random Variable Networks. Springer, 2020..
- Uncertainty was measured by risk function for additive loss function.
- To calculate the risk function one needs assumption on true network structure.

Introduction

- In *P. A. Koldanov, A. P. Koldanov, D. P. Semenov. Confidence bounds for threshold similarity graph in random variable network // Statistical Analysis and Data Mining. 2023. Vol. 16. No. 6. P. 583-595.* new concept of uncertainty was proposed.
- The concept has an advantage over classical definitions of uncertainty as a risk function.
- Namely the uncertainty could be calculated from observed data as the difference between the upper and lower bounds which does not require knowledge of the true network structure.

Introduction

New concept of uncertainty is applied and further developed to make comparative analysis of conclusions uncertainty on connections between stocks of stock markets.

- New concept is based on the construction of upper and low bounds for threshold similarity graph.
- These bounds allows to separate all conclusions on threshold similarity graph by significant (reliable) and unreliable.
- Uncertainty is defined by number of unreliable conclusions.

Random variables network

- Let $X = (X_1, \ldots, X_N)$ is a random vector.
- Let $\gamma_{i,j} = \gamma(X_i, X_j)$ is a measure of similarity between $X_i, X_j, i, j = 1, \ldots, N$.
- As measure of similarity Pearson correlation and Kendall correlation will be considered.
- A pair (X, γ) is called random variables network.[1]
- The matrix

$$\Gamma = ((\gamma_{i,j}))_{N \times N}, \quad i, j = 1, \ldots, N \tag{1}$$

describe all pairwise similarities between components of the vector X.

[1] Kalyagin V. A., Koldanov A. P., Koldanov P., Pardalos P. M. Statistical Analysis of Graph Structures in Random Variable Networks. Springer, 2020.

Network model, threshold similarity graph

- The random variables network (X, γ) generates a network model (V, Γ) — the complete weighted graph with N nodes, where the weights of edges (i, j) is given by $\gamma_{i,j} = \gamma(X_i, X_j)$. (V, Γ) — true network model.
- Threshold similarity graph is a pair (V, E) where $V = \{1, \ldots, N\}$ is the set of nodes which correspond to the random variables $X_i, i = 1, \ldots, N$ and E is the set of unweighted edges between nodes in V.
- Edge $(i, j) \in E$ if $\gamma_{i,j} = \gamma(X_i, X_j) > \gamma_0$, where γ_0 is a given threshold.
- Let $J_e(\gamma_0) = \{(i, j) : \gamma_{i,j} > \gamma_0\}$ - set of edges of the true threshold graph at threshold γ_0.

Approach description

Let $X = (X(1), \ldots, X(n))$ be sample from vector X where $X(i) = (X_1(i), \ldots, X_N(i)), \quad i = 1, \ldots, n.$
In [2] to identify market graph by observations it was proposed to construct two set of edges $L_e(x, \gamma_0, P^\star)$ and $U_e(x, \gamma_0, P^\star)$ satisfying the condition:

$$P(L_e(x, \gamma_0, P^\star) \subset J_e(\gamma_0) \subset U_e(x, \gamma_0, P^\star)) \geq P^\star. \qquad (2)$$

The condition (2) means that with a given probability $P^\star$ the set $U_e(x, \gamma_0, P^\star)$ contains a set of edges of the true market graph and at the same time the set $L_e(x, \gamma_0, P^\star)$ lies in the set of edges of the true market graph.
The sets $U_e(x, \gamma_0, P^\star), L_e(x, \gamma_0, P^\star)$ was called simultaneous upper and lower bounds of the level $P^\star$ for the set edges $J_e(\gamma_0)$ of the true market graph.

[2] P. A. Koldanov, A. P. Koldanov, D. P. Semenov. Confidence bounds for threshold similarity graph in random variable network // Statistical Analysis and Data Mining. 2023. Vol. 16. No. 6. P. 583-595.

Methodology for $L_e(x, \gamma_0, P^\star)$ and $U_e(x, \gamma_0, P^\star)$ construction

Methodology for $U_e(x, \gamma_0, P^\star), L_e(x, \gamma_0, P^\star)$ construction is based on the tests

$$\varphi^e_{i,j}(x) = \begin{cases} 1, & T_{i,j}(x) < c^e_{i,j} \\ 0, & T_{i,j}(x) \geq c^e_{i,j} \end{cases} \tag{3}$$

$$\varphi^n_{i,j}(x) = \begin{cases} 1, & T_{i,j}(x) > c^n_{i,j} \\ 0, & T_{i,j}(x) \leq c^n_{i,j} \end{cases} \tag{4}$$

for testing hypotheses $h^e_{i,j} : \gamma_{i,j} > \gamma_0$ and $h^n_{i,j} : \gamma_{i,j} \leq \gamma_0$, $i, j = 1, \ldots, N$ respectively. Critical values $c^n_{i,j}, c^e_{i,j}$ are defined from

$$P_{\gamma_0}(T_{i,j}(x) > c^n_{i,j}) = P_{\gamma_0}(T_{i,j}(x) < c^e_{i,j}) = \frac{\alpha}{M} \tag{5}$$

where $\alpha = 1 - P^\star$

Methodology for $L_e(x, \gamma_0, P^\star)$ and $U_e(x, \gamma_0, P^\star)$ construction

Methodology for $L_e(x, \gamma_0, P^\star)$ and $U_e(x, \gamma_0, P^\star)$ construction has the form:

- test hypotheses $h^e_{i,j}$ и $h^n_{i,j}$, $i, j = 1, \ldots, N$ on the same level $\frac{\alpha}{M}$.
- Pairs (i, j), such that hypotheses $h^e_{i,j}$ are accepted are included to the $U_e(x, \gamma_0, P^\star)$
- Pairs (i, j), such that hypotheses $h^n_{i,j}$ are rejected are included to the $L_e(x, \gamma_0, P^\star)$.

Results of the approach

This method, with a fixed γ_0, allows us to divide all conclusions about connections between stocks of the stock market into three types:

- significant conclusions about the presence of a connection — set $L_e(x, \gamma_0, P^\star)$,
- significant conclusions about absence of connection — set $L_n(x, \gamma_0, P^\star) = J \setminus U_e(x, \gamma_0, P^\star)$,
- not significant but admissible conclusions — set

$$G(x, \gamma_0, P^\star) = U_e(x, \gamma_0, P^\star) \setminus L_e(x, \gamma_0, P^\star) = U_n(x, \gamma_0, P^\star) \setminus L_n(x, \gamma_0, P^\star) \quad (6)$$

The set (6) consists of those pairs of stocks between which, for a given volume of observations and a fixed threshold γ_0 at a given significance level $1 - P^\star$, it is possible to either draw or not draw an edge.

To assess the uncertainty of conclusions about connections in the analyzed market, it was proposed to use $E(|G(x, \gamma_0, P^\star)|)$.

Classical approach to measure of uncertainty

- Classical approach to measure uncertainty is based on the risk function for given loss function [3].
- Uncertainty of statistical procedures for network structure identification in the framework of classical approach under additive loss function [4] was discussed in [5].
- It was shown that under additive loss function such uncertainty is related with difference between edges in the true and sample network structures.

---

[3] A. Wald, Statistical decision functions. (1950).

[4] E.L. Lehmann, *A theory of some multiple decision problems, I.* The Annals of Mathematical Statistics. 1957.Vol.28. Pp. 1-25.

[5] Kalyagin V. A., Koldanov A. P., Koldanov P., Pardalos P. M. Statistical Analysis of Graph Structures in Random Variable Networks. Springer, 2020.

Procedure δ^n for market graph identification

In [6] it was proposed procedure $\delta^n = (\varphi^n_{1,2}, \ldots, \varphi^n_{N-1,N})$ based on the tests

$\varphi^n_{i,j} = \begin{cases} 1, & T_{i,j} > c^n_{i,j} \\ 0, & T_{i,j} \leq c^n_{i,j} \end{cases}$ for testing hypotheses

$$h^n_{i,j} : \gamma_{i,j} \leq \gamma_0 \text{ versus } k^n_{i,j} : \gamma_{i,j} > \gamma_0.$$

In [7] it was proposed to measure uncertainty of the procedure δ^n by the risk function under additive loss function with components $a^n_{i,j}$; $b^n_{i,j}$ where
$a^n_{i,j}$ — loss from decision $\varphi^n_{i,j} = 1$ for the case $\gamma_{i,j} \leq \gamma_0$,
$b^n_{i,j}$ — loss from decision $\varphi^n_{i,j} = 0$ for the case $\gamma_{i,j} > \gamma_0$.
Moreover it was shown that risk function of procedure δ^n has the form:

$$Risk(\theta, \delta^n) = \\ = \sum_{(i,j):\gamma_{i,j} \leq \gamma_0} a^n_{i,j} P(T_{i,j} > c^n_{i,j}) + \sum_{(i,j):\gamma_{i,j} > \gamma_0} b^n_{i,j} P(T_{i,j} \leq c^n_{i,j}) \quad (7)$$

[6] Koldanov A. P., Kalyagin V. A., Koldanov P.A., Pardalos P. M. Statistical procedures for the market graph construction // Computational Statistics & Data Analysis. 2013. Vol. 68. P. 17-29.

[7] Kalyagin V.A., Koldanov A.P., Koldanov P.A., Pardalos P.M. Zamaraev V.A Measures of uncertainty in market network analysis, Physica A 413 (2014) 59–70

Procedure δ^e for market graph identification

In the same way one can construct procedure $\delta^e = (\varphi^e_{1,2}, \ldots, \varphi^e_{N-1,N})$ for market graph identification based on the tests $\varphi^e_{i,j} = \begin{cases} 1, & T_{i,j} \leq c^e_{i,j} \\ 0, & T_{i,j} > c^e_{i,j} \end{cases}$ for testing hypotheses

$$h^e_{i,j} : \gamma_{i,j} > \gamma_0 \text{ versus } k^e_{i,j} : \gamma_{i,j} \leq \gamma_0.$$

It is easy to prove that risk function of the procedure δ^e under additive loss function with components $a^e_{i,j}; b^e_{i,j}$ where
$a^e_{i,j}$ — loss from decision $\varphi^e_{i,j} = 1$ for the case $\gamma_{i,j} > \gamma_0$,
$b^e_{i,j}$ — loss from decision $\varphi^e_{i,j} = 0$ for the case $\gamma_{i,j} \leq \gamma_0$.
has the form:

$$\begin{aligned} &Risk(\theta, \delta^e) = \\ &= \sum_{(i,j):\gamma_{i,j} > \gamma_0} a^e_{i,j} P(T_{i,j} < c^e_{i,j}) + \sum_{(i,j):\gamma_{i,j} \leq \gamma_0} b^e_{i,j} P(T_{i,j} > c^e_{i,j}) \end{aligned} \tag{8}$$

Procedure $\delta = (\delta^e, \delta^n)$ for market graph identification

If both procedures δ^n, δ^e are applied simultaneously to identify market graph by observations the risk function of procedure $\delta = (\delta^n, \delta^e)$ could be defined as:

$$Risk(\theta, \delta) = Risk(\theta, \delta^e) + Risk(\theta, \delta^n) \tag{9}$$

Since

$$P(T_{i,j} > c^e_{i,j}) = P(c^e_{i,j} < T_{i,j} \leq c^n_{i,j}) + P(T_{i,j} > c^n_{i,j})$$

$$P(T_{i,j} \leq c^n_{i,j}) = P(T_{i,j} < c^e_{i,j}) + P(c^e_{i,j} < T_{i,j} \leq c^n_{i,j})$$

then (9) could be written as

$$\begin{aligned} Risk(\theta, \delta) = & \sum_{(i,j):\gamma_{i,j} > \gamma_0} (a^e_{i,j} + b^n_{i,j}) P(T_{i,j} < c^e_{i,j}) + \\ & + \sum_{(i,j):\gamma_{i,j} \leq \gamma_0} (a^n_{i,j} + b^e_{i,j}) P(T_{i,j} > c^n_{i,j}) + \\ & + \sum_{(i,j):\gamma_{i,j} \leq \gamma_0} b^e_{i,j} P(c^e_{i,j} < T_{i,j} \leq c^n_{i,j}) + \\ & + \sum_{(i,j):\gamma_{i,j} > \gamma_0} b^n_{i,j} P(c^e_{i,j} < T_{i,j} \leq c^n_{i,j}) \end{aligned} \tag{10}$$

Relations between measures of uncertainty

Let

$$a^e_{i,j} = a^n_{i,j} = a; b^e_{i,j} = b^n_{i,j} = b; a + b = 1, \quad \forall i,j = 1,\ldots,N; i \neq j.$$

Then

$$Risk(\theta,\delta) = \sum_{(i,j):\gamma_{i,j}>\gamma_0} P(T_{i,j} < c^e_{i,j}) + \sum_{(i,j):\gamma_{i,j}\leq\gamma_0} P(T_{i,j} > c^n_{i,j}) + \\ + b\sum_{(i,j)} P(c^e_{i,j} < T_{i,j} \leq c^n_{i,j}) \tag{11}$$

Relations between measures of uncertainty

Since

$$|G(x, \gamma_0, P^\star)| = \sum_{i,j=1}^{N} I(\varphi_{i,j}^e = 0, \varphi_{i,j}^n = 0)$$

then

$$E(|G(x, \gamma_0, P^\star)|) = \sum_{(i,j)} P(c_{i,j}^e < T_{i,j} \leq c_{i,j}^n)$$

and from (11) one has

$$Risk(\theta, \delta) = \sum_{(i,j):\gamma_{i,j} > \gamma_0} P(T_{i,j} < c_{i,j}^e) + \sum_{(i,j):\gamma_{i,j} \leq \gamma_0} P(T_{i,j} > c_{i,j}^n) + \\ + bE(|G(X, \gamma_0, P^\star)|) \quad (12)$$

Therefore measure of uncertainty $E\left(|G(x, \gamma_0, P^\star)|\right)$ is a part of classical measure of uncertainty for additive loss function. Note that first and second terms from (12) are bounded and $Risk(\theta, \delta)$ is defined by expectation of the gap (6).

First indicator of uncertainty

It is obvious that measure of uncertainty $E\left(|G(x,\gamma_0,P^\star)|\right)$ depends on the size of the market being analyzed.
To assess the uncertainty of conclusions about connections in the market at the threshold γ_0, independent of the market size, the following coefficient can be proposed:

$$K_1(x,\gamma_0,P^\star) = \frac{|G(x,\gamma_0,P^\star)|}{M}^{8} \qquad (13)$$

The coefficient shows ratio of admissible conclusions about connections among all conclusions.

[8] $M = C_N^2$ - number of all possible pairs

Second indicator of uncertainty

Another indicator of uncertainty in conclusions about market connections represents the ratio of the number of admissible conclusions about market connections to the number of significant conclusions about market connections at a fixed threshold γ_0, i.e.

$$K_2(x, \gamma_0, P^\star) = \frac{|G(x, \gamma_0, P^\star)|}{|L_e(x, \gamma_0, P^\star) \cup L_n(x, \gamma_0, P^\star)|} \tag{14}$$

Comparison and weak points of the indicators

- The indicator $K_1^i(x, \gamma_0, P^\star)$ ranges from 0 to 1, whereas the indicator $K_2^i(x, \gamma_0, P^\star)$ ranges from 0 to $+\infty$.
- The values of $K_2^i(x, \gamma_0, P^\star)$ depend on the size of the analyzed market. In particular, if at a certain threshold γ_0 the number of significant conclusions is equal to 1, then the number of admissible conclusions, and consequently the value of $K_2^i(x, \gamma_0, P^\star)$, is equal to $M - 1$.

1. Comparison of the uncertainties in different markets could be done using the indicators $K_1^i(x, \gamma_0, P^\star)$ and $K_2^i(x, \gamma_0, P^\star)$, which indicate the uncertainties of conclusions about connections in the i-th stock market at the threshold γ_0.
2. Such a comparison is of interest, but the results may depend on the strength of the connections, i.e., on the value of γ_0.

Aggregate indicators of the uncertainty of conclusions about connections in the market i

$$K_1^i(x, P^\star) = \sum_{s=1}^{K} c_s \left[K_1^i(x, \gamma_s, P^\star)\right] \tag{15}$$

$$K_2^i(x, P^\star) = \sum_{s=1}^{K} c_s \left[K_2^i(x, \gamma_s, P^\star)\right] \tag{16}$$

$$c_s \geq 0, \sum_{s=1}^{K} c_s = 1.$$

where $\{\gamma_1, \gamma_2, \ldots, \gamma_K\}$ are the values of the threshold γ_0, c_s are the weigths of the threshold γ_s.

Comparative aggregated indicators of uncertainty in conclusions about connections in markets i and j

$$K_1^{ij}(x, P^\star) = K_1^i(x, P^\star) - K_1^j(x, P^\star) = \sum_{s=1}^{K} c_s \left[K_1^i(x, \gamma_s, P^\star) - K_1^j(x, \gamma_s, P^\star) \right] \tag{17}$$

$$K_2^{ij}(x, P^\star) = K_2^i(x, P^\star) - K_2^j(x, P^\star) = \sum_{s=1}^{K} c_s \left[K_2^i(x, \gamma_s, P^\star) - K_2^j(x, \gamma_s, P^\star) \right] \tag{18}$$

- Smaller absolute value of the indicators (17) signifies that the ratios of admissible conclusions about connections in the corresponding stock markets are close.
- Smaller absolute value of the indicators (18) indicates that the ratios of admissible and significant conclusions about connections in the corresponding stock markets are close.

Comparative aggregated indicators of uncertainty in conclusions about connections in markets i and j

$$K_3^{ij}(x, P^\star) = \frac{K_1^i(x, P^\star)}{K_1^j(x, P^\star)} \tag{19}$$

$$K_4^{ij}(x, P^\star) = \frac{K_2^i(x, P^\star)}{K_2^j(x, P^\star)} \tag{20}$$

When using the indicators (19) and (20), the proximity of these indicators to 1 serves as an indicator of the proximity of uncertainty in conclusions about connections in the considered stock markets.

Asymptotic analysis of the introduced uncertainty indicators

This section discusses and analyzes the comparative aggregate indicators of uncertainty introduced by the relations (17), (18), (19), (20) under conditions of consistency of tests $\varphi^e_{l,s}(x), \varphi^n_{l,s}(x)$under $n \to \infty$.
By formula (15) from [9]

$$E_{\gamma_i}|G(x, \gamma_i, P^\star)| = |K_i|(1 - \frac{2\alpha}{M}) + \sum_{l,s} \beta_{l,s}$$

where $K_i = \{(l, s) : \gamma_{l,s} = \gamma_i\}$,

$$\beta_{l,s} = P_{\gamma_{l,s} \neq \gamma_i}(\varphi^e_{l,s}(x) = 0, \varphi^n_{l,s}(x) = 0).$$

[9] P. A. Koldanov, A. P. Koldanov, D. P. Semenov. Confidence bounds for threshold similarity graph in random variable network // Statistical Analysis and Data Mining. 2023. Vol. 16. No. 6. P. 583-595.

Asymptotic analysis of the introduced uncertainty indicators

From conditions of consistency of tests $\varphi^e_{l,s}(x), \varphi^n_{l,s}(x)$ it follows $\sum_{l,s} \beta_{l,s} \to 0$ for $n \to \infty$. Then one has:

$$E\left(K_1(x, \gamma_i, P^\star)\right) = \frac{|K_i|}{M}(1 - \frac{2\alpha}{M}) + \frac{\sum_{l,s} \beta_{l,s}}{M} \to \frac{|K_i|}{M}.$$

Therefore from (15) one has

$$E\left(K_1^i(x, P^\star)\right) = \sum_{t=1}^{K} c_t E\left(K_1^i(x, \gamma_t, P^\star)\right) \to \sum_{t=1}^{K} c_t \frac{|K_t^i|}{M_i}$$

where M_i is the size of market i, K_t^i—number of pairs (l, s) from market i such that $\gamma_{l,s} = \gamma_t$.
Note that $\frac{|K_t^i|}{M_i}$ is the value of histogram $H^i(\gamma_t)$ of correlation coefficient of the market i at the threshold γ_t.

Asymptotic analysis of the introduced uncertainty indicators

Therefore, for the expectation of comparative aggregate indicator of uncertainty introduced by the relation (17) one has:

$$E\left(K_1^{ij}(x,P^\star)\right) \to \sum_{t=1}^{K} c_t \left(\frac{|K_t^i|}{M_i} - \frac{|K_t^j|}{M_j}\right) \tag{21}$$

In addition, for the expectation of comparative aggregated indicator of uncertainty introduced by the relation (19) one has:

$$E\left(K_3^{ij}(x,P^\star)\right) \to \frac{\sum_{t=1}^{K} c_t \left(\frac{|K_t^i|}{M_i}\right)}{\sum_{t=1}^{K} c_t \left(\frac{|K_t^j|}{M_j}\right)} \tag{22}$$

Asymptotic analysis of the introduced uncertainty indicators

By Chebishev inequality

$$|G(x, \gamma_0, P^\star)| \xrightarrow{P} C_N^2 \times H(\gamma_0)$$

and

$$K_1(x, \gamma_0, P^\star) \xrightarrow{P} H(\gamma_0).$$

Since

$$K_2(x, \gamma_0, P^\star) = \frac{|G(x, \gamma_0, P^\star)|}{|L_e(x, \gamma_0, P^\star) \cup L_n(x, \gamma_0, P^\star)|} = \frac{K_1(x, \gamma_0, P^\star)}{1 - K_1(x, \gamma_0, P^\star)}$$

then by the theorem 2.1.3 of [10] for the case $H(\gamma_0) \neq 1$ one has

$$K_2(x, \gamma_0, P^\star) \xrightarrow{P} \frac{H(\gamma_0)}{1 - H(\gamma_0)}.$$

[10] E.L. Lehmann Elements of Large-Sample Theory.

Asymptotic analysis of the introduced uncertainty indicators

It follows that

$$K_2^i(x, P^\star) = \sum_{s=1}^{K} c_s \left[K_2^i(x, \gamma_s, P^\star)\right] \xrightarrow{P} \sum_{s=1}^{K} c_s \left[\frac{H^i(\gamma_s)}{1 - H^i(\gamma_s)}\right]$$

$$K_2^{ij}(x, P^\star) = K_2^i(x, P^\star) - K_2^j(x, P^\star) \xrightarrow{P} \sum_{s=1}^{K} c_s \left[\frac{H^i(\gamma_s)}{1 - H^i(\gamma_s)} - \frac{H^j(\gamma_s)}{1 - H^j(\gamma_s)}\right]$$

$$K_4^{ij}(x, P^\star) = \frac{K_2^i(x, P^\star)}{K_2^j(x, P^\star)} \xrightarrow{P} \frac{\sum_{s=1}^{K} c_s \left[\frac{H^i(\gamma_s)}{1-H^i(\gamma_s)}\right]}{\sum_{s=1}^{K} c_s \left[\frac{H^j(\gamma_s)}{1-H^j(\gamma_s)}\right]}$$

Proposed method is used to analyze and compare the uncertainty of conclusions about the relationships between the returns of selected stocks in three markets:

- the Russian market (25 most liquid shares of the Moscow Stock Exchange),
- the French market (39 out of 40 shares included in the CAC-40 index of the Paris Stock Exchange),
- the US market (all 30 stocks included in the Dow-Jones index of the NASDAQ stock exchange).

Period of observations - 254 trading days from 01/01/2021 to 01/01/2022. The strength of the connection was measured at 22 values of the threshold γ_0 from the segment $[-0.1; 0.95]$ with a step of 0.05.

Comparison by size of $G(x, \gamma_0, P^\star)$

Dependence of $|G(x, \gamma_0, P^\star)|$ in stock markets of Russia, France and USA from γ_0 for two correlations. One make the following conclusions:

- the most definitive is the Russian market;
- the most uncertain market is France;

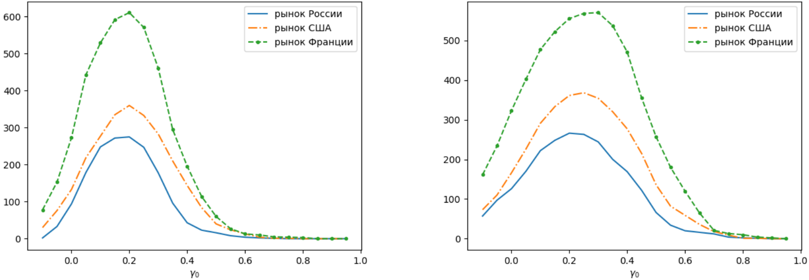

Figure: Dependence of $|G(x, \gamma_0, P^\star)|$ from γ_0 for Kendall (left) and Pearson (right) correlations, $P^\star = 0.9$. Stock markets of Russia, France and USA.

Comparison by the shares of valid conclusions

Analysis of Fig. 2 shows that by the share of valid conclusions (coefficient $K_1(x, \gamma_0, P^\star)$) the markets in question differ only slightly at almost all thresholds, regardless of the correlation coefficient used.

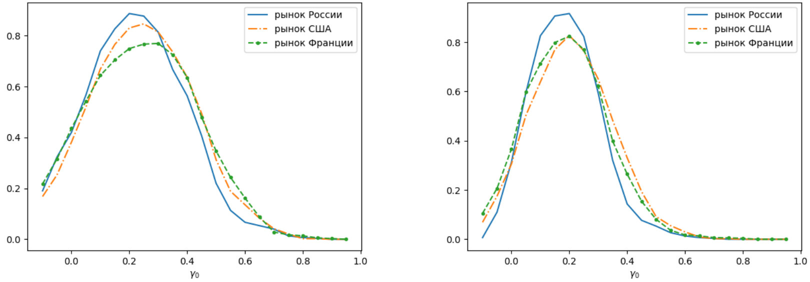

Figure: Dependence of $K_1(x, \gamma_0, P^\star)$ from γ_0, Pearson correlation (left) and Kendall correlation (right), $P^\star = 0.9$. Stock markets of Russia, France and USA.

Comparison by ratio of the number of admissible conclusions to the number of significant conclusions

In terms of the ratio of the number of admissible conclusions to the number of significant conclusions (Fig. 3), markets differ significantly at thresholds in the interval $(0.1; 0.3)$.

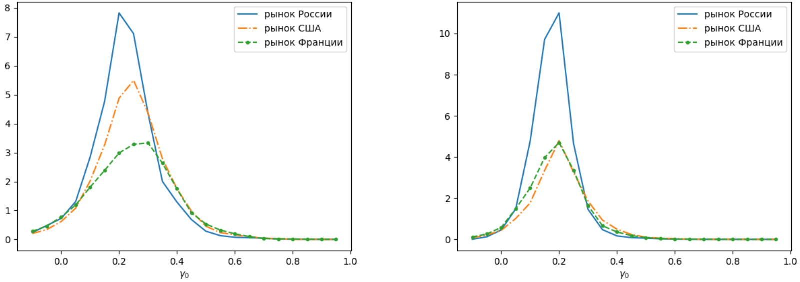

Figure: Dependence of $K_2(x, \gamma_0, P^\star)$ from γ_0, Pearson correlation (left) and Kendall correlation (right), $P^\star = 0.9$. Stock markets of Russia, France and USA.

Comparison by aggregated indicators with equal weights

In terms of the share of admissible conclusions, the considered indicators differ slightly also in terms of the aggregated indicator.

	Pearson correlation			Kendall correlation		
	Russia	France	USA	Russia	France	USA
$K_1^i(x, P^\star)$	0.354	0.359	0.359	0.261	0.272	0.268
$K_2^i(x, P^\star)$	1.556	1.045	1.308	1.568	0.909	0.848

Table: Values of $K_1^i(x, P^\star)$ and $K_2^i(x, P^\star)$ for $P^\star = 0.9, c_s = \frac{1}{K}, \forall s = 1, \ldots, K; K = 22,$ $(\gamma_1, \gamma_2, \ldots, \gamma_K) = (-0.1, -0.05, \ldots, 0.95)$. Pearson and Kendall correlations.

Comparison by comparative aggregated indicators with equal weights. Pearson correlation

	Russia	France	USA
Russia	-	0.9881	0.9871
France	1.0121	-	0.999
USA	1.0131	1.001	-

Table: Value of $K_3^{ij}(x, P^\star)$ for $P^\star = 0.9, c_s = \frac{1}{K}, \forall s = 1, \ldots, K$; $K = 22, (\gamma_1, \gamma_2, \ldots, \gamma_K) = (-0.1, -0.05, \ldots, 0.95)$. Pearson correlation

	Russia	France	USA
Russia	-	1.4884	1.1892
France	0.6718	-	0.799
USA	0.8409	1.2516	-

Table: Value of $K_4^{ij}(x, P^\star)$ for $P^\star = 0.9, c_s = \frac{1}{K}, \forall s = 1, \ldots, K; K = 22$, $(\gamma_1, \gamma_2, \ldots, \gamma_K) = (-0.1, -0.05, \ldots, 0.95)$. Pearson correlation.

Comparison by comparative aggregated indicators with equal weights. Kendall correlation

	Russia	France	USA
Russia	-	0.9581	0.9717
France	1.0438	-	1.0143
USA	1.0291	0.9859	-

Table: Value of $K_3^{ij}(x, P^\star)$ for $P^\star = 0.9, c_s = \frac{1}{K}, \forall s = 1, \dots, K$; $K = 22, (\gamma_1, \gamma_2, \dots, \gamma_K) = (-0.1, -0.05, \dots, 0.95)$. Kendall correlation.

	Russia	France	USA
Russia	-	1.7253	1.8488
France	0.5796	-	1.0716
USA	0.5409	0.9332	-

Table: Value of $K_4^{ij}(x, P^\star)$ for $P^\star = 0.9, c_s = \frac{1}{K}, \forall s = 1, \dots, K$; $K = 22, (\gamma_1, \gamma_2, \dots, \gamma_K) = (-0.1, -0.05, \dots, 0.95)$. Kendall correlation.

Conclusion

- The approach proposed in [11] is applied to make comparative analysis of conclusions uncertainty on connections between several stocks of stock markets USA, France and Russia.
- Several simple indicators are proposed. First type of these indicators depends from strengths of connections. Second type of these indicators has aggregated character.
- Asymptotic properties of the indicators are investigated. The properties allows the indicators to be simple calculated using histogram of correlation coefficient.
- Pearson and Kendall correlations are used.

[11] P. A. Koldanov, A. P. Koldanov, D. P. Semenov. Confidence bounds for threshold similarity graph in random variable network // Statistical Analysis and Data Mining. 2023. Vol. 16. No. 6. P. 583-595.

Conclusion. Discussion of obtained results

- The number of stocks does not taken into account - the most definite is the Russian market, the most uncertain market is France.
- By the share of valid conclusions - the markets in question differ only slightly at almost all thresholds, regardless of the correlation coeffcient used.
- By the ratio of the number of admissible conclusions to the number of significant conclusions - the most definite is the France market, the most uncertain market is Russia.
- In terms of the aggregated indicators - the considered markets differ slightly.

THANK YOU FOR YOUR ATTENTION!

Histogram of Pearson correlations

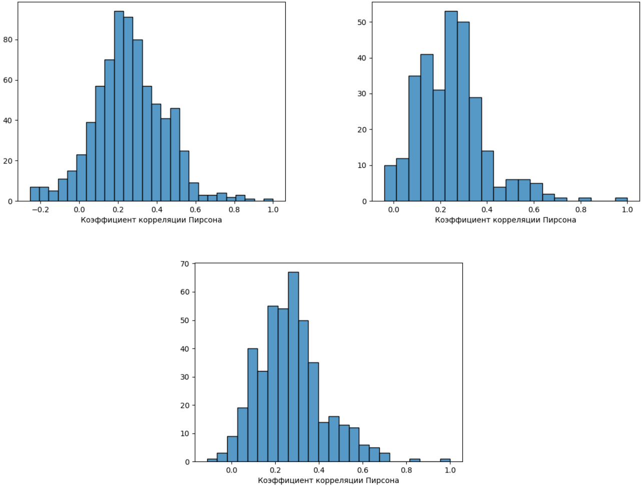

Figure: Histogram of Pearson correlations. Left - France, right - Russia, down - USA

Histogram of Kendall correlations

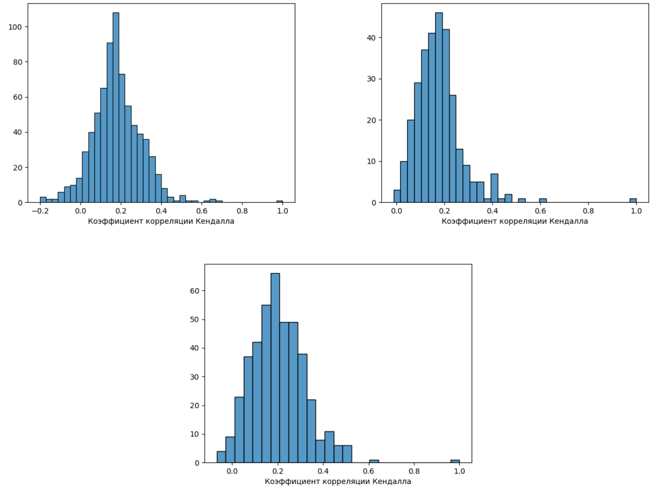

Figure: Histogram of Kendall correlations. Left - France, right - Russia, down - USA