# МУЛТИПЛИКАТИВЕН МОДЕЛ НА SOFIX[*]

**ВИКТОР Г. ГЕРЧЕВ, ПАВЛИНА К. ЙОРДАНОВА**

# MULTIPLICATIVE MODEL OF SOFIX

**VIKTOR G. GERCHEV, PAVLINA K. JORDANOVA**

***ABSTRACT:*** *In this paper we consider the main index on Bulgarian Stock Exchange, the index SOFIX. In order to obtain stationary sequence, we make log-return transformation of SOFIX daily closing price and prepare a multiplicative model of the data. It turns out that the error term has logarithmic stable distribution. Using the inverse transformations we analyse the initial data. We propose the Monte-Carlo method for preparing forecasts. Due to the large variation of stable distribution these forecasts have too wide confidence intervals.*

***KEYWORDS:*** *SOFIX, Stable distribution*

## 1. История на проблема и новости при решаването му

Основният борсов индекс на регулираните пазари на Българската Фондова Борса (БФБ) е SOFIX. Неговото изчисление започва на 20 октомври 2000 г. при базисна стойност от 100 пункта. Той е базиран на пазарната капитализация на включените 15 емисии обикновени акции, коригирана с фрий-флоута на всяка от тях. Под фрий-флоут се разбира делът на акциите, които са налични за покупка, за инвеститорите. За по-подробно описание виж Малинов и Канарян [3]. Емисиите в SOFIX е необходимо да отговарят на определени изисквания за ликвидност, пазарна капитализация не по-малко от 50 млн.лв., фрий-флоут и брой акционери (поне 500) и обхваща само емисии, които са допуснати до търговия на тези пазари. Дружествата, чиито емисии се търгуват са публични, не са обявени в ликвидация или несъстоятелност и не изпълняват оздравителна програма. Формула, по която се изчислява този индекс е

$$\mathrm{SOFIX}(t) = \mathrm{SOFIX}(t-1)\left[\frac{\sum_{i=1}^{n} N_i(t)P_i(t)FF_i(t)W_i(t)D_i(t)}{\sum_{i=1}^{n} N_i(t-1)P_i(t-1)FF_i(t-1)W_i(t-1)}\right]K,$$

където

$N_i(t)$ е броят на акциите в емисията на i-тото дружество в деня t,
$P_i(t)$ е цената на последната сделка за i-тата ценна книга в деня t,
$FF_i(t)$ е фрий-флоутът на i-тата ценна книга за деня t,
$W_i(t)$ е тегловият коефициент на i-тата ценна книга за деня t,
$D_i(t)$ е делителят за текущата търговска сесия на i-тата ценна книга за деня t,
n е броят на емисиите, включени в индекса,
i е номерът на конкретната ценна книга,
К е коригиращ фактор, който е различен от единица само в случаите на промяна на базата на индекса,

[*] Статията е финансирана от НИП „Математически модели за оценка на риска“, № РД-08-230/12.03.2014 г. на Шуменски университет.

t е денят, за който се изчислява индекса.

Със събирането на все повече данни за търговията на БФБ тенденциите в поведението на отделните индекси предизвиква все по-голям интерес след изследователите статистици. Един от първите много добри анализи е направен от Patev и Kanaryan [9]. Те правят преглед на изследванията на аналогични въпроси, възникнали при търговията с други международни индекси и анализират риска от инвестиции на БФБ като прилагат предивмно метода на експоненциално претеглените плъзгащи се средни като частен случай на GARCH(1,1) моделирането. То от своя страна е възможно само при наличие на дълъг ред от данни. Посочват, че рискът от инвестиции на фондовите борси се свързва предимно с волатилността на разпределението на възвращаемостта, и че основният недостатък на този метод е, че той приема хомоскедастичност в данните, която в повечето случаи не е на лице. Отбелязват, че изменението на волатилността във времето се взема в предвид при GARCH моделирането. Patev и Kanaryan [9] не разглеждат устойчивото разпределение като модел на логаритмичната възвращаемост, което както се вижда от настоящата статия, е много близо до действителното разпределение на наблюдаваната величина. Аналогично изследване е правено и през 2010 г. от Петков [11].

През 2006 г. Милинов и Канарян анализират връзката между рискът от съответната инвестиция на БФБ и обемът на търговията. Един изчерпателен анализ на динамиката на търговията на фондовите борси на Балканите до 2006 г. е направен в Samitas и др. [10]. С нарастването на обема на данните става възможно моделирането на логаритмичната възвръщаемост с модел с тежка опашка на разпределението на наблюдаваната величина. През последните години Ломев и съавтори [2, 7, 8] прилагат GARCH моделите с тежки опрашки за SOFIX. През 2009 г. те наблюдават отношението на сумата и максимума на логаритмичната възвръщаемост и посредством функцията на средните стойности на надхвърлянията показват, че логаритчимната възвръщаемост се моделира с разпределение, което няма вариация и е с тежка опашка, прилагат GARCH и FARIMA моделите, но не използват $\alpha$ - устойчиво разпределение, което има достатъчно параметри и поради това предоставя възможност за добър модел на данните. Обобщаване на фактите около възвращаемостта на капиталите при инвестиране в търговия на БФБ до 2009 г. може да бъде намерено в Ценков [5] . Един много добър метод за предвиждане на поведението на SOFIX е чрез невронни мрежи. Подобно изследване е правено от Shahpazov през 2013. Виж [11].

В настоящото изследване е обхванат периодът от 16 март 2005 до 19 септември 2014 г. Данните са взети от сайта quotenet.com.

**http://www.quotenet.com/index/historical-prices/SOFIX/1.1.2000_20.9.2014**

Показано е, че разпределението на логаритмичната възвръщаемост се моделира много добре с $\alpha$-устойчиво разпределение и чрез симулации и Монте-Карло метод е направена прогноза. За изчертаването на графикити и обработването на данните, са използвани функции от софтуера R. Това е софтуер с отворен код, предназначен предимно за статистическа обработка. Инсталационни файлове, описание и упътвания за работа с R могат да бъдат намерени в [12].

## 2. Описание на модела

Нека Y(1), Y(2), ..., Y(n) са данни от динамичен ред, с равно отдалечени интервали от време. В класическия, адитивен модел за анализ на динамични редове

(1) $$Y(t) = \mu(t) + c(t) + \varepsilon(t)$$

имаме следните компоненти

$\mu(t)$ – тренд в момента t;

$c(t)$ – циклични компоненти в момента $t$;
$\varepsilon(t)$ – стохастична грешка в момента $t$, като тази грешка има нулево математическо очакване и за редицата $\varepsilon$ (1), $\varepsilon$ (2), ..., $\varepsilon$ (n) се предполага, че е от независими наблюдения върху случайната грешка $\varepsilon$. Този модел не е подходящ за моделиране на времеви ред, ако вариацията на случайната му грешка нараства с времето защото се нарушава условието за еднаква разпределеност. Ако изходните данни могат да бъдат трансформирани така, че този модел да бъде адекватен за тях ще можем да използваме добре развитите техники за работа с него.

Един алтернативен модел е мултипликативният модел

(2) $$Y(t) = \mu(t).c(t).\varepsilon(t).$$

При него цикличната и случайната компоненти са пропорционални на средното ниво на процеса. Чрез логаритмуване на модел (2), той се свежда до адитивния модел

(3) $$\log Y(t) = \log \mu(t) + \log c(t) + \log \varepsilon(t).$$

Ще означаваме с $Y_L(t) = \log Y(t)$, $\mu_L(t) = \log \mu(t)$, $c_L(t) = \log c(t)$ и $\varepsilon_L(t) = \log \varepsilon(t)$.

Грешката може да има различни типове разпределения. Най-често се разглеждат модели с нормална грешка. За динамичния ред на SOFIX получаваме, че грешката има устойчиво разпределение. Ще използваме (M) параметризацията на Золотарев [1] на това разпределение, т.е. грешката $\varepsilon$ има следната характеристична функция

$$Ee^{i\varepsilon y} = e^{-\lambda|y|^{\alpha} + i\lambda y(\gamma + \omega(y,\alpha,\beta))},$$

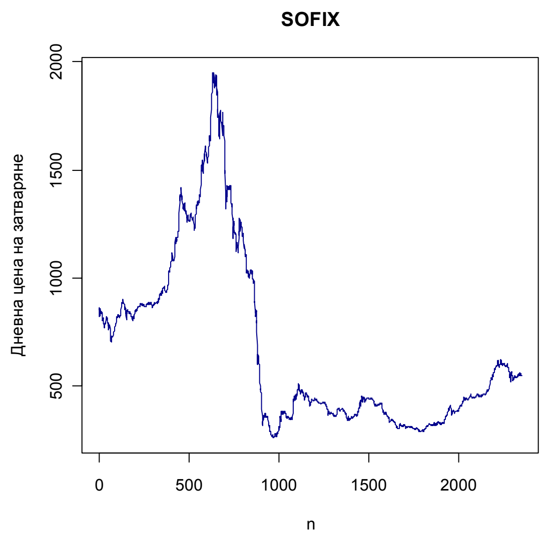
където

$$\omega(y,\alpha,\beta) = \begin{cases} (|y|^{\alpha-1} - 1)\beta.tg\left(\frac{\pi}{2}\alpha\right) & , \quad \alpha \neq 1 \\ -\beta\frac{2}{\pi}\ln|y| & , \quad \alpha = 1 \end{cases}.$$

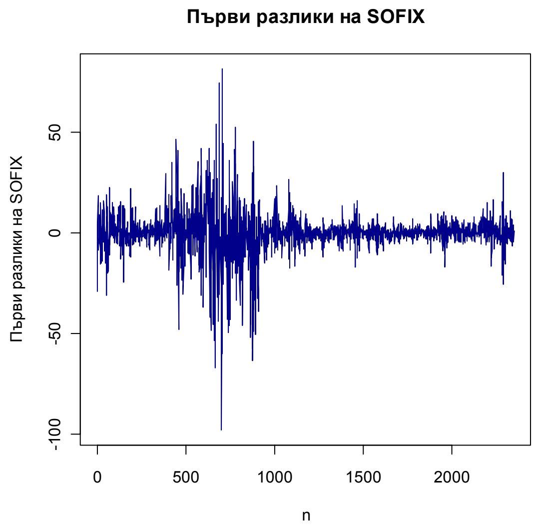
Накратко ще означаваме това с $\varepsilon \sim S_{\alpha}(\lambda, \beta, \gamma)$ .

## 3. Анализ на SOFIX

В настоящото изследване е обхванат периодът от 16 март 2005 до 19 септември 2014 г.

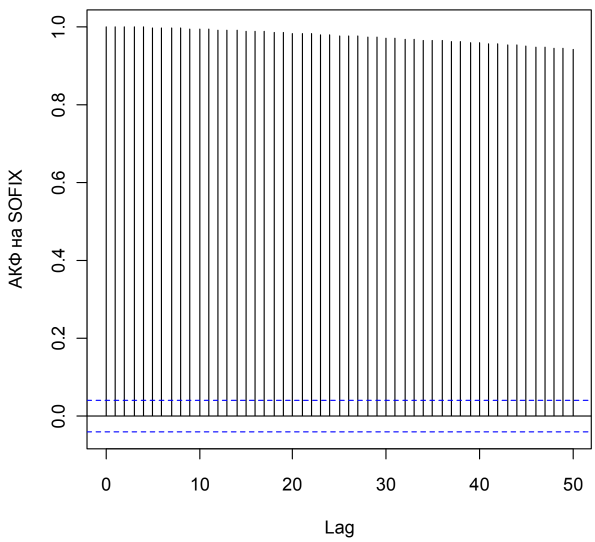



**Фиг. 1. Дневни цени на затваряне на SOFIX, за периода от 16 март 2005 до 19 септември 2014 г.**

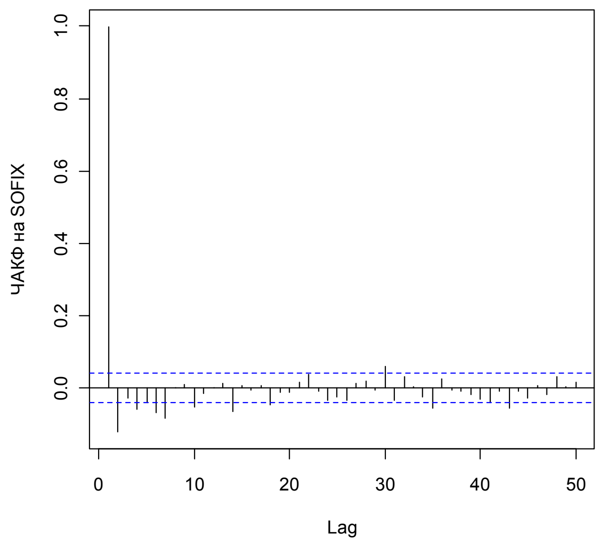



**Фиг. 2. Парви разлики на дневните цени на затваряне на SOFIX.**

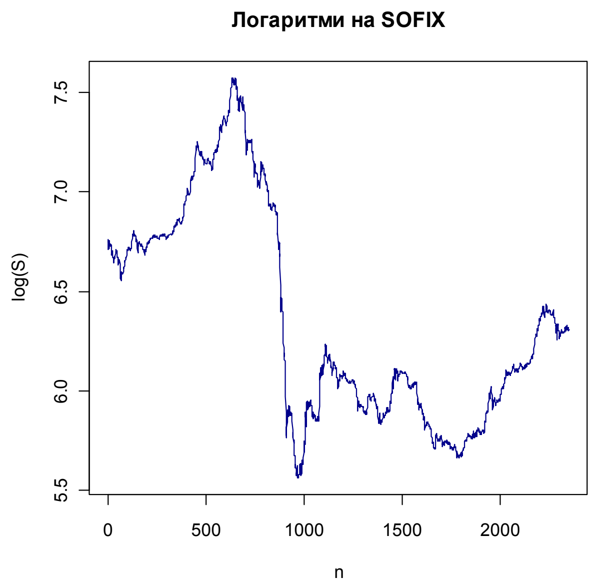
Използвани са 2349 дневни двойки наблюдения, чиято първа компонента е датата на наблюдението, втората е цената на затваряне на този индекс - S, а третата компонента е максималната дневна цена на SOFIX, M. Данните са за цената на затваряне на този индекс, визуализирани на фиг. 1. Те показват, че редът от данни не е стационарен. Разглеждаме техните първи разлики. Те са изобразени на фиг. 2 и показват, че адитивният модел не е подходящ за моделирането им поради липса на хомоскедастичност. Автокорелационната и частната автокорелационна функция на S, които са изобразени съответно на фиг. 3 и фиг. 4 показват, че SOFIX е динамичен ред с дълга памет, в който няма статистически значима цикличност, т.е. $c(t) = 0$.



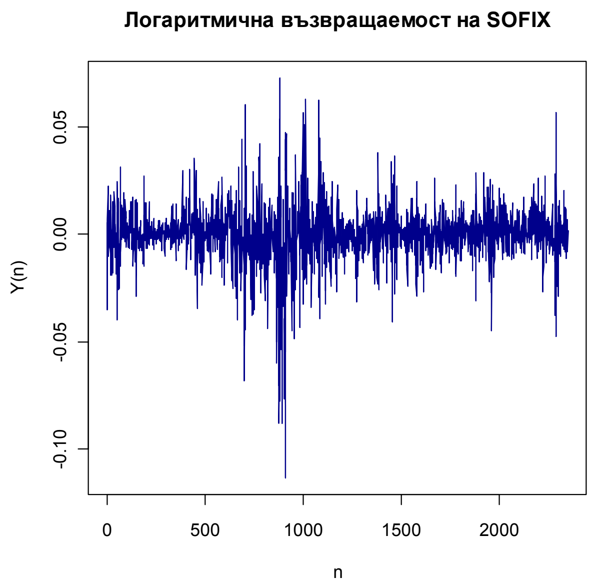

**Фиг. 3. Автокорелационна функция на SOFIX.**




**Фиг. 4. Частна автокорел. функция на SOFIX.**




**Фиг. 5. Логаритми на дневните цени на затваряне на SOFIX.**

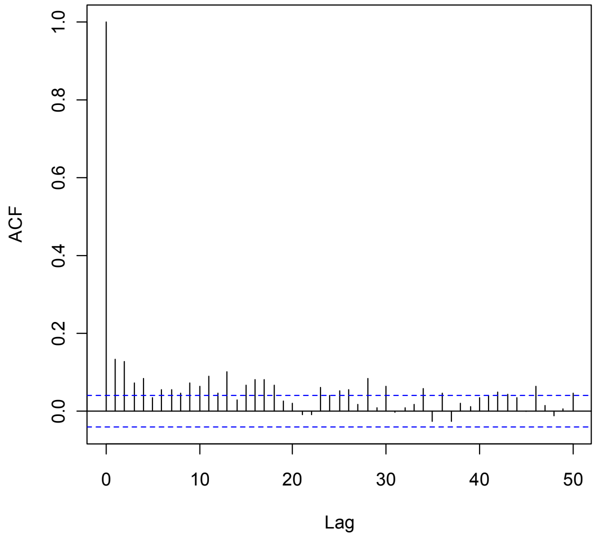



**Фиг. 6. Парви разлики на логаритмите на дневните цени на затваряне на SOFIX.**

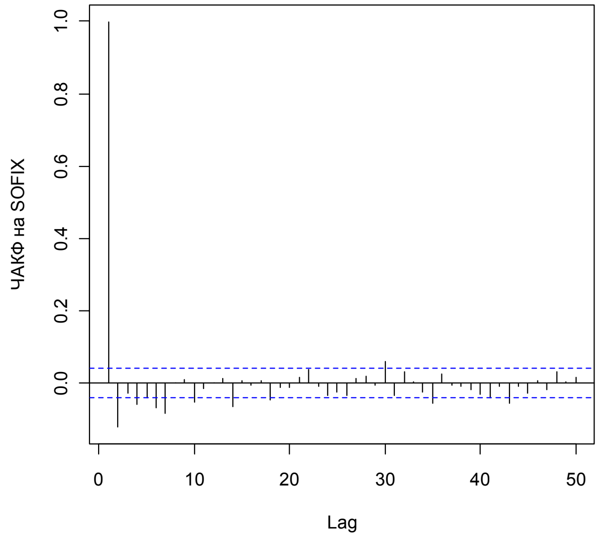
За да намалим хетероскедастичността, прагаме мултипликативния модел (2). За целта логаритмуваме данните. Графиката на логаритмите е дадена на фиг. 5, а техните първи разлики на фиг. 6. Тези първи разлики представляват дневната логаритмична възвращаемост на SOFIX, да я означим с

$$Y_L(t) = \log \frac{S(t)}{S(t-1)}.$$

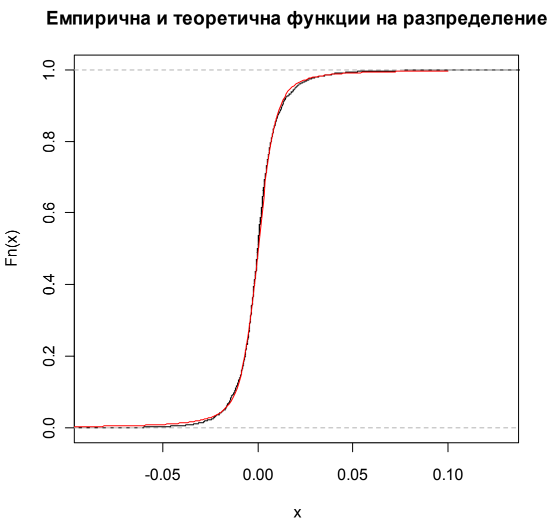
Забелязва се, че разглеждането на така трансформираните данни повишава хомоскедастичността им. Средната логаритмична възвръщаемост е 0.0001945376, при стандартното отклонение 0.01304423, т.е. можем да приемем хипотезата, че $\mu_L(t) = E(Y_L(t)) = 0$. Корелационната и автокорелационната функция на първите разлики на логаритмите са дадени на фиг. 7 и фиг. 8. Те показват, че в реда $Y_L$ не се наблюдава статистически значима цикличност.



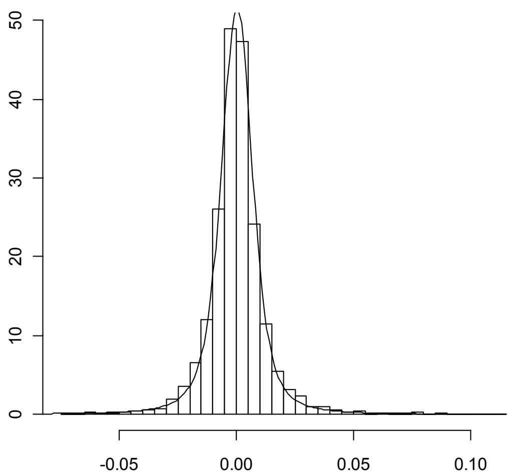

**Фиг. 7. Автокорелационна функция на първите разлики на логаритмита на SOFIX.**




**Фиг. 8. Частна автокорелационна функция на първите разлики на логаритмите на SOFIX.**




**Фиг. 9.**

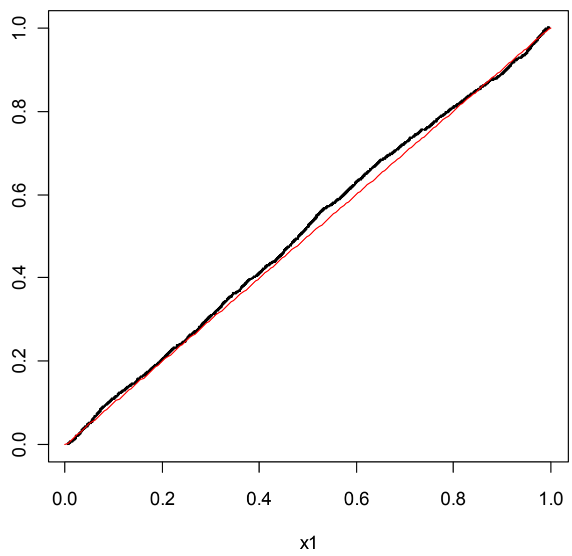


**Фиг. 10.**

Да моделираме разпределението на дневната логаритмична възвращаемост $Y_L(t)$. Използваме тестът на Колмогоров-Смирнов и Jarque-Berra статистиката и отхвърляме нулевата хипотезатезата за нормалност. Аналогичен резултат е получен и от Patev и Kanaryan [9], за периода от 24 октомври 2000 г. до 19 ноември 2004 г.

Основният принос на настоящата статия е, че правим $\alpha$-устойчив модел на грешката и получаваме, че $\varepsilon_L \sim S_{1,45}(0.0001945376,-0.0345, 0.0056)$. Проверяваме тази хипотеза с теста на Колмогоров-Смирнов и получаваме, че нямаме основание да я отхвърлим. Максималната разлика, по абсолютна стойност, между тестваната и емпиричната функция на разпределение е 0.031. Графиките на двете функции на разпределение са дадени на фиг. 9.

На фиг. 10. може да бъде наблюдавана разликата между теоретичната плътност и емпиричната плътност на разпределение. Близостта на двете разпределения се наблюдава и от графиката на съответната pp-plot, фиг.11.




**Фиг. 11.**

Така получихме модела

$$S(t) = S(t-1)\, e^{\varepsilon_L},$$

където $\varepsilon_L \sim S_{1,45}(0.0001945376,-0.0345, 0.0056)$.

**Табл. 1. Прогнозни стойности за SOFIX**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Дата | (1) | 20.09.2014 | 21.09.2014 | 22.09.2014 | 23.09.2014 | 24.09.2014 | 25.09.2014 |
| Очаквана стойност | (2) | 547.2843 | 547.2886 | 547.2929 | 547.2972 | 547.3015 | 547.3058 |
| Долна граница на доверителния интервал | (3) | 543.3509 | 538.9254 | 535.6396 | 529.7749 | 526.2876 | 522.2700 |
| Горна граница на доверителния интервал | (4) | 551.2304 | 555.7466 | 559.1539 | 565.3587 | 569.1002 | 573.4821 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (1) | 26.09.2014 | 27.09.2014 | 28.09.2014 | 29.09.2014 | 30.09.2014 | 01.10.2014 | 02.10.2014 |
| (2) | 547.3101 | 547.3144 | 547.3187 | 547.3230 | 547.3273 | 547.3316 | 547.3359 |
| (3) | 514.9156 | 511.0820 | 501.1929 | 495.1532 | 487.7901 | 483.6873 | 479.4714 |
| (4) | 581.6841 | 586.0370 | 597.5826 | 604.8729 | 613.9880 | 619.1923 | 624.6288 |

Устойчивостта на разпределението на логаритмичната възвръщаемост и липсата на крайна вариация на това разпределение показват, че този пазар е високо рисков, което е в синхрон с изследванията направени в Patev и Kanaryan [9].

Използваме този модел и правим по 100 000 симулации на този процес. На базата на тези симулации и като използваме формално Метода Монте Карло определяме прогнозната стойност за цените на затваряне на SOFIX за следващите 14 дни. Получаваме втория ред от табл. 1. Долните и горните граници на 0,95 доверителните интервали (получени по метода Монте Карло) са дадени съответно в третия и четвъртия ред на табл. 1. Поради несъществуването на дисперсия на разпределението на $\varepsilon_L$ данните не удовлетворяват изискванията за нормалност и от там границите на доверителните интервали, получени по метода Монте Карло подценяват действителните стойности на горните граници и надценяват действителните стойности на долните граници на 0,95 доверителните интервали на цените на затваряне като се използва $\alpha$ - устойчивото разпределение.

## 4. Изводи

В настоящото изследване е потвърдено, че търгодията със SOFIX е високо рискова, логаритмичната възвръщаемост, не се подчинява на условието за нормалност на разпределението и това е пазар с дълга памет. Приносът на изследването е, че доказваме, че $\alpha$-устойчивото разпределение е един много добър модел на логаритмичната възвръщаемост на SOFIX. Това поражда още много възможности за построяване на нови по-добри модели, например експоненциално претеглени модели или по-общи GARCH модели с $\alpha$-устойчивото разпределение на грешката. Тези модели са обект на бъдещи изследвания. Друга възможност за бъдещи изследвания е прилагане на обобщените разпределения на екстремалните стойности към разпределението на надхвърлянията на дадено ниво от логаритмичната възвръщаемост и оценка на квантили извън обхвата на данните. Аналогични техники към други екстремални събития и величини са използвани от Embrechts, Kluepelberg Mikosch [6].

**ЛИТЕРАТУРА**

1. **Золотарев**, В. М. Одномерные устойчивые распределения. // Наука, Москва, 1983.
2. **Ломев**, Б., И. Иванов, А. Мицанов. Изследване на наличието на „тежки опашки” и моделиране на GARCH модели за SOFIX и други световни борсови индекси. // Сборник от доклади на Международната научна конференция УНИНЕХ’08, Габрово, 2008, стр. III – 57-62.
3. **Милинов**, В., Н. Канарян. Връзката риск-обем на търговия на българския фондов пазар.// Бизнес управление, 2006, № 1.
http://www.uni-svishtov.bg/dialog_old/2006/1.06Milinov-Kanaryan-2.pdf
4. **Петков,** Пл. Иконометричен анализ на българските фондови борсови индекси с GARCH(1, 1) модели.// http://homeworking.hit.bg/Economics2010_PPetkov.pdf
5. **Ценков**, Вл. Обобщени факти около капиталовата възвръщаемост: по примера на индекса SOFIX.// Икономика и управление, год. V, № 1, 2009, с. 34-48.
6. **Embrechts**, P., Cl. Klüppelberg, Th. Mikosch. //Modelling Extremal Events for Insurance and Finance, 1997, Springer Verlag.

7. **Lomev**, B., I. Ivanov. Self-similarity, heavy tails and long-range dependence as measures for financial market inefficiency-the case of Bulgaria.//The 9th Hellenic European Research on Computer Mathematics and its Applications Conference, Athens,(a conference talk). 2009. http://www.aueb.gr/pympe/hercma/proceedings2009/H09-FULL-PAPERS-1/LOMEV-IVANOV-1.pdf
8. **Lomev**, B., I. Ivanov, B. Bogdanova. What can wavelets reveal about SOFIX?.// Journal of Engineering Science and Technology Review 4.3, 2011, 233-236.
9. **Patev** Pl., N. Kanaryan. Modelling and forecasting the volatility of thin emerging stock markets: the case of Bulgaria.// http://bnb.bg/bnbweb/groups/public/documents/bnb_download/p_e_sc_third_annual_conf_a6_bg.pdf
10. **Samitas**, A., D. Kenourgios, Nikos P. Short and long run parametric dynamics in the Balkans stock markets. // Available at SSRN 993756, 2006.
11. **Shahpazov**, V., L. Forecasting Price Movement of SOFIX Index on the Bulgarian Stock Excange – Sofia using an Artificial Neural Network Model.// SCITEPRESS Digital Library, http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0004776802980302
12. URL: http://www.r-project.org