ISSN 1648-8776
JAUNŲJŲ MOKSLININKŲ DARBAI. Nr. 1 (45). 2016 Doi: 10.21277/jmd.v1i45.43

# DAUGIAMAČIO $\alpha$-STABILIOJO DĖSNIO PARAMETRŲ VERTINIMAS DIDŽIAUSIO TIKĖTINUMO METODU

**Leonidas Sakalauskas**[1], **Ingrida Vaičiulytė**[2]
*[1]Šiaulių universitetas, [2]Šiaulių valstybinė kolegija*

## Įvadas

Daugelis statistinių išvadų remiasi prielaida, kad stebimas atsitiktinis dydis turi normalųjį skirstinį – tai vienas plačiausiai naudojamų matematinėje statistikoje skirstinių. Tačiau nustatyta, kad finansiniai duomenys dažnai yra leptokurtotiniai ir pasižymi išskirtimis, tad ekonominių rodiklių analizei taip dažnai taikomas normalusis dėsnis nėra pats geriausias variantas, todėl siūloma jį keisti bendresniais, pavyzdžiui, stabiliaisiais modeliais (Rachev ir Mittnik, 2000; Kabašinskas ir kt., 2012; Sakalauskas ir kt., 2013). Skirstinių, susijusių su $\alpha$-stabiliaisiais, tyrimai šiuo metu yra ypač aktualūs, nes jie dažnai pasitaiko analizuojant verslo ir finansų duomenis ar informacinius srautus kompiuterių tinkluose. Kadangi Rachevo ir Mittniko (2000), Kabašinsko ir kt. (2012) bei Sakalausko ir kt. (2013) darbuose $\alpha$-stabiliojo dėsnio parametrų vertinimas yra ištirtas tik vienmačiu atveju, praktikoje iškilo problema, kaip vertinti daugiamačius duomenis. Nors problema tiriama ir nagrinėjama jau kelis dešimtmečius, tačiau iki galo ji dar neišsemta (Press, 1972; Davydov ir Paulauskas, 1999; Nolan, 1998; Kring ir kt., 2009; Ogata, 2013). Šiame darbe sudarytas didžiausio tikėtinumo metodas daugiamačio $\alpha$-stabiliojo dėsnio parametrų įvertiniams gauti pritaikius EM algoritmą.

Vienmačiu atveju yra žinoma, kad

$$s = \sqrt{s_1} \cdot s_2, \tag{1}$$

čia

$s_1$– stabilusis atsitiktinis dydis, kurio asimetrijos parametras $\beta = 1$, o formos parametras $\alpha_1 < 1$;

$s_2$ – kitas stabilusis atsitiktinis dydis su nuliniu vidurkiu, nepriklausomas nuo $s_1$, kurio asimetrijos parametras $\beta = 0$, o formos parametras $\alpha_2$;

$s$ – stabilusis atsitiktinis dydis, kurio asimetrijos parametras $\beta = 0$, o formos parametras $\alpha = \alpha_1 \cdot \alpha_2$ (Rachev ir Mittnik, 1993; Samorodnitsky ir Taqqu, 1994; Ravishanker ir Qiou, 1999).

Taikant šį būdą paprastai parenkama, kad $s_2$ būtų atsitiktinis dydis, pasiskirstęs pagal normalųjį dėsnį, t. y. $\alpha_1 = \frac{\alpha}{2}$ ir $\alpha_2 = 2$. Atsitiktinis stabilusis dydis, kai $\alpha_1 < 1$, o $\beta = 1$, yra vadinamas subordinatoriumi ir jis visada įgyja tik teigiamas reikšmes (Rachev ir Mittnik, 1993; Ravishanker ir Qiou, 1999). Pritaikius šį metodą daugiamačiu atveju gaunamas su priklausomomis komponentėmis atsitiktinis vektorius, kuriuo galima modeliuoti duomenis su didelėmis išskirčių tikimybėmis (Nolan, 2007; Sakalauskas ir Vaičiulytė, 2014).

Taikant šį metodą galima gauti daugiamatį stabilųjį simetrinį vektorių (Rachev ir Mittnik, 1993; Ravishanker ir Qiou, 1999):

$$X = \mu + \sqrt{s_1} \cdot s_2, \tag{2}$$

čia

$s_1$ – subordinatorius su parametru $\alpha$,

$s_2$ – atsitiktinis vektorius, pasiskirstęs pagal normalųjį dėsnį $N(0,\Omega)$,

$\mu$ – atsitiktinis vidurkio vektorius.

## Didžiausio tikėtinumo metodu gaunami įvertiniai

Tegu $X = (X^1, X^2, \ldots, X^K)$ yra nepriklausomų $d$-mačių stabiliųjų vektorių imtis. Šios imties tikėtinumo funkcija užrašoma taip (Ravishanker ir Qiou, 1999):

$$f(\mu,\Omega,x,y,s) = \frac{s^{-\frac{p}{2}}}{(2\pi)^{\frac{p}{2}} \cdot |\Omega|^{\frac{1}{2}}} \frac{\alpha \cdot s^{\frac{2}{\alpha-2}}}{2\cdot(2-\alpha)} \cdot U_\alpha(y) \cdot \exp\left[\frac{(x-\mu)^T \cdot \Omega^{-1} \cdot (x-\mu)}{2\cdot s} - s^{\frac{\alpha}{\alpha-2}} \cdot U_\alpha(y)\right], \tag{3}$$

čia

$$U_\alpha(y)=\frac{\sin\left(\frac{\pi}{4}\cdot\alpha\cdot(y+1)\right)^{\frac{\alpha}{2-\alpha}}\cdot\cos\left(\frac{\pi}{4}\cdot(\alpha-(2-\alpha)\cdot y)\right)}{\cos\left(\pi\cdot\frac{y}{2}\right)^{\frac{2}{2-\alpha}}\cdot\cos\left(\frac{\pi\cdot\alpha}{4}\right)^{\frac{2}{2-\alpha}}}. \tag{4}$$

Įvedus keitinį

$$z_i=|s_i|^{\frac{\alpha}{\alpha-2}}\cdot U_\alpha(y_i), \tag{5}$$

logaritminė tikėtinumo funkcija yra:

$$L(X,\mu,\Omega)=\sum_{i=1}^{K}\ln\left(\int_0^\infty \exp\{-z_i\}\int_{-1}^{1}B(X^i,y_i,z_i,\mu,\Omega)dy_i dz_i\right) \tag{6}$$

čia

$$B(X^i,y_i,z_i,\mu,\Omega)=\frac{1}{2\cdot(2\pi)^{\frac{p}{2}}\cdot|\Omega|^{\frac{1}{2}}}\cdot\left(\frac{z_i}{U_\alpha(y_i)}\right)^{\frac{p\cdot(2-\alpha)}{2\cdot\alpha}}\cdot\exp\left\{-\frac{(X^i-\mu)^T\Omega^{-1}(X^i-\mu)}{2}\cdot\left(\frac{z_i}{U_\alpha(y_i)}\right)^{\frac{2-\alpha}{\alpha}}\right\}.$$

Norint rasti daugiamačio $\alpha$-stabiliojo dėsnio parametrų $\mu$, $\Omega$ įverčius didžiausio tikėtinumo metodu, reikia išspręsti šią lygčių sistemą:

$$\begin{cases}\dfrac{\partial L(X,\mu,\Omega)}{\partial\mu}=-\displaystyle\sum_{i=1}^{K}\frac{\partial f(X^i,\mu,\Omega)}{\partial\mu}\cdot\frac{1}{f(X^i,\mu,\Omega)}=0,\\ \dfrac{\partial L(X,\mu,\Omega)}{\partial\Omega}=-\displaystyle\sum_{i=1}^{K}\frac{\partial f(X^i,\mu,\Omega)}{\partial\Omega}\cdot\frac{1}{f(X^i,\mu,\Omega)}=0.\end{cases} \tag{7}$$

Pagal fiksuoto taško metodą parametrų įverčiai gaunami iš lygčių:

$$\hat{\mu}=\frac{\sum_{i=1}^{K}\frac{X^i\cdot g_i}{f_i}}{\sum_{i=1}^{K}\frac{g_i}{f_i}} \tag{8}$$

$$\hat{\Omega}=\frac{1}{K}\cdot\sum_{i=1}^{K}\frac{(X^i-\hat{\mu})(X^i-\hat{\mu})^T g_i}{f_i} \tag{9}$$

čia

$$g(X,\hat{\mu},\hat{\Omega})=\int_0^\infty\int_{-1}^{1}\left(\frac{z_i}{U_\alpha(y_i)}\right)^{\frac{2-\alpha}{\alpha}}\cdot B(X,y,z,\hat{\mu},\hat{\Omega})\cdot\exp\{-z\}dydz \tag{10}$$

$$f(X,\hat{\mu},\hat{\Omega})=\int_0^\infty\int_{-1}^{1}B(X,y,z,\hat{\mu},\hat{\Omega})\cdot\exp\{-z\}dydz \tag{11}$$

Integralai, įeinantys į įvertinių išraiškas, darbe apskaičiuojami pasinaudojus Gauso ir Lagero bei Gauso kvadratūrinėmis formulėmis (Ehrich, 2002; Stoer ir Bulirsch, 2002; Kovvali, 2012; Casio Computer co., 2015):

$$\int_0^\infty x^\alpha e^{-x} f(x)dx \cong \sum_{i=1}^{n} \omega_i f(x_i), \tag{12}$$

$$\int_{-1}^{1} f(\chi)d\chi \cong \sum_{i=1}^{m} \vartheta_i f(\chi_i), \tag{13}$$

čia $f(x_i)$ – integruojama funkcija, $n$, $m$ – mazgų skaičius, $x_i$, $\chi_i$ – integravimo mazgai, $\omega_i$, $\vartheta_i$ – fiksuoti svoriai.

**Kompiuterinis modeliavimas**

Parametrų didžiausio tikėtinumo įvertinių apskaičiavimas, sprendžiant skaitiniais metodais, yra iteracinis procesas, kuriam reikia parinkti pradines įvertinių reikšmes ir atlikti tam tikrą iteracijų kiekį, kol gretimų iteracijų įverčiai pradės skirtis pakankamai mažai. Tiriant algoritmo elgseną buvo atlikti eksperimentai su parinktais pradiniais duomenimis:

$$p = 2, \quad \alpha = 1{,}5, \quad \mu = \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \quad \Omega = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}. \tag{14}$$

Norint palyginti kelis integralų tikėtinumo funkcijos išraiškoje (10), (11) apskaičiavimo metodus, buvo sugeneruota imtis su $K$ = 50 $\alpha$-stabiliojo dėsnio reikšmių ir atlikti skaičiavimai pagal aprašytą metodą trimis variantais, t. y.:

I variantas – kai (10), (11) integralai skaičiuojami pasinaudojus matematine MathCad integralų apskaičiavimo programa;

II variantas – kai integralai skaičiuojami pasinaudojus MathCad programa bei Gauso ir Lagero kvadratūrinėmis formulėmis;

III variantas – kai integralai skaičiuojami pasinaudojus Gauso bei Gauso ir Lagero kvadratūrinėmis formulėmis.

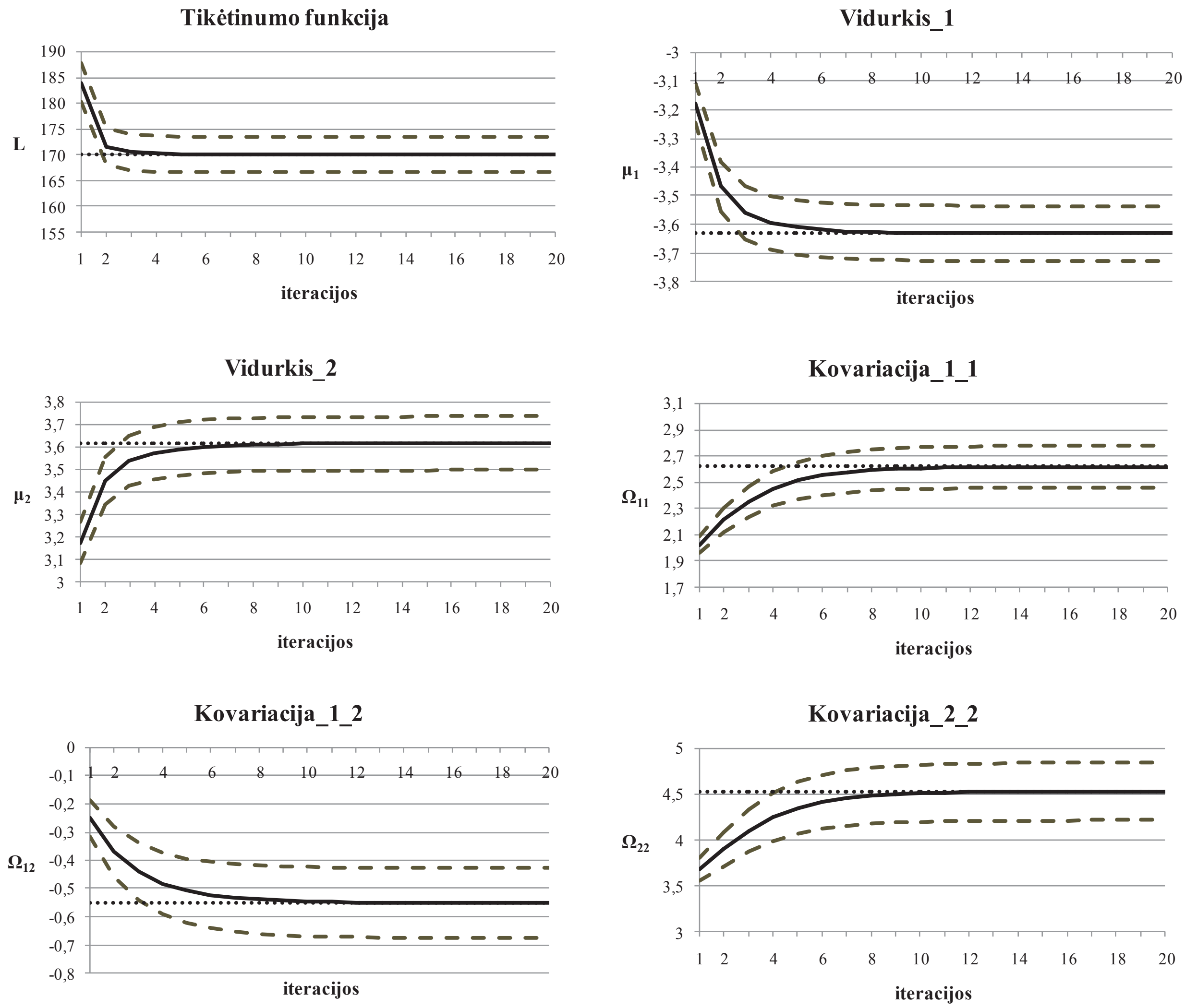
1 lentelėje pateikti gauti $\alpha$-stabiliojo dėsnio parametrų įverčiai ir kiekvieno algoritmo skaičiavimo trukmės palyginimas.

1 lentelė. ***α*-stabiliojo dėsnio parametrų skaičiavimo algoritmų trukmės palyginimas**

| Variantas | $\mu_1$ | $\mu_2$ | $\Omega_{11}$ | $\Omega_{12}$ | $\Omega_{22}$ | Tikėtinumo funkcija | Trukmė |
|---|---|---|---|---|---|---|---|
| I | -4,142 | 3,541 | 3,045 | -0,797 | 6,343 | 176,333 | 6 val. 10 min. |
| II | -4,150 | 3,546 | 3,060 | -0,800 | 6,408 | 176,336 | 27 min. |
| III | -4,150 | 3,546 | 3,060 | -0,800 | 6,408 | 176,336 | 8 min. |

Atlikus šiuos skaičiavimo eksperimentus nuspręsta pasirinkti trečią skaičiavimo variantą, nes paklaida yra priimtina (tikėtinumo funkcijos paklaida yra šeštame ženkle), o skaičiavimo trukmė trumpiausia.

Žinoma, kad didžiausio tikėtinumo metodu gauti stabiliojo dėsnio įverčiai asimptotiškai yra pasiskirstę pagal normalųjį dėsnį (DuMouchel, 1973). Tokiu būdu šiems įverčiams tirti Monte Karlo metodu pakanka sugeneruoti $M$ = 100 imčių. Kiekviena imtis sudaryta iš $K$ = 100 $\alpha$-stabiliojo dėsnio reikšmių. Pasinaudojus (12) ir (13) kvadratūrinėmis formulėmis apskaičiuoti (10) ir (11) integralai. Pritaikius MathCad minimizavimo paprogramę buvo rasti nagrinėjamo modelio didžiausio tikėtinumo įverčiai kiekvienai imčiai. Po to buvo atlikta po 20 iteracijų sukurtuoju metodu su pradinėmis reikšmėmis (14) ir gauti keli grafikai (žr. 1 pav.), kuriuose parametrų įverčiai pavaizduoti ištisine linija, taškiukais – optimali reikšmė, o 95 proc. pasikliautinasis intervalas, apskaičiuotas pasinaudojus normaline aproksimacija, – punktyrinėmis linijomis. 1 pav. matome, kad tikėtinumo funkcijos reikšmė ir parametrų įverčiai jau po kelių iteracijų konverguoja į vertes, apskaičiuotas su MathCad minimizavimo paprograme.




**1 pav.** Dvimačio $\alpha$-stabiliojo dėsnio parametrų įverčiai

## Išvados

1. Darbe sudarytas daugiamačio $\alpha$-stabiliojo dėsnio didžiausio tikėtinumo metodas, leidžiantis įvertinti šio skirstinio parametrus, pasinaudojant EM algoritmu.
2. Skaitinio modeliavimo metodu gauti $\alpha$-stabiliojo dėsnio parametrų įvertiniai yra statistiškai adekvatūs, kadangi tikėtinumo funkcijos reikšmė ir parametrų įverčiai po kelių iteracijų konverguoja į didžiausio tikėtinumo vertes.
3. Sudarytas metodas gali būti taikomas atliekant akcijų rinkų duomenų analizę, sudarant finansinius modelius.

## Literatūra


1. Casio Computer co., 2015. Prieiga per internetą: <http://keisan.casio.com/exec/system/1281279441>. [Žiūrėta 2015-08-06].
2. Davydov Yu., Paulauskas V., 1999, On the Estimation of the Parameters of Multivariate Stable Distributions. *Acta Applicandae Mathematica*. Vol. 58. No. 1. P. 107–124.
3. DuMouchel W. H., 1973, On the Asymptotic Normality of the Maximum-Likelihood Estimate when Sampling from a Stable Distribution. *The Annals of Statistics*. Vol. 1. No. 5. P. 948–957.
4. Ehrich S., 2002, On stratified extensions of Gauss-Laguerre and Gauss-Hermite quadrature formulas. *Journal of Computational and Applied Mathematics*. Vol. 140. No. 1–2. P. 291–299.
5. Kabašinskas A., Sakalauskas L., Sun E. W., Belovas I., 2012, Mixed-Stable Models for Analyzing High-Frequency Financial Data. *Journal of Computational Analysis and Applications*. Vol. 14. No. 7. P. 1210–1226.
6. Kovvali N., 2012, *Theory and Applications of Gaussian Quadrature Methods*. Morgan & Claypool Publishers.
7. Kring S., Rachev S. T., Hochstotter M., Fabozzi F. J., 2009, Estimation of $\alpha$-Stable Sub-Gaussian Distributions for Asset Returns. *Risk Assessment: Decisions in Banking and Finance*. P. 111–152.
8. Nolan J., 1998, Multivariate stable distributions: approximation, estimation, simulation and identification. *A practical guide to heavy tails*. P. 509–525.

9. Nolan J. P., 2007, *Stable Distributions – Models for Heavy Tailed Data*. Boston: Birkhauser.
10. Ogata H., 2013, Estimation for multivariate stable distributions with generalized empirical likelihood. *Journal of Econometrics*. Vol. 172. No. 2. P. 248–254.
11. Press S. J., 1972, Estimation in Univariate and Multivariate Stable Distributions. *Journal of the American Statistical Association*. Vol. 67. No. 340. P. 842–846.
12. Rachev S. T., Mittnik S., 1993, Modeling Asset Returns with Alternative Stable Distributions. *Econometric Reviews*. Vol. 12. No. 3. P. 261–330.
13. Rachev S. T., Mittnik S., 2000, *Stable Paretian Models in Finance*. New York: Wiley.
14. Ravishanker N., Qiou Z., 1999, Monte Carlo EM Estimation for Multivariate Stable Distributions. *Statistics & Probability Letters*. Vol. 45. No. 4. P. 335–340.
15. Sakalauskas L., Kalsyte Z., Vaiciulyte I., Kupciunas I., 2013, The application of stable and skew *t*-distributions in predicting the change in accounting and governance risk ratings. *Proceedings of the 8th International Conference „Electrical and Control Technologies“*. P. 53–58.
16. Sakalauskas L., Vaičiulytė I., 2014, Subgausinio vektoriaus skirstinio parametrų vertinimas Monte Karlo Markovo grandinės metodu. *Jaunųjų mokslininkų darbai*. Nr. 1 (41). P. 104–107.
17. Samorodnitsky G., Taqqu M. S., 1994, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman and Hall.
18. Stoer J., Bulirsch R., 2002, *Introduction to Numerical Analysis*. Springer Science & Business Media.

## Summary

## MAXIMUM LIKELIHOOD METHOD FOR THE ESTIMATION OF MULTIVARIATE ALPHA-STABLE DISTRIBUTION

***L. Sakalauskas, I. Vaičiulytė***

Research of alpha-stable distributions is especially important nowadays because they often occur in the analysis of financial data and information flows along computer networks. It has been found that financial data are often leptokurtic with heavy-tailed dsitributions; many authors, e.g., Rachev, Mittnik (2000), Kabašinskas *et al.* (2012), Sakalauskas *et al.* (2013) have proved that the most often used normal distribution is not the most suitable way to analyse economic indicators and suggested to replace it with more general, for example stable distributions. Since Rachev, Mittnik (2000), Kabašinskas *et al.* (2012), Sakalauskas *et al.* (2013) have estimated one-dimensional alpha-stable distributions a problem arises how to estimate multidimensional data. Although the problem has been investigated and analyzed for several decades it has not been solved yet (Press, 1972; Davydov, Paulauskas, 1999; Nolan, 1998; Kring *et al.* 2009; Ogata, 2013). Maximum likelihood method for the estimation of multivariate alpha-stable distributions by using the EM algorithm is presented in this work. Integrals included in the expressions of the estimates have been calculated using the Gaussian and Legendre-Gauss quadrature formulas. The constructed model can be used in stock market data analysis.

**Keywords:** maximum likelihood method, alpha-stable distribution, Monte Carlo method, statistical modeling.

## Santrauka

## DAUGIAMAČIO $\alpha$-STABILIOJO DĖSNIO PARAMETRŲ VERTINIMAS DIDŽIAUSIO TIKĖTINUMO METODU

***L. Sakalauskas, I. Vaičiulytė***

Skirstinių, susijusių su $\alpha$-stabiliaisiais, tyrimai šiuo metu yra ypač aktualūs, nes jie dažnai pasitaiko analizuojant verslo ir finansų duomenis ar informacinius srautus kompiuterių tinkluose. Nustatyta, kad finansiniai duomenys dažnai yra leptokurtotiniai ir pasižymi išskirtimis, tad daugelis autorių (Rachev ir Mittnik, 2000; Kabašinskas ir kt., 2012; Sakalauskas ir kt., 2013) parodė, jog dažnai taikomas normalusis skirstinys nėra pats geriausias variantas ekonominių rodiklių analizei atlikti, todėl siūloma jį keisti bendresniais, pavyzdžiui, stabiliaisiais, modeliais. Kadangi Rachevo ir Mittniko (2000), Kabašinsko ir kt. (2012) bei Sakalausko ir kt. (2013) darbuose $\alpha$-stabiliojo dėsnio parametrų vertinimas yra ištirtas tik vienmačiu atveju, praktikoje iškilo problema, kaip vertinti daugiamačius duomenis. Nors problema tiriama ir nagrinėjama jau kelis dešimtmečius, tačiau iki galo ji dar neišsemta (Press, 1972; Davydov ir Paulauskas, 1999; Nolan, 1998; Kring ir kt., 2009; Ogata, 2013). Šiame darbe sudarytas didžiausio tikėtinumo metodas daugiamačio $\alpha$-stabiliojo dėsnio parametrų įvertiniams gauti pritaikius EM algoritmą. Integralai, įeinantys į įvertinių išraiškas, apskaičiuojami pasinaudojus Gauso bei Gauso ir Lagero kvadratūrinėmis formulėmis. Sukonstruotas modelis gali būti taikomas akcijų rinkų duomenų analizei atlikti.

**Prasminiai žodžiai:** didžiausio tikėtinumo metodas, $\alpha$-stabilusis skirstinys, Monte Karlo metodas, statistinis modeliavimas.

Įteikta 2016-01-15
Priimta 2016-06-07