Дослідження в математиці і механіці. – 2023. – Т. 28, вип. 1–2 (41–42). – С. 159–171

УДК 519.213

Є. В. Турчин

Дніпровський національний університет імені Олеся Гончара

ОДНЕ НОВЕ УЗАГАЛЬНЕННЯ РОЗПОДІЛУ КОШІ

У роботі розглядається новий розподіл із важкими хвостами — p-узагальнений розподіл Коші. Ця нова 5-параметрична сім'я розподілів є значно гнучкішою порівняно з класичним розподілом Коші, зокрема, до неї входять і асиметричні розподіли. Розглянуті різні числові характеристики нових розподілів, зокрема, моменти дробового порядку та "подвійно неповні" моменти. Також отримані (числовими методами) значення нормованих центральних моментів вищих порядків, що базуються на квантилях — асиметрії Bowley та ексцеса Moors. Придатність p-узагальненого розподілу Коші до моделювання реальних даних підтверджена підгонкою цього розподілу до ряду приростів логарифмів цін акцій. При цьому для p-узагальненого розподілу Коші отримано менше значення статистики AIC, ніж для звичайного розподілу Коші, асиметричного розподілу Коші, узагальненого логістичного розподілу та гіперболічного розподілу.

MSC: 62E10, 62E15.

Ключові слова: розподіл Коші, неповні моменти, асиметрія Bowley, ексцес Moors.

DOI: https://doi.org/10.18524/2519-206X.2023.1-2(41-42).296368.

Вступ

Розподіл Коші часто відносять до "екзотичних" розподілів, які, по суті, не зустрічаються на практиці (хоча насправді це не зовсім так — див., наприклад, [7], гл. 9). Тому виправданими є дослідження, присвячені аналогам чи узагальненням класичного розподілу Коші. Сюди відносяться як різні варіанти асиметризації розподілу Коші, так і інші його модифікації. Серед робіт, присвячених цій тематиці, варто згадати, зокрема, монографію [2], статті [1], [3], [8], [10], [11], та [15]. Нові узагальнення розподілу Коші залишаються актуальними, враховуючи важливість розподілів із важкими хвостами для моделювання різноманітних реальних явищ (в особливості це стосується економічних даних).

У даній статті вводиться ще один аналог розподілу Коші — так званий p-узагальнений розподіл Коші. Метою роботи є дослідження властивостей цього розподілу.

Надійшла 05.03.2023

Основні результати

1. Щільність та функція розподілу.

Означення 1. Щільність p-узагальненого розподілу Коші з параметрами $(1, 0, c, \gamma_1, \gamma_2)$, де $c \geq 0, \gamma_1 \geq 0$, $\gamma_2 \geq 0$, означається наступним чином:

$$p_{1,0,c,\gamma_1,\gamma_2}(x) = \begin{cases} \dfrac{1}{A(1+x^2)} \left(1 + \dfrac{c}{(1+x^2)^{\gamma_1}}\right), & x \leq 0; \\ \dfrac{1}{A(1+x^2)} \left(1 + \dfrac{c}{(1+x^2)^{\gamma_2}}\right), & x > 0, \end{cases} \tag{1}$$

де

$$A = \pi + c \cdot \frac{\sqrt{\pi}}{2} \left(\frac{\Gamma(1/2+\gamma_1)}{\Gamma(1+\gamma_1)} + \frac{\Gamma(1/2+\gamma_2)}{\Gamma(1+\gamma_2)} \right). \tag{2}$$

Щільність p-узагальненого розподілу Коші з параметрами $(a, b, c, \gamma_1, \gamma_2)$, де
$a > 0$, означимо так:

$$p_{a,b,c,\gamma_1,\gamma_2}(x) = \frac{1}{a} p_{1,0,c,\gamma_1,\gamma_2}((x-b)/a).$$

Зауваження 1. Далі вживатимемо позначення $pgC(a, b, c, \gamma_1, \gamma_2)$ для p-узагальненого розподілу Коші з параметрами $(a, b, c, \gamma_1, \gamma_2)$.

Зауваження 2. Коректність означення випливає з формули

$$\int_0^\infty \frac{dx}{(1+x^2)^\gamma} = \frac{\sqrt{\pi}}{2} \frac{\Gamma(1/2+\gamma)}{\Gamma(1+\gamma)}$$

Графіки щільностей p-узагальненого розподілу Коші з параметрами $(1, 0, 10, 1, 5)$, $(1, 0, 5, 0.2, 0.8)$ та $(1, 0, 100, 2, 7)$ наведені відповідно на рис. 1, 2 та 3.

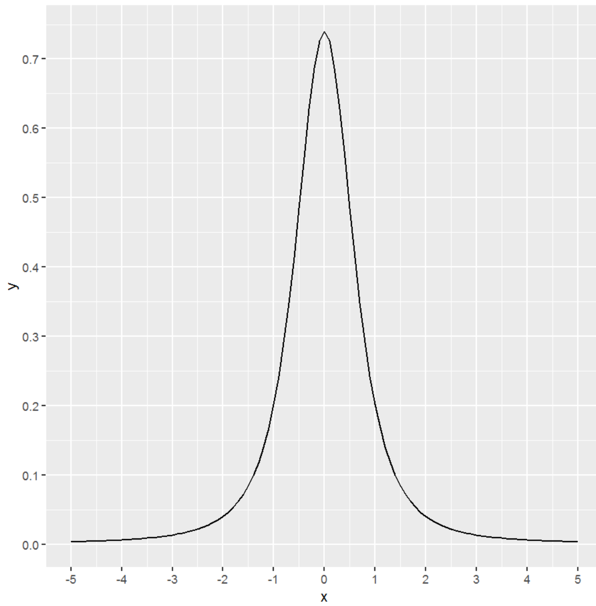

Рис. 1: Щільність $pgC(1, 0, 10, 1, 5)$.

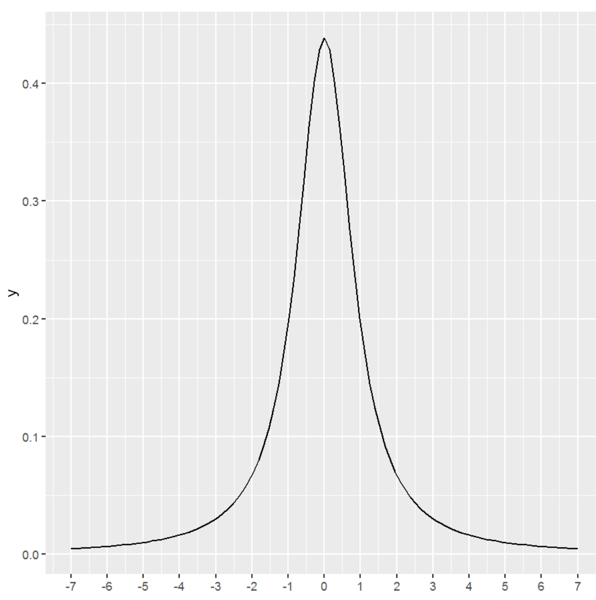

Рис. 2: Щільність $pgC(1, 0, 5, 0.2, 0.8)$.

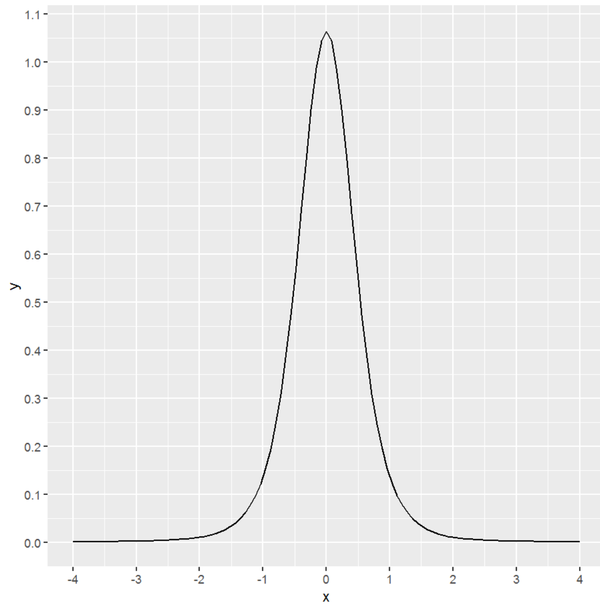

Рис. 3: Щільність $pgC(1, 0, 100, 2, 7)$.

Теорема 1. *Функція розподілу для p-узагальненого розподілу Коші з параметрами* $(a, b, c, \gamma_1, \gamma_2)$

$$F_{a;b;c,\gamma_1,\gamma_2}(x) = \begin{cases} c_1 + \dfrac{1}{A}\operatorname{arctg}\left(\dfrac{x-b}{a}\right) + \\ \qquad + \dfrac{c(x-b)}{Aa}\ {}_2F_1\left(\dfrac{1}{2}, 1+\gamma_1; \dfrac{3}{2}; -(\dfrac{x-b}{a})^2\right), x \le b; \\ c_1 + \dfrac{1}{A}\operatorname{arctg}\left(\dfrac{x-b}{a}\right) + \\ \qquad + \dfrac{c(x-b)}{Aa}\ {}_2F_1\left(\dfrac{1}{2}, 1+\gamma_2; \dfrac{3}{2}; -(\dfrac{x-b}{a})^2\right), x > b. \end{cases} \tag{3}$$

де A означено у (2),

$$c_1 = A^{-1}\pi/2 + cA^{-1}\sqrt{\pi}\frac{\Gamma(1/2+\gamma_1)}{2\Gamma(1+\gamma_1)},$$

${}_2F_1$ *— гіпергеометрична функція.*

Доведення. Достатньо перевірити (3) для $a = 1$ та $b = 0$, тобто переконатися, що

$$F_{1;0;c,\gamma_1,\gamma_2}(x) = \begin{cases} c_1 + A^{-1}\operatorname{arctg} x + \\ \qquad + cA^{-1} x\, {}_2F_1(1/2, 1+\gamma_1; 3/2; -x^2), x \le 0; \\ c_1 + A^{-1}\operatorname{arctg} x + \\ \qquad + cA^{-1} x\, {}_2F_1(1/2, 1+\gamma_2; 3/2; -x^2), x > 0. \end{cases} \tag{4}$$

Розглянемо спочатку випадок $x \le 0$. Маємо:

$$F_{1;0;c,\gamma_1,\gamma_2}(x) = A^{-1}\operatorname{arctg} x\Big|_{-\infty}^{x} + cA^{-1}\int_{-\infty}^{x} \frac{dy}{(1+y^2)^{1+\gamma_1}}.$$

Але

$$\int_{-\infty}^{x} \frac{dy}{(1+y^2)^{1+\gamma_1}} = \frac{\sqrt{\pi}}{2}\frac{\Gamma(1/2+\gamma_1)}{\Gamma(1+\gamma_1)} + x\, {}_2F_1(1/2, 1+\gamma_1; 3/2; -x^2)$$

(інтеграл знайдено за допомогою *Mathematica* 11.1) і формулу (4) для випадку $x \le 0$ доведено.

Випадок $x > 0$ розглядається аналогічно. $\square$

2. Моменти.

2.1. Моменти дробового порядку. Очевидно, для p-узагальненого розподілу Коші не існує навіть перший момент μ_1'. Але існують і мають досить простий вигляд абсолютні моменти дробового порядку, тобто

$$M_r' = \int\limits_{\mathbb{R}} |x|^r p_{a,b,c,\gamma_1,\gamma_2}(x)dx,$$

де $r \in (0;1)$.

Теорема 2. *Нехай $r \in (0;1)$. Абсолютний момент порядку r p-узагальненого розподілу Коші з параметрами $(1, 0, c, \gamma_1, \gamma_2)$ дорівнює*

$$M_r' = \frac{A^{-1}\pi}{\cos(\pi r/2)} + \frac{cA^{-1}}{2}\left(B\left(\frac{r+1}{2}, \frac{1}{2}+\gamma_1-\frac{r}{2}\right) + B\left(\frac{r+1}{2}, \frac{1}{2}+\gamma_2-\frac{r}{2}\right)\right), \tag{5}$$

де A означається рівністю (2).

Доведення. Маємо:

$$M_r' = 2A^{-1}\int\limits_0^\infty \frac{x^r dx}{1+x^2} + cA^{-1}\left(\int\limits_0^\infty \frac{x^r dx}{(1+x^2)^{1+\gamma_1}} + \int\limits_0^\infty \frac{x^r dx}{(1+x^2)^{1+\gamma_2}}\right) \tag{6}$$

Але для $\gamma \geq 0$

$$\int\limits_0^\infty \frac{x^r}{(1+x^2)^{1+\gamma}}dx = \frac{1}{2}B\left(\frac{r+1}{2}; \frac{1}{2} - \frac{r}{2} + \gamma\right), \tag{7}$$

що випливає з рівності 3.259(3^{11}) у [9]. Зокрема,

$$\int\limits_0^\infty \frac{x^r dx}{1+x^2} = \frac{1}{2}B\left(\frac{r+1}{2}; 1 - \frac{r+1}{2}\right) = \frac{\pi}{2\sin\left(\pi(r+1)/2\right)} = \frac{\pi}{2\cos\left(\pi r/2\right)} \tag{8}$$

(скористались рівністю $B(x, 1-x) = \pi/\sin \pi x$).

Підставляючи вирази з (7) і (8) у (6), отримуємо (5). □

2.2. Узагальнені неповні моменти. Розглянемо також і "узагальнені" неповні моменти.

Означення 2. Подвійно неповним моментом порядку n розподілу F із щільністю $p(x)$ називатимемо число

$$m_n^{r,s} = \int\limits_r^s x^n p(x)dx.$$

Теорема 3. *Подвійно неповний момент* $m_n^{r,s}$ *p-узагальненого розподілу Коші з параметрами* $(1, 0, c, \gamma_1, \gamma_2)$

1) *для* $0 < r < s$ *дорівнює*

$$m_n^{r,s} = \tfrac{1}{A(n+1)}\Big[s^{n+1}\left({}_2F_1\left(1, \tfrac{n+1}{2}; \tfrac{n+3}{2}; -s^2\right) + c\,{}_2F_1\left(1+\gamma_2, \tfrac{n+1}{2}; \tfrac{n+3}{2}; -s^2\right)\right) -$$

$$-r^{n+1}\left({}_2F_1\left(1, \tfrac{n+1}{2}; \tfrac{n+3}{2}; -r^2\right) + c\,{}_2F_1\left(1+\gamma_2, \tfrac{n+1}{2}; \tfrac{n+3}{2}; -r^2\right)\right)\Big]; \tag{9}$$

2) *для* $r < 0 < s$ —

$$m_n^{r,s} = \tfrac{1}{A(n+1)}\Big[(-1)^n|r|^{n+1}\left({}_2F_1\left(1, \tfrac{n+1}{2}; \tfrac{n+3}{2}; -r^2\right) + \right.$$

$$+c\,{}_2F_1\left(1+\gamma_1,\frac{n+1}{2};\frac{n+3}{2};-r^2\right)\right)+$$

$$\left.+s^{n+1}\left({}_2F_1\left(1,\tfrac{n+1}{2};\tfrac{n+3}{2};-s^2\right)+c\,{}_2F_1\left(1+\gamma_2,\tfrac{n+1}{2};\tfrac{n+3}{2};-s^2\right)\right)\right]; \qquad (10)$$

3) *для* $r < s < 0$ —

$$m_n^{r,s}=\tfrac{(-1)^n}{A(n+1)}\left[|r|^{n+1}\left({}_2F_1\left(1,\tfrac{n+1}{2};\tfrac{n+3}{2};-r^2\right)+c\,{}_2F_1\left(1+\gamma_1,\tfrac{n+1}{2};\tfrac{n+3}{2};-r^2\right)\right)+\right.$$

$$\left.-|s|^{n+1}\left({}_2F_1\left(1,\tfrac{n+1}{2};\tfrac{n+3}{2};-s^2\right)+c\,{}_2F_1\left(1+\gamma_1,\tfrac{n+1}{2};\tfrac{n+3}{2};-s^2\right)\right)\right]. \qquad (11)$$

Доведення. Обмежимося доведенням для випадку $0 < r < s$ (в інших випадках доведення аналогічне).

Позначимо через $p(x)$ щільність p-узагальненого розподілу Коші з параметрами $(1,0,c,\gamma_1,\gamma_2)$. Спочатку знайдемо вираз для інтеграла

$$\int\limits_0^t \frac{x^n dx}{(1+x^2)^{1+\gamma}}, \qquad (12)$$

де $\gamma \ge 0$, $t > 0$. Користуючись рівністю 3.194(1) з [9], отримаємо:

$$\int\limits_0^t \frac{x^n dx}{(1+x^2)^{1+\gamma}}=\frac{1}{2}\int\limits_0^{t^2}\frac{y^{(n-1)/2}dy}{(1+y)^{1+\gamma}}=\frac{t^{n+1}}{n+1}\,{}_2F_1\left(1+\gamma,\tfrac{n+1}{2};\tfrac{n+3}{2};-t^2\right). \qquad (13)$$

А тому

$$\int\limits_0^t x^n p(x)dx=\frac{1}{A}\int\limits_0^t\frac{x^n dx}{1+x^2}+\frac{c}{A}\int\limits_0^t\frac{x^n dx}{(1+x^2)^{1+\gamma_2}}=$$

$$=\tfrac{t^{n+1}}{A(n+1)}\left[{}_2F_1\left(1,\tfrac{n+1}{2};\tfrac{n+3}{2};-t^2\right)+c\,{}_2F_1\left(1+\gamma_2,\tfrac{n+1}{2};\tfrac{n+3}{2};-t^2\right)\right]. \qquad (14)$$

І тепер рівність (9) випливає з (14). □

2.3. Робастні моменти 3-го та 4-го порядків. Нагадаємо поняття асиметрії Bowley та ексцеса Moors.

Нехай F — розподіл із неперервною строго зростаючою функцією розподілу $F(x)$,

$$Q(u) = F^{-1}(u), \quad u \in [0;1],$$

— його квантильна функція.

Означення 3. ([6]). Число

$$\frac{Q(3/4) + Q(1/4) - 2Q(1/2)}{Q(3/4) - Q(1/4)}$$

називається асиметрією Bowley розподілу F.

Означення 4. ([13]). Число

$$\frac{(Q(7/8) - Q(5/8)) + (Q(3/8) - Q(1/8))}{Q(6/8) - Q(2/8)}$$

називається ексцесом Moors розподілу F.

У випадку p-узагальненого розподілу Коші для асиметрії Bowley та ексцеса Moors не існує аналітичного виразу. Але за допомогою числового обертання функції розподілу можна наближено знайти відповідні квантилі, а за ними — ці робастні моменти.

Наближені значення асиметрії Bowley та ексцеса Moors для p-узагальненого розподілу Коші з параметрами $(1, 0, c, \gamma_1, \gamma_2)$ для різних значень c, γ_1, γ_2 наведені у табл. 3.Квантилі знаходились за допомогою пакету

`GoFKernel` середовища `R`.

c	γ_1	γ_2	**Асиметрія Bowley**	**Ексцес Moors**
1	5	0,2	0,1264	2,0283
1	5	0,5	0,0731	2,0369
1	10	0,5	0,0606	2,0795
1	5	1	0,0316	2,0927
1	7	7	0	2,3332
1	1	10	-0,0123	2,1278
1	10	10	0	2,3125
2	5	0,2	0,1947	1,9143
2	5	5	0	2,4767
2	10	5	-0,0095	2,5943
2	0,2	10	-0,2144	2,0177
2	7	10	0,0106	2,6429
5	5	0,2	0,2409	1,7002
5	10	0,2	0,2630	1,7154
5	10	5	0,0445	2,4860
5	7	7	0	2,4966
5	10	7	0,0216	2,7507
5	10	10	0	3,0221
10	5	0,2	0,2544	1,6435
10	10	0,2	0,2753	1,6471
10	10	0,5	0,2162	1,5199
10	5	1	0,1354	1,4798
10	10	1	0,1694	1,4821
10	10	10	0	2,0604

Табл. 3: Значення асиметрії Bowley та ексцеса Moors

3. Підгонка до реальних даних.

Покажемо, що p-узагальнений розподіл Коші може використовуватися для моделювання реальних даних.

Розглянемо набір даних вартості акцій фірми PDD (див. [14]) з 15.07.2022 р.

до 08.11.2022 р. Прологарифмуємо цей ряд і знайдемо його прирости, отриманий часовий ряд позначимо PDD. Значущі автокореляції у ряді PDD відсутні.

Порівняємо за допомогою статистики AIC якість підгонки для p-узагальненого розподілу Коші та наступних конкуруючих сімей розподілів: розподілу Коші; асиметричного розподілу Коші (skew-Cauchy distribution), див. [2]; узагальненого логістичного розподілу I-го типу (type I generalized logistic distribution), див. [4]; а також гіперболічного розподілу, див. [5] та [12] (гл. 6). Нами використовується варіант статистики AIC, який означається формулою

$$\mathrm{AIC} = -2(l - p),$$

де l — логарифмічна функція максимальної правдоподібності, p — кількість параметрів розподілу. Результати наведені у табл. 4. (При підгонці цих розподілів були використані наступні пакети середовища `R`: `fitdistrplus`, `sn`, `glogis`, `GeneralizedHyperbolic`.)

Розподіл	**AIC**
p-узагальнений Коші	-254,596
Коші	-241,144
Асиметричний Коші	-239,388
Узагальнений логістичний	-253,699
Гіперболічний	-252,133

Табл. 4: Результати підгонки.

Таким чином, p-узагальнений розподіл Коші по якості підгонки перевершує усіх своїх конкурентів.

Висновки

Нами введено новий імовірнісний розподіл — p-узагальнений розподіл Коші (який являє собою модифікацію класичного розподілу Коші) і розглянуто його властивості. Зокрема, знайдено моменти дробового порядку та певні узагальнені неповні моменти. Досліджено робастні стандартизовані моменти 3-го та 4-го порядку. Продемонстровано, що p-узагальнений розподіл Коші може використовуватися для моделювання приростів логарифмів часових рядів цін акцій.

Список літератури

1. **Alzaatreh A.** An alternative to the Cauchy distribution / A. Alzaatreh // MethodsX. — 2019. — Vol. 6. — P. 938–952.
2. **Azzalini A.** The Skew-normal and Related Families / A. Azzalini. — N. Y.: Cambridge University Press, 2013. — 262 p.
3. **Bahrami W.** A two-parameter generalized skew-Cauchy distribution / W. Bahrami, H. Rangin, K. Rangin // Journal of Statistical Research of Iran. — 2010. — Vol. 7(1). — P. 61–72.
4. **Balakrishnan N.** Handbook of the Logistic Distribution / N. Balakrishnan. — N. Y.: Marcel Dekker, Inc., 1992. — 586 p.
5. **Barndorff-Nielsen O.** Exponentially decreasing distributions for the logarithm of particle size / O. Barndorff-Nielsen // Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences. — 1977. — Vol. 353, № 1674. — P. 401–419.
6. **Bowley A. L.** Elements of Statistics / A. L. Bowley. — N. Y.: Scribner's, 1920. — 454 p.
7. **Chattamvelli R.** Continuous distributions in engineering and the applied sciences. (Part I) / R. Chattamvelli, R. Shanmugam. — Cham: Springer International Publishing, 2021. — 152 p.
8. **Dorić D.** New generalizations of Cauchy distribution / D. Dorić // Communications in Statistics. Theory and Methods. — 2011. — Vol. 40(21). — P. 3764–3776.
9. **Gradshteyn I. S.** Table of Integrals, Series, and Products / I. S. Gradshteyn, I. M. Ryzhik. — Amsterdam: Elsevier, 2007. — 1171 p.
10. **Huang W. J.** Generalized skew-Cauchy distribution / W. J. Huang, Y. H. Chen // Statistics & Probability Letters. — 2007. — Vol. 77. — P. 1137–1147.
11. **Jayakumar K.** On a new generalization of Cauchy distribution / K. Jayakumar, K. Fasna // Asian Journal of Statistical Sciences. — 2022. — Vol. 2(1). — P. 61–81.
12. **Handbook of Heavy Tailed Distributions in Finance, Book 1** / editor: S.T Rachev. — Amsterdam: Elsevier, 2003. — 704 p.
13. **Moors J. J. A.** A quantile alternative for kurtosis / J. J. A. Moors // Journal of the Royal Statistical Society: Series D (The Statistician). — 1988. — Vol. 37(1). — P. 25–32.
14. **PDD.** PDD Holdings Inc. (PDD). — Regime of access: https://finance.yahoo.com/quote/PDD,
15. **Arnold B. C.** Univariate and Bivariate Models Related to the Generalized Epsilon–Skew–Cauchy Distribution / B. C. Arnold, H. W. Gómez, H. Varela, I. Vidal // Symmetry. — 2019. — Vol. 11(6). — P. 794.

Turchyn I.

A NEW GENERALIZATION OF CAUCHY DISTRIBUTION

Summary

A new heavy-tailed distribution is considered in this paper: the p-generalized Cauchy distribution. This new 5-parameter distribution family is substantially more flexible than than the classical Cauchy distribution, it contains, in particular, asymmetric distributions. Various numerical characteristics of the new distribution are considered, among them are fractional order moments and "twice incomplete" moments. There were obtained also (using numerical methods) values of quantile-based standardized moments of higher orders: Bowley's skewness and Moors' kurtosis. Suitability of the p-generalized Cauchy distribution for modeling real data was confirmed by fitting this distribution to a series of log returns of stock prices. The p-generalized Cauchy distribution had the smaller value of the AIC statistic than the Cauchy distribution, the skew-Cauchy distribution, the generalized logistic distribution and the hyperbolic distribution.

Key words: Cauchy distribution, incomplete moments, Bowley's skewness, Moors' kurtosis.

REFERENCES

1. Alzaatreh, A. (2019). An alternative to the Cauchy distribution. *MethodsX*, Vol. 6, P. 938–952.
2. Azzalini, A. (2013). *The Skew-Normal and Related Families.* N. Y.: Cambridge University Press, 262 p.
3. Bahrami, W., Rangin, H. and Rangin, K. (2010). A two-parameter generalized skew-Cauchy distribution. *Journal of Statistical Research of Iran,* Vol. 7(1), P. 61–72.
4. Balakrishnan, N. (1992). *Handbook of the Logistic Distribution.* N. Y.: Marcel Dekker, Inc., 586 p.
5. Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences,* Vol. 353, №1674, P. 401–419.
6. Bowley, A. L. (1920). *Elements of Statistics.* N. Y.: Scribner's, 454 p.
7. Chattamvelli, R., Shanmugam, R. (2021) *Continuous Distributions in Engineering and the Applied Sciences, Part I.* Cham: Springer International Publishing, 152 p.

8. Dorić, D. (2011). New generalizations of Cauchy distribution. *Communications in Statistics — Theory and Methods,* Vol. 40(21), P. 3764–3776.

9. Gradshteyn, I. S., Ryzhik, I. M. (2007). *Table of Integrals, Series, and Products.* Amsterdam: Elsevier, 1171 p.

10. Huang, W. J., Chen, Y. H. (2007). Generalized skew-Cauchy distribution. *Statistics & Probability Letters,* Vol. 77, P. 1137–1147.

11. Jayakumar, K., Fasna, K. (2022). On a new generalization of Cauchy distribution. *Asian Journal of Statistical Sciences,* Vol. 2(1), P. 61–81.

12. Rachev, S. T. (ed.) (2003) *Handbook of Heavy Tailed Distributions in Finance, Book 1.* Amsterdam: Elsevier, 2003, 704 p.

13. Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician),* Vol. 37(1), P. 25–32.

14. PDD Holdings Inc. (PDD). https://finance.yahoo.com/quote/PDD

15. Arnold, B. C., Gómez, H. W., Varela, H. and Vidal, I. (2019). Univariate and Bivariate Models Related to the Generalized Epsilon–Skew–Cauchy Distribution. *Symmetry,* Vol. 11(6), 794 p.