# МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ЕСТЕСТВЕННЫХ, ТЕХНИЧЕСКИХ И СОЦИАЛЬНЫХ НАУКАХ

**УДК 513.015.2**

**В.М. Галкин, Л.Н. Ерофеева, С.В. Лещева**

## ОЦЕНКИ ПАРАМЕТРА РАСПРЕДЕЛЕНИЯ КОШИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Известно, что вероятностное распределение Коши не имеет ни математического ожидания, ни дисперсии. Поэтому для оценки его параметров приходится модифицировать излагаемые в учебниках методы. Так, в [3] использовался метод моментов с дробными степенями. В данной статье приводится еще ряд оценок, а также даются результаты численного эксперимента.

*Ключевые слова:* вероятностное распределение, состоятельные и эффективные оценки, неравенство Рао-Крамера.

**1.** Под однопараметрическим распределением Коши понимается вероятностное распределение с плотностью

$$f(x)=\frac{1}{\pi}\cdot\frac{a}{a^2+x^2},\ \ (-\infty<x<+\infty), \tag{1}$$

зависящее от параметра $a$. Это распределение довольно экзотично по сравнению с известными часто встречающимися распределениями, например нормальным. Тем не менее оно имеет интересные свойства и находит приложение в экономике. Отметим некоторые из этих свойств ([1], [2]):

1) распределение Коши безгранично делимо;

2) отношение независимых нормально распределенных случайных величин распределено по Коши;

3) если случайная величина $X$ распределена по Коши, то $Y=\frac{2}{\pi}\mathrm{arctg}\frac{X}{a}$ и $|Y|$ распределены равномерно на интервалах $[-1,1]$ и $[0,1]$ соответственно.

В [3] отмечено, что второе свойство используется при моделировании статистического поведения биржевого курса валют и, следовательно, нахождение оценок параметра $a$ по статистическим данным $x_1,\ x_2,...,x_n$ реализации $X$ имеет и прикладное значение.

При исследовании этого вопроса мы сразу же встречаемся с той трудностью, что $X$ не имеет ни математического ожидания, ни дисперсии, поскольку соответствующие интегралы для $M(X)$ и $D(X)$ расходятся. Это обстоятельство заставляет, по крайней мере, на первых порах исключить из рассмотрения классические оценки типа среднего арифметического $\bar{x}=\frac{1}{n}(x_1+x_2+...+x_n)$ и среднего квадратичного $\overline{x^2}=\frac{1}{n}\left(x_1^2+x_2^2+...+x_n^2\right)$.

© Галкин В.М., Ерофеева Л.Н., Лещева С.В., 2014.

**2.** Сделаем несколько замечаний о стандартных методах нахождения оценок – методе максимального правдоподобия и методе моментов. В первом из них значение параметра ищется из требования максимальности выражения

$$L = f(x_1) f(x_2) ... f(x_n) \tag{2}$$

для экспериментальных данных $x_1, x_2, ..., x_n$.

Отсюда получаем $\frac{\partial L}{\partial a} = 0$ и, используя (1), сведем последнее равенство к уравнению

$$\sum_{i=1}^{n} \frac{a^2}{a^2 + x_i^2} = \frac{n}{2}. \tag{3}$$

Каждое слагаемое в левой есть возрастающая функция $a > 0$, а потому можно заключить, что уравнение (3) имеет единственное решение относительно $a$, которое принимается за оценку $a$. Однако явно решить уравнение (3) не удается, а потому вопрос о точности оценки в этом методе остается открытым.

В методе моментов вычисляется величина

$$M_k = \frac{1}{n} \sum_{i=1}^{n} |x_i|^k, \tag{4}$$

при заданном $k$. Математическое ожидание этой величины содержит нужный параметр и, приравнивая $M_k$ ее математическому ожиданию, находим оценку параметра. Однако для распределения Коши моменты целых порядков не существуют (точнее они равны $\infty$).
М.Л. Шинкеев ([3]) обошел эту трудность, рассматривая (4) при $0 < k < 1$, где математическое ожидание существует и равно $\frac{a^k}{\cos \frac{\pi k}{2}}$. Для $a$ оценка получается в виде

$$a \cong a_k^* = \left( M_k \cos \frac{\pi}{2} \right)^{1/k}. \tag{5}$$

К изложенному можно добавить, что $M_k$ можно рассматривать и для отрицательных $k$ из интервала $(-1, 0)$. Следующее соображение показывает, что выигрыша при этом не происходит. Дело в том, что распределение Коши инвариантно при преобразовании $X \to \frac{1}{X}$, $a \to \frac{1}{a}$ и (5) дает

$$\frac{1}{a} \cong a_{-k}^* = \left( M_{-k} \cos \frac{\pi k}{2} \right)^{1/k}.$$

Чтобы исключить множитель $\cos \frac{\pi k}{2}$, можно предложить оценку

$$a \cong \sqrt{\frac{a_k^*}{a_{-k}^*}} = \left( \frac{M_k}{M_{-k}} \right)^{\frac{1}{2k}}.$$

**3.** Иные типы оценок можно получать трансформируя исходную случайную величину $X$ в другую, имеющую более удобный для исследования закон распределения. Проиллюстрируем это двумя примерами. Как отмечено ранее, величина $Y=\frac{2}{\pi}\operatorname{arctg}\frac{|X|}{a}$ имеет равномерное распределение на интервале $[0,1]$. Для математического ожидания $Y$, равного $\frac{1}{2}$, имеется оценка ([4]) $\frac{1}{2}(y_{\min}+y_{\max})$, где $y_i=\frac{2}{\pi}\operatorname{arctg}\frac{x_i}{a}$.

Следовательно, имеем приближенное равенство

$$\frac{1}{2}\cong\frac{1}{\pi}\left(\operatorname{arctg}\left|\frac{x_{\min}}{a_i}\right|+\operatorname{arctg}\left|\frac{x_{\max}}{a}\right|\right), \tag{6}$$

которое легко преобразуется к виду

$$a\cong a^*=\sqrt{|x_{\min}|\cdot|x_{\max}|}\,. \tag{7}$$

Оценка (6) относится к типу так называемых «сверхоценок», у которых дисперсия на порядок меньше той, которая следует из неравенства Рао-Крамера ([1,2,6]). Однако утверждать, что таким же свойством обладает и оценка (7), нельзя.

Второй пример трансформации $X$ рассмотрим более подробно. От $X$ перейдем к величине $Z=\ln|X|$.

Функция распределения $F_Z(x)$ вычисляется стандартным образом:

$$F_Z(x)=p(Z<x)=p\left(|X|<e^{\alpha}\right)=\int_{-e^x}^{e^x}\frac{2}{\pi}\cdot\frac{a}{a^2+x^2}dx\,.$$

После вычисления интеграла и дифференцирования $F_Z(x)$ получается выражение для плотности распределения величины $Z$:

$$f_Z(x)=\frac{2}{\pi}\cdot\frac{ae^x}{a^2+e^{2x}},\ (-\infty<x<+\infty). \tag{8}$$

У $Z$ существуют моменты любого неотрицательного порядка.

Математическое ожидание $M(Z)=\int_{-\infty}^{+\infty}xf_Z(a)dx=\frac{2}{\pi}\int_{-\infty}^{+\infty}x\frac{ae^x}{a^2+e^{2x}}$, если положить $a=e^{\alpha}$ и сделать замену $x\to x-\alpha$ в интеграле, получается равным

$$M(Z)=\frac{2}{\pi}\int_{-\infty}^{+\infty}(x+\alpha)\frac{dx}{e^x+e^{-x}}=\alpha\,. \tag{9}$$

Для центрального момента порядка $k$ получаем интегральное представление

$$M_k(Z)=\frac{2}{\pi}\int_{-\infty}^{+\infty}x^k\,\frac{dx}{e^x+e^{-x}}=\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{x^k dx}{\operatorname{ch}x}.$$

Очевидно, что моменты целого нечетного порядка $k>0$ обращаются в ноль. Явные выражения для моментов четного порядка получаются, если рассмотреть известный интеграл([5])

$$\int_0^{\infty}\frac{\cos px}{\operatorname{ch}x}dx=\frac{\pi}{2}\cdot\frac{1}{ch\frac{p\pi}{2}}\,.$$

Используя степенные разложения

$$\cos t=1-\frac{1}{2!}t^2+\frac{1}{4!}t^4-\frac{1}{6!}t^6+\cdots,$$

$$\frac{1}{\operatorname{ch} t}=1-\frac{E_2}{2!}t^2+\frac{E_4}{4!}t^4-\frac{E_6}{6!}t^6+\cdots,$$

где $E_2=1,\ E_4=5,\ E_6=61-$ так называемые числа Эйлера, легко найти, что

$$M_{2k}(Z)=\left(\frac{\pi}{2}\right)^{2k}E_{2k}.$$

В частности, дисперсия $D(Z)=\frac{\pi^2}{4}$.

В качестве оценки $\alpha=\ln a$ возьмем среднее арифметическое значение $z_i=\ln|x_i|$ величины $Z$:

$$\alpha\cong\alpha^*=\frac{1}{n}\sum_i \ln|x_i|. \tag{10}$$

Эта оценка состоятельна и несмещена. Чтобы оценить ее эффективность, обратимся к неравенству Рао-Крамера, для несмещенной оценки $\theta$:

$$D(\theta)\geq\frac{1}{\left\|\frac{\partial\ln L}{\partial\alpha}\right\|^2},$$

где $L$ дается формулой (2), в которой $f(x)$ надо заменить на $f_Z(x)$ из (8). Квадрат нормы $\|u\|^2$ дается кратным интегралом

$$\|u\|^2=\int_{-\infty}^{\infty}\cdots\int u^2 L dx_1...dx_n.$$

Несколько громоздкие, но нетрудные вычисления дают $\left\|\frac{\partial\ln L}{\partial\alpha}\right\|^2=\frac{n}{2}$ и значение $\frac{2}{n}$ для нижней границы дисперсии. Но $D(\alpha^*)=\frac{1}{n}D(Z)=\frac{2,4...}{n}$ мало отличается от $\frac{2}{n}$. Возможно оценка $\alpha^*$ является эффективной, так как нижняя граница $\frac{2}{n}$ не может быть достигнута для распределения (8). Для того чтобы перейти к оценке параметра $a$, надо взять экспоненту от $\alpha^*$, т.е. среднее геометрическое $\sqrt[n]{|x_1|\cdot|x_2|\cdot...\cdot|x_n|}$.

**4.** В статистическом пакете программ из Maple есть программа, генерирующая значения случайной (квазислучайной) величины, распределенной по Коши. Был проведен численный эксперимент и подсчитывались значения рассмотренных ранее оценок.

Моделировалось распределение с параметром $a=2$ для значений $n=10,100,1000$. Полученные результаты даны далее (табл. 1, 2, 3). В каждой из них приводятся результаты по 10 реализаций *X*. Через $a_1,\ a_2,\ a_3,\ a_4$ обозначены оценки:

$a_1=\sqrt{|x_{\min}|\cdot|x_{\max}|}$;

$a_2=\left(\frac{M_k}{M_{-k}}\right)^{\frac{1}{2k}},\ \ k=\frac{1}{4}$, где $M_k$ – дается формулой (4);

$a_3=\sqrt[n]{|x_1|\cdot|x_2|\cdot...\cdot|x_n|}$;

$a_4-$ корень уравнения (3).

**Таблица 1**

$(a = 2,\ n = 10)$

| № п/п | $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|---|
| 1 | 3,854 | 2,050 | 1,925 | 1,071 |
| 2 | 1,789 | 1,143 | 1,125 | 1,057 |
| 3 | 2,179 | 3,091 | 3,111 | 3,197 |
| 4 | 1,437 | 1,780 | 1,794 | 1,982 |
| 5 | 9,530 | 3,062 | 2,927 | 2,155 |
| 6 | 1,126 | 2,421 | 2,927 | 2,822 |
| 7 | 1,496 | 1,866 | 1,870 | 1,945 |
| 8 | 0,691 | 1,345 | 1,408 | 1,951 |
| 9 | 9,374 | 5,959 | 5,883 | 4,759 |
| 10 | 5,212 | 3,632 | 3,564 | 2,649 |

**Таблица 2**

$(a = 2,\ n = 100)$

| № п/п | $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|---|
| 1 | 2,322 | 2,459 | 2,454 | 2,429 |
| 2 | 4,738 | 2,554 | 2,504 | 2,334 |
| 3 | 2,928 | 1,949 | 1,958 | 1,926 |
| 4 | 0,691 | 1,719 | 1,752 | 1,804 |
| 5 | 1,106 | 1,916 | 1,899 | 1,761 |
| 6 | 8,884 | 2,054 | 1,981 | 1,819 |
| 7 | 1,499 | 1,210 | 1,224 | 1,424 |
| 8 | 1,036 | 1,971 | 1,994 | 2,094 |
| 9 | 11,10 | 2,722 | 2,539 | 2,350 |
| 10 | 2,358 | 2,262 | 2,248 | 2,097 |

**Таблица 3**

$(a = 2,\ n = 1000)$

| № п/п | $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|---|
| 1 | 2,6740 | 2,9631 | 1,9615 | 1,9443 |
| 2 | 4,3306 | 2,0016 | 1,9872 | 1,9750 |
| 3 | 1,7889 | 2,0634 | 2,0616 | 2,0259 |
| 4 | 1,2766 | 2,0209 | 2,0228 | 2,0211 |
| 5 | 1,6135 | 1,9583 | 1,9545 | 2,0212 |
| 6 | 1,6135 | 1,9583 | 1,9545 | 1,9204 |
| 7 | 3,5532 | 2,1465 | 2,1474 | 2,1560 |
| 8 | 0,8521 | 1,8769 | 1,8946 | 1,9273 |
| 9 | 4,1821 | 2,2677 | 2,2433 | 2,2129 |
| 10 | 2,5099 | 1,9825 | 1,9993 | 2,0494 |

Как видно из таблиц, оценка $a_1$ очень грубая. Небольшую точность, но лучшую, чем у $a_1$ дают и другие оценки при $n = 10$. Но оценки $a_2$, $a_3$, $a_4$ становятся весьма удовлетворительными при $n = 100$ и $n = 1000$. Практически можно рекомендовать работу с экспериментальными данными в объеме нескольких десятков.

### Библиографический список

1. **Крамер, Г.** Математические методы статистики / Г. Крамер. – М.: Мир, 1975.
2. **Ван дер Варден, Б.Л.** Математическая статистика / Б.Л. Ван дер Варден. – М.: Иностр. лит., 1960.
3. **Шинкеев, М.Л.** Оценка параметров распределения Коши / М.Л. Шинкеев; Национальный исследовательский Томский политехнический университет. – Томск.
4. Математическая энциклопедия // Оценка статистическая. – М.: Советская энциклопедия, 1984. Т. 4.
5. **Рыжик, И.М.** Таблицы интегралов, сумм, рядов и произведений / И.М. Рыжик, И.С. Градштейн. – М.–Л.: Госиздат, 1951.
6. **Аниковский, В.В.** Математическая статистика / В.В. Аниковский, Л.Н. Ерофеева. – Н. Новгород, 2013.

*Дата поступления*
*в редакцию: 22.04.2014*

**V.M. Galkin, L.N. Erofeeva, S.V. Leshcheva**

## THE PARAMETER ESTIMATES OF CAUCHY DISTRIBUTION

**Purpose:** The design is the methods of the estimates finding, of the probability distributions which have not the moments of integer orders to give

**Design/methodology/approach:** The transformations of the distributions to ones that having the arbitrary order moments are used.

**Findings:**The new estimates are finded.

**Research limitation/ implications:** The results of this paper will beuseful for the students and aspirants.

**Originality/value:** The paper is useful for not enough papers of this sort in literature are.

*Key words:* probability distribution, consistent and efficient estimates, Rao-Kramer unequality.