УДК 513.015.2

Е.Д. Галкина[2], С.В. Лещева[1], Н.С. Лукичев[1], В.Е. Рыков[1]

НЕКОТОРЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ КОШИ

Нижегородский государственный технический университет им. Р.Е. Алексеева[1]
ООО «Телека»[2]

Для двухпараметрического вероятностного распределения Коши получены три типа оценок параметров. Один из них использует понятие медианы, а другие – сведение к подходящим функциям от исходной случайной величины. В частности, применяются комплекснозначные функции.

Ключевые слова: оценка параметра, вероятностное распределение, медиана, среднеквадратическое отклонение.

Распределение Коши – это вероятностное распределение с плотностью

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

где a и b $(b>0)$ – параметры. Его обычно относят к числу классических (наряду с нормальным, пуассоновским и др.) распределений. Ввиду его так называемой безграничной делимости [1]:

$$X=X_1+X_2+...+X_n, \tag{2}$$

где X распределение по Коши с параметрами a и b, а независимые случайные величины в правой части распределены также по Коши, но с параметрами $\frac{a}{n}$ и $\frac{b}{n}$. В частности, среднее $\frac{1}{n}(X_1+X_2+...+X_n)$ имеет те же параметры, что и каждое из X_i.

Имеются связи распределения Коши с другими известными распределениями. Так нетрудно показать, что $Y=\operatorname{arctg}\frac{x-a}{b}$ распределено равномерно на $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$. Отношение $Z=\frac{X_1}{X_2}$ независимых нормально распределенных величин распределено по Коши с параметрами 0 и $\sigma=\frac{\sigma_2}{\sigma_1}$, если $M(X_1)=M(X_2)=0$, $D(X_i)=\sigma_i^2$. Это свойство, как отмечено в [2], используется в экономических моделях биржевой практики.

Нахождение оценок параметров a и b в (1), исходя из множества независимых измерений случайной величины X,

$$x_1,x_2,...,x_n \tag{3}$$

представляет некоторые трудности. Это объясняется тем, что стандартные методы получения оценок – метод максимального правдоподобия и метод моментов почти не работают.

Первый приводит к не решаемому в явной форме алгебраическому уравнению, а второй без модификации неприменим из-за несуществования моментов целого порядка.

Поучительно рассмотреть оценку момента $M_1(X)=\int\limits_{-\infty}^{+\infty}xf(x)dx$. Интеграл расходится, но существует в смысле главного значения. Если оценку a взять в виде $a\cong a^*=\frac{1}{n}(x_1+x_2+...+x_n)$, то можно обнаружить, что a^* имеет то же самое распределение, что и X. Таким образом, количество измерений n в данном случае не влияет на точность оценки.

В свое время А.Н. Колмогоров [3] обратил внимание на полезность использования в статистике такой характеристики случайной величины, как медиана. По определению медиана $m(X)$ случайной величины X есть такое число m, что

$$p(X < m) = p(X \geq m) = \frac{1}{2}. \tag{4}$$

Предложение 1. Пусть X распределено по Коши с параметрами a и b, $Y = |X - a|$. Тогда $m(X) = a,\ m(Y) = b$.

Проверка здесь тривиальна. Надо лишь заметить, что Y имеет плотность $f_Y(x)$, равную $\frac{2}{\pi} \cdot \frac{b}{x^2 + b^2}$ при $x > 0$ и 0 при $x \leq 0$.

В [2] для оценки параметра a предложено использовать медиану, но подробности не приведены. Далее приводятся соответствующие результаты.

Если экспериментальные данные значений какой-либо случайной величины упорядочить по возрастанию: $x_1 < x_2 < ... < x_i < x_{i+1} < ... < x_k$, то выборочная медиана m^* принимается равной x_k при $n = 2k - 1$ и $\frac{1}{2}(x_k + x_{k+1})$ при $n = 2k$. При нечетном n закон распределения выборочной медианы легко описать.

Предложение 2. Для случайной величины X с плотностью распределения $f(x)$ выборочная медиана $m^* = m(x_1, x_2, ..., xn)$ имеет плотность $nC_{n-1}^{k-1} F^{k-1}(x)(1 - F(x))^{k-1} f(x)$. Здесь $n = 2k - 1$, а $F(x)-$ функция распределения X.

Доказательство можно найти в книге Ван-дер Вардена [1], где рассматривается случайная величина с произвольным законом распределения. Там же устанавливается асимптотическая нормальность распределения выборочной медианы при $k \to \infty$. Если соответствующие формулы использовать для распределения Коши (1), то получим:

Предложение 3. При $n \to \infty$ выборочная медиана распределения Коши распределена асимптотически нормально с математическим ожиданием a и среднеквадратическим отклонением $\sigma = \frac{\pi b}{2\sqrt{n}}$.

Интуитивно ясно, что при $n = 2k$ четном ситуация аналогична, хотя в предложении 2 формулы значительно усложняются.

По существу, тот же результат получается и для величины $Y = |X - a|$ из предложения 1, а именно выборочная медиана здесь распределена асимптотически нормально с математическим ожиданием b и среднеквадратическим отклонением $\sigma = \frac{\pi b}{2\sqrt{n}}$.

Конечно, здесь предполагается, что параметры a и b известны. Можно ожидать, что замена их на значения выборочных медиан вносит незначительные изменения.

Итак, предлагаются следующие оценки параметров a и b в (1) по экспериментальным данным (3):

$$a \cong a^* = \text{median}(x_1, x_2, ..., x_n),$$
$$b \cong b^* = \text{median}(|x_1 - a^*|, |x_2 - a^*|, ..., |x_n - a^*|).$$

Дисперсии оценок при больших n приблизительно равны $\frac{\pi b^*}{2\sqrt{n}}$.

Другой тип оценок может быть получен методом, который назовем *методом сжатия* *экспериментальных данных*. Он заключается в переходе от случайной величины X к случайной величине $Y = g(X)$, имеющей моменты нужных порядков. По существу, этот метод использовался в [2 и 4].

В качестве примера возьмем $Y = e^{i\lambda X}$, где $i = \sqrt{-1}$, $\lambda -$ действительный коэффициент. Математическое ожидание Y есть не что иное, как характеристическая функция X. Она легко подсчитывается с помощью вычетов и результат таков:

$$M\left(e^{i\lambda X}\right) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{i\lambda x} \frac{b}{(x-a)^2 + b^2}\,dx = e^{(ia-b)\lambda}. \tag{5}$$

Аппроксимируя левую часть в (5) средним арифметическим $\frac{1}{n}\sum_{k=1}^{n} = e^{i\lambda x_k}$, получаем следующие оценки:

$$\begin{aligned} a &\cong \frac{1}{\lambda}\arg\left(\frac{1}{n}\sum_{k=1}^{n} e^{i\lambda x_k}\right), \\ b &\cong -\frac{1}{2}\ln\left|\frac{1}{n}\sum_{k=1}^{n} e^{i\lambda x_k}\right|. \end{aligned} \tag{6}$$

Еще примеры доставляют функции $Z = \frac{1}{(X-a_1)^2 + b_1^2}$ и $Z_1 = \frac{X - a_1}{(X-a_1)^2 + b_1^2}$, где a_1 и b_1 заданы. Математические ожидания здесь равны:

$$\begin{aligned} M(Z) &= \frac{b_1 + b}{b_1}\cdot\frac{1}{(a-a_1)^2 + (b_1+b)^2}, \\ M(Z_1) &= \frac{a - a_1}{b_1}\cdot\frac{1}{(a-a_1)^2 + (b_1+b)^2}. \end{aligned} \tag{7}$$

Вновь аппроксимации левых частей в (7) средними арифметическими позволяют выразить a и b через экспериментальные данные. Мы не приводим выражение для дисперсий (или среднеквадратических отклонений для оценок из (6) и (7)). Отметим только, что их порядки по n равны $\frac{1}{\sqrt{n}}$.

В заключение приведем данные по численным экспериментам.

Таблица 1

a_1^*	2,220	2,642	2,159	1,225	1,946
	2,297	2,410	1,991	4,064	3,029
	2,970	0,619	1,921	2,744	2,127
	1,676	2,832	3,565	1,726	1,275
a_2^*	1,112	2,714	3,047	0,898	1,735
	2,210	2,715	1,579	4,326	3,254
	2,139	1,454	1,829	2,455	2,296
	1,024	2,168	2,706	2,413	1,259
a_3^*	2,089	2,2462	1,971	1,338	1,770
	2,728	2,472	1,895	3,884	3,318
	2,496	0,982	2,028	2,685	2,145
	1,664	2,185	2,835	2,149	1,439

Моделировалось распределение Коши с параметрами $a = 2, b = 5$. Число n бралось равным 100, 200 и 500. Для каждого из этих n значений брались по 20 реализаций модельного процесса. В прилагаемых табл. 1 и 2 данные соответствуют $n = 100$. В табл. 1 помещены значения оценок параметра a, а в табл. 2 – параметра b. При этом a_1^*, b_1^* – оценки, полученные с использованием медианы, a_2^*, b_2^* – оценки из (6), а a_3^*, b_3^* – оценки из (7). В расчетах по формулам (6) значение $\lambda = 0{,}1$. В расчетах по формулам (7) надо было сначала задать значения a_1 и b_1. В качестве таковых были взяты a_1^* и b_1^*.

Таблица 2

b_1^*	4,928	5,525	5,171	5,473	4,506
	4,162	4,163	4,130	3,856	4,599
	6,382	5,800	5,057	3,739	4,717
	3,896	5,417	4,700	4,775	4,926
b_2^*	5,148	5,656	5,394	5,034	5,166
	5,256	4,567	5,548	4,251	4,286
	5,023	5,566	6,827	3,168	5,477
	4,632	6,068	5,107	4,541	4,448
b_3^*	5,067	5,478	5,272	5,351	4,463
	4,325	4,159	4,131	3,980	4,699
	6,253	5,811	5,119	3,834	4,825
	3,940	5,428	4,649	4,735	4,847

Библиографический список

1. **Ван дер Варден**. Математическая статистика / Ван дер Варден. – М., 1960.
2. **Шинкеев, М.Л.** Оценка параметров распределения Коши / М.Л. Шинкеев // Национальный исследовательский Томский политехнический университет. – Томск.
3. **Колмогоров, А.Н.** Метод медианы в теории ошибок // Теория вероятностей и математическая статистика. С. 11–112.
4. **Галкин, В.М.** Оценки параметров распределения Коши / В.М. Галкин, Л.Н. Ерофеева, С.В. Лещева // Труды НГТУ им Р.Е. Алексеева. – Н. Новгород, 2014. №2. С. 314–319.

Дата поступления
в редакцию 18.06.2015

E. D. Galkina[2], S.V. Leshcheva[1], N.S. Lukichev[1], V.E Rykov[1]

SOME ESTIMATES OF CAUCHY DISTRIBUTION PARAMETERS

Nizhny Novgorod state technical university n.a. R.E. Alexeev[1],
Symphony «Teleca»[2]

Purpose: New estimates of Cauchy distribution parameters are to build.
Design/methodology/approach: The median and random variable transformation are used for the estimate construction.
Findings: Three new types of estimates are found.
Research limitation/ implications: There methods can be used for the other non standard distribution.
Originality/value: The results of this paper are new.

Key words: parameter estimate, probability distribution, median, standard deviation.