

## ОБРАБОТКА ИНФОРМАЦИИ

УДК 519.2

И.А. Багрова

### О ДАТЧИКЕ СЛУЧАЙНЫХ ВЕЛИЧИН, ИМЕЮЩИХ УСТОЙЧИВОЕ РАСПРЕДЕЛЕНИЕ С ХАРАКТЕРИСТИЧЕСКИМ ПОКАЗАТЕЛЕМ БОЛЬШЕ ЕДИНИЦЫ

Описывается работа датчика, генерирующего устойчивые распределения. В основу моделирования положена обобщенная ЦПТ. Предложены формулы для определения параметров датчика.

**Ключевые слова:** *моделирование случайной величины, устойчивые, распределения, распределение Парето.*

При изучении и построении моделей различных явлений часто используется нормальное распределение, зависящее от двух параметров, описывающих его. Но нормальное распределение, в свою очередь, принадлежит к более широкому классу устойчивых распределений, обладающих набором 4 параметров, благодаря чему при их применении возможна более тонкая настройка модели под реальные данные. Кроме того, использование нормального распределения не приводит к удовлетворительным результатам для описания явлений, имеющих импульсный характер, когда вероятность появления экстремального значения отлично от нуля. Поэтому устойчивые распределения, главной особенностью которых являются «тяжелые хвосты», нашли широкое применение во многих областях [1–3]. Например, при описании и получении характеристик процессов в теории лазерного охлаждения атомов [4] и радиотехнике [5]. В [6] устойчивые распределения использовались для получения оптимального портфеля акций на основе методологии VaR. В работе [7], применяя устойчивые распределения, авторы предлагают алгоритм оценивания параметров регрессионных уравнений, обеспечивающий максимально правдоподобное оценивание даже в ситуациях, когда распределение случайных ошибок имеет большую дисперсию.

Таким образом, имеющаяся потребность построения моделей с использованием устойчивых распределений делает актуальной задачу разработки датчиков для моделирования этих законов. Существует несколько подходов [8, 9] к моделированию одномерных устойчивых распределений. Наиболее часто применяется алгоритм, основанный на интегральном представлении Золотарева [10, 11]. Эти методы нельзя использовать для построения датчиков в многомерном случае, поскольку потребуются применение теоремы Коши в  $Z^d$ , что является практически невыполнимой задачей. Указанная проблема не возникает, если разработка датчика основана на обобщенной ЦПТ [5]:

$$S_n = \frac{X_1 + \dots + X_n}{bn^{1/\alpha}} \xrightarrow{d} Y \text{ при } n \rightarrow \infty, \alpha \in (1, 2), \quad (1)$$

где  $X_j$  – центрированные независимые одинаково распределенные случайные величины со значениями из  $d$ -мерного евклидова пространства  $R^d$ , принадлежащие области притяжения устойчивых законов, а нормирующий множитель  $b$  зависит от формы параметризации устойчивой случайной величины  $Y$ .

Первым шагом в применении (1) является разработка датчика для одномерного случая. Таким образом, стоит задача получить такие значения параметра  $b$ , чтобы распределение суммы  $S_n$  сходилось к распределению устойчивой случайной величины  $Y$ . Эту задачу можно решить через рассмотрение соответствующих им характеристических функций. Характеристическая функция устойчивого распределения в форме (А) представляется следующим образом [3]:

$$f_Y(t) = \exp\left\{-(\sigma_A |t|)^\alpha \left(1 - i\beta_A \operatorname{tg} \frac{\pi\alpha}{2} \operatorname{sign}(t)\right)\right\}, \quad (2)$$

где  $Y$  – устойчивая случайная величина.

Следовательно, необходимо получить характеристическую функцию суммы  $S_n$ .

### 1. Характеристическая функция суммы $S_n$

Для моделирования двухстороннего устойчивого распределения использовалась смесь распределений Парето, которые являются наиболее простыми случайными величинами из области притяжения  $Y$ :

$$X_j = pX_j^+ + qX_j^-, \quad p + q = 1, \quad (3)$$

где

$$X_j^+ = \tilde{X}_j^+ - M\tilde{X}^+, \quad X_j^- = \tilde{X}_j^- - M\tilde{X}^-,$$

а  $\tilde{X}_j^+$  и  $\tilde{X}_j^-$  имеют следующие функции распределения:

$$F_{\tilde{X}^+}(x) = \begin{cases} 1 - \frac{r^\alpha}{x^\alpha}, & x \geq r, \\ 0, & x < r, \end{cases} \quad \text{и} \quad F_{\tilde{X}^-}(x) = \begin{cases} 1 - \frac{l^\alpha}{|x|^\alpha}, & x \leq -l, \\ 0, & x > -l, \end{cases}$$

где  $r$  и  $l$  – граничные значения, являющиеся параметрами масштаба.

Математические ожидания вычисляются по формулам

$$M\tilde{X}^+ = \frac{\alpha r}{\alpha - 1}, \quad M\tilde{X}^- = -\frac{\alpha l}{\alpha - 1}.$$

Так как  $X_j$  – независимые одинаково распределенные по Парето случайные величины, то характеристическая функция их суммы равна произведению их характеристических функций  $f_{S_n}(t) = [f_{X_j}(t/(bn^{1/\alpha}))]^n$ . Поэтому сначала была вычислена характеристическая функция нормированной суммы  $S_n$  из (1), в которой отценрированные математическим ожиданием слагаемые  $X_j$  имеют вид (3)

$$f_{S_n}(t) = \left[ pf_{X^+}\left(\frac{t}{bn^{1/\alpha}}\right) + qf_{X^-}\left(\frac{t}{bn^{1/\alpha}}\right) \right]^n.$$

Таким образом, было получено следующее выражение для характеристической функции суммы  $S_n$  смеси распределений Парето:

$$f_{S_n}(t) = \exp \left\{ p \left( \frac{t^2 r^2}{2b^2(\alpha-1)^2 n^{(2/\alpha-1)}} + \frac{(-itr)^\alpha \Gamma(2-\alpha, \frac{|t|r}{bn^{1/\alpha}})}{(\alpha-1)b^\alpha} + \frac{-i|t|^2 r^2 \frac{\pi}{2} \text{sign}(t)}{b^2 n^{(2/\alpha-1)}} \right) + q \left( \frac{t^2 l^2}{2b^2(\alpha-1)^2 n^{(2/\alpha-1)}} + \frac{(itl)^\alpha \Gamma(2-\alpha, \frac{|t|l}{bn^{1/\alpha}})}{(\alpha-1)b^\alpha} + \frac{i|t|^2 l^2 \frac{\pi}{2} \text{sign}(t)}{b^2 n^{(2/\alpha-1)}} \right) \right\} + R_n(t),$$

где  $\Gamma(\cdot, \cdot)$  – верхняя неполная гамма-функция, а  $R_n(t)$  – остаточный член и  $R_n(t) \rightarrow 0$  при  $n \rightarrow \infty$ .

## 2. Моделирование устойчивых случайных чисел в форме А

Теперь необходимо установить взаимосвязь между параметрами смеси распределений Парето и устойчивого закона. При  $n \rightarrow \infty$  характеристическая функция суммы  $S_n$  приобретает вид

$$f_{S_n}(t) \rightarrow \exp \left\{ \frac{-\Gamma(1-\alpha)}{b^\alpha} (p(-itr)^\alpha + q(itl)^\alpha) \right\}, \quad n \rightarrow \infty, \quad \forall t \in R.$$

Применяя формулу для главного значения степени комплексного числа, получим

$$f_Y(t) = \exp \left\{ \frac{-\Gamma(1-\alpha)(pr^\alpha + ql^\alpha) \cos \frac{\pi\alpha}{2} |t|^\alpha}{b^\alpha} \left( 1 - i \frac{pr^\alpha - ql^\alpha}{pr^\alpha + ql^\alpha} \text{tg} \frac{\pi\alpha}{2} \text{sign}(t) \right) \right\}. \quad (4)$$

Для того чтобы характеристическая функция суммы  $S_n$  имела в пределе вид характеристической функции устойчивого распределения (2), необходимо определить параметр  $b$  следующим образом:

$$b_k^\alpha = \frac{1}{\sigma_A^\alpha} \left( \frac{pr^\alpha \cos \frac{\pi\alpha}{2}}{1-\alpha} \Gamma(2-\alpha, \frac{|t|r}{b_{k-1} n^{1/\alpha}}) + \frac{ql^\alpha \cos \frac{\pi\alpha}{2}}{1-\alpha} \Gamma(2-\alpha, \frac{|t|l}{b_{k-1} n^{1/\alpha}}) \right). \quad (5)$$

Как видно, параметр  $b$  зависит от  $t$ . Необходимо выбрать единственное значение  $b$ , поэтому потребовалось ввести меру качества аппроксимации  $f_Y(t)$  функцией  $f_{S_n}(t)$ . Размеры существенной области  $D = \{t: t \in [-t_m, t_m]\}$ , в которой сравниваются характеристические функции, находятся на основании равенства Парсевала, применяемого в теории цифровой обработки сигналов [5], при этом должно выполняться

$$\frac{\gamma(1/\alpha, 2\sigma_A \cdot t_m^\alpha)}{\Gamma(1/\alpha)} < 1 - \varepsilon_{t_m},$$

где  $\gamma(\cdot, \cdot)$  – нижняя неполная гамма-функция, а параметр  $\varepsilon_{t_m}$  выбран равным  $10^{-3}$  и отвечает за уровень энергии отброшенных высокочастотных составляющих функции плотности  $f_Y(x)$ . Значения границ существенной области помещены в табл. 1.

Таблица 1

**Границы существенной области характеристической функции при  $\sigma_A = 1$**

$\alpha$	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	1,95	1,99	1,999
$t_m$	3	2,7	2,5	2,3	2,1	2	1,9	1,8	1,8	1,7	1,7	1,7

Порядок нахождения  $b$  следующий: для каждого  $t_j$  из  $D$  по рекуррентной формуле (5) методом простой итерации до достижения точности  $\epsilon_b = 10^{-3}$  вычисляется  $b(t_j)$ :

$$b = \{b_{k^*} : |b_{k^*} - b_{k^*-1}| \leq \epsilon_b\}.$$

Затем из всех полученных  $b(t_j)$  выбирается такое, которое удовлетворяет

$$b_{\text{опт}} = \arg \min_{b(t_j)} \max_{t_j \in D} |f_{S_n}(t_j) - f_Y(t_j)|.$$

Из (2) и (4) видно, что параметры смеси распределений Парето и параметры устойчивого распределения связаны соотношениями

$$\begin{cases} p + q = 1, \\ pr^\alpha + ql^\alpha = \sigma_A^\alpha, \\ pr^\alpha - ql^\alpha = \sigma_A^\alpha \cdot \beta_A. \end{cases} \quad (6)$$

**3. Результаты моделирования**

Для моделирования устойчивых случайных величин задаются следующие параметры:  $\beta_A \in (-1; 1)$ ,  $\sigma_A > 0$ ;  $p < 1$ . Можно показать, что разные  $p$  и  $q$ ,  $r$  и  $l$ , удовлетворяющие системе (6), дают хорошее приближение для теоретической функции плотности.

При сравнении эмпирической и теоретической функций плотности  $f_{S_n}(x)$  и  $f_Y(x)$  необходимо учитывать специфику поведения устойчивых распределений – «тяжелые хвосты». Поэтому было наложено следующее ограничение на диапазон  $[x_l, x_r]$ , в котором проверяется работа датчика. Граничные значения  $x_l, x_r$  должны удовлетворять уравнению

$$f_Y(x) = \epsilon_f.$$

Параметр  $\epsilon_f$  отвечает за размеры существенной области  $f_Y(x)$ . При реализации алгоритма был выбран уровень  $\epsilon_f = 0,01$ . В этой области необходимо определить шаг дискретизации. Для восстановления функции плотности в теории цифровой обработки сигналов используют теорему Котельникова [5], которая устанавливает границы для шага по оси  $x$ :  $\Delta x \leq 1/(2t_m)$ . Полагая  $\Delta x = 1/(2t_m)$ , получим сетку значений  $\{x_j = x_l + \Delta x \cdot j, j = 0, \dots, N\}$ , где  $N = [(x_r - x_l)/\Delta x]$ .

Для оценки качества аппроксимации  $f_Y(x)$  эмпирической функцией плотности  $f_{S_n}(x)$  вычисляются следующие показатели:

1) максимальное отклонение функций

$$\max AE = \max_{x_j} |f_Y(x_j) - f_{S_n}(x_j)|;$$

2) средняя абсолютная ошибка

$$MAE = \frac{1}{N+1} \sum_{j=0}^N |f_Y(x_j) - f_{S_n}(x_j)|;$$

3) средняя абсолютная процентная ошибка

$$MAPE = \frac{1}{N+1} \sum_{j=0}^N \frac{|f_Y(x_j) - f_{S_n}(x_j)|}{f_Y(x_j)} \cdot 100 \% .$$

В табл. 2 представлены параметры смеси распределений Парето и значения функционалов ошибки для различных значений параметров  $\alpha$  и  $\beta_A$  при  $p = 0,5$ ,  $K = 10^5$ ,  $n = 10^4$ . Отметим, что для отрицательных  $\beta_A$  необходимо поменять места значения  $r$  и  $l$ .

Таблица 2

**Параметры смеси распределений Парето и значения функционалов ошибки для различных значений параметров  $\alpha$  и  $\beta_A$ ,  $\sigma_A=1$ ,  $K = 10^5$ ,  $n = 10^4$ ,  $p=0,5$  при  $\beta \neq 1$ ,  $p=1$  при  $\beta=1$**

$\alpha$	$\beta$	$r$	$l$	$b$	$\max AE$	$MAE$	$MAPE, \%$
1,2	0	1	1	1,6264	0,009878	0,001894	0,0019
	0,25	1,204	0,787	1,6263	0,008067	0,008067	3,9253
	0,75	1,594	0,315	1,6249	0,009280	0,002038	3,4554
	1	1	-	1,6265	0,008247	0,002122	3,8062
1,5	0	1	1	1,7745	0,010086	0,003207	4,7693
	0,25	1,16	0,825	1,7735	0,012438	0,003472	5,5148
	0,75	1,452	0,397	1,7647	0,013317	0,004831	4,9774
	1	1	-	1,7745	0,015652	0,005233	5,3763
1,7	0	1	1	1,8979	0,011477	0,003999	4,8542
	0,25	1,140	1,140	1,8957	0,013808	0,005252	6,8994
	0,75	1,390	0,442	1,8766	0,024535	0,008694	7,9243
	1	1	-	1,9031	0,022852	0,010870	9,4221
1,9	0	1	1	2,0063	0,019069	0,008836	7,6308
	0,25	1,125	0,859	2,0027	0,022860	0,009066	7,8948
	0,75	1,343	0,482	1,9755	0,032926	0,015570	11,0892
	1	1	-	2,0495	0,041990	0,019988	13,9168
1,95	0	1	1	2,0378	0,013389	0,007777	6,5597
	0,25	1,121	0,863	2,0338	0,020848	0,008471	7,0071
	0,75	1,332	0,491	2,0134	0,042061	0,018752	12,4505
	1	1	-	2,0889	0,043147	0,022561	14,8206

Работа датчика продемонстрирована на рис. 1 и 2, где, кроме того, для сравнения представлена теоретическая функция плотности  $f_Y(x)$  и указана информация о количестве слагаемых  $X_j - n$ , параметрах устойчивого закона, количестве сгенерированных устойчивых чисел  $K$ .

Из графиков, представленных на рис. 2, видно, что при  $\alpha \rightarrow 2$  предельные формулы для характеристической функции суммы  $S_n$  и полученное из них выражение для поиска параметра  $b$  не дают необходимой точности аппроксимации, а следовательно, требуется уточнение формул при конечном  $n$ .

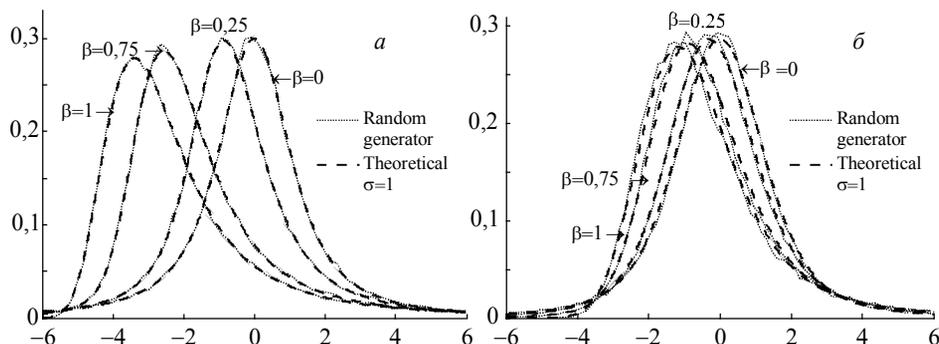


Рис.1. График функций плотностей сгенерированных чисел и устойчивой случайной величины при разных  $\beta_A$  в форме А при  $n = 10^4$ ,  $K = 10^5$ ,  $\sigma_A = 1$ ;  $\alpha = 1, 2$  (а),  $\alpha = 1, 5$  (б)

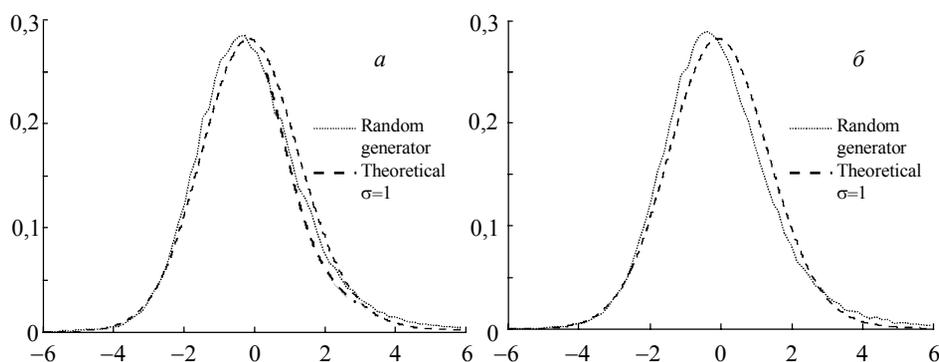


Рис.2. График функций плотностей сгенерированных чисел и устойчивой случайной величины в форме А при  $n = 10^4$ ,  $K = 10^5$ ,  $\beta_A = 0, 75$ ,  $\sigma_A = 1$ ;  $\alpha = 1, 9$  (а),  $\alpha = 1, 95$  (б)

### Заключение

В работе получены формулы для моделирования случайных величин, имеющих устойчивое распределение с характеристическим показателем  $\alpha \in (1, 2)$  для формы параметризации А. Работа датчика демонстрируется на графиках.

### ЛИТЕРАТУРА

1. Золотарев В.М. Одномерные устойчивые распределения. М.: Наука, 1983. 304 с.
2. Rachev S., Mittnik S. Stable Paretian Models in Finance. Wiley, 2000. 855 p.
3. Samorodnitsky G. and Taqqu M.S. Stable Non-Gaussian Random Processes. New York: Chapman and Hall, 1994. 632 p.
4. Барду Ф., Бушо Ж.-Ф., Аспе А., Коэн-Таннуджи К. Статистика Леви и лазерное охлаждение. Как редкие события останавливают атомы: пер. с англ. / под ред. В.П. Яковлева. М.: ФИЗМАТЛИТ, 2006. 216 с.
5. Маслов О.Н. Устойчивые распределения и их применение в радиотехнике. М: Радио и связь, 1994. 152 с.
6. Mittnik S., Rachev S., and Schwartz E. Value-at-risk and asset allocation with stable return distributions // Allgemeines Statistisches Archiv. 2002. V. 86. No. 1. P. 53–68.
7. Денисов В.И., Тимофеев В.С. Устойчивые распределения и оценивание параметров регрессионных зависимостей // Известия Томского политехнического университета. Томск: Изд-во ТПУ, 2011. Т. 318. № 2. С. 10–15.

8. *Uchaikin V.V., Zolotarev V.M.* Chance and Stability. Stable Distributions and their Applications. Utrecht: VSP, 1999. 594 p.
9. *Janicki A., Weron A.* Simulation and Chaotic Behavior of  $\alpha$ -Stable Stochastic Processes. New York: Marcel Dekker, 1994. 355 p.
10. *Chambers J., Mallows C., Stuck B.* A method for simulating stable random variables // Journal of the American Statistical Association. Theory and Methods Section. 1976. V. 71. No. 354. P. 340–344.
11. *Nolan J.P.* Numerical calculation of stable densities and distribution functions // Commun. Statist. Stochastic Models. 1997. V.13. P. 759–774.

*Багрова Инна Александровна*  
Тверской государственный университет  
E-mail: inna@tversu.ru

Поступила в редакцию 27 апреля 2012 г.

*Bagrova Inna A.* (Tver State University). **On the stable-number generator design with characteristic exponent greater than one.**

Keywords: modeling random variable, stable distributions, Pareto distribution.

The stable distribution random number generator with characteristic factor greater than one is developed. Simulation algorithm is based on the General Central Limit Theorem according to that stable distributions be limit distributions for normalised and centralised sum of independent and identically distributed random variables from the domain of attraction for stable distributions:

$$S_n = \frac{X_1 + \dots + X_n}{bn^{1/\alpha}} \xrightarrow{d} Y \text{ if } n \rightarrow \infty, \alpha \in (1, 2).$$

As summands we use the Pareto mixtures whose carriers belong to the positive and negative semi-axes. The limit equation for characteristic function of sum  $S_n$  was obtained. The relation for parameters of stable distribution in (A) parameterization and Pareto distribution parameters was obtained, as well as the expression for normalising factor  $b$ . Resulted characteristics – values for three error functionals corresponding to Pareto mixtures with specified mixture parameters ( $r, l$ ) and  $(\alpha, \beta_A)$  parameters are given in the table. The simulation results are illustrated by the figures 1 and 2. Note that limit equations for  $S_n$ , characteristic function and corresponding expression for  $b$  when  $\alpha \rightarrow 2$  do not provide sufficient approximation accuracy.