Безгранично делимые распределения

Случайная величина ξ называется *безгранично делимой*, если для любого $N > 1$ существуют независимые одинаково распределенные случайные величины $\{z_n(N)\}_1^N$, такие что $\xi = \sum_1^N z_n(N)$. Это понятие было введено де Финетти (Bruno de Finetti) в 1929 г. В терминах характеристических функций свойство безграничной делимости означает, что для любого N существует характеристическая функция $h_{z,N}(u)$ такая, что характеристическая функция $h_\xi(u)$ равна $\left(h_{z,N}(u)\right)^N$. Из леммы Шура следует, что безгранично делимые распределения образуют алгебру относительно свертки.

Извлечение корня по стандартному правилу:

$$(h_\xi(u))^{1/n} \stackrel{\mathrm{def}}{=} |h_\xi(u)|^{1/n} e^{i\mathrm{Arg}(h_\xi(u))/n}$$

не нарушает непрерывность в нуле и обеспечивает равенство $h_{z,N}(0) = (h_\xi(0))^{1/n} = 1$. Поэтому содержательная часть условия на класс безгранично делимых распределений состоит в требовании положительной определенности функции $(h_\xi(u))^{1/n}$. Если функции такого вида, соответствующие h_ξ, положительно определены, то из теоремы Шура следует положительная определенность функций вида $(h_\xi(u))^a$, $a = m/n$, где m, n – целые. По непрерывности это свойство распространяется на все вещественные $a \in \mathbb{R}_+$.

Пример 1

Класс б. д. распределений достаточно узок. Равномерное распределение не является безгранично делимым. Для $\xi \in U[-10, 10]$, *х. ф.* $h_\xi(u)$ *равна* $\sin 10u/10u$. *Для 3-х пар случайно выбранных точек*

$$z = \{0.1249 + 0.9142i,\ 0.1696 + 0.6263i,\ 0.5606 + 0.8953i\},$$
$$u = \{0.6432,\ 0.7062,\ 0.0859\}$$

сумма $\sum_{j,k=1}^{3} h_\xi^{1/2}(u_j - u_k)\bar{z}_j z_k = 3.536 + 0.757i$ *не является положительной при естественном выборе ветви корня.*

На сопряженной ветви нарушается непрерывность в нуле, а сдвиги и изменения масштаба $\xi \to a + b\xi$ не могут создать или нарушить б. д. свойство. Поэтому равномерные распределения не являются безгранично делимыми.

Объяснение состоит в том, что х. ф. безгранично делимых законов, в отличие от функции

$$h_\xi(u) = \frac{\sin \lambda u}{\lambda u} = \int_{-\lambda}^{\lambda} e^{iux}\, P_\xi(dx), \quad P_\xi(dx) = \frac{dx}{2\lambda}, \quad |\xi| \le \lambda,$$

не могут обращаться в нуль (см. Б. В. Гнеденко, Курс теории вероятностей, М:, Наука, 1988, стр. 265–266). Так что равномерное распределение – простейший пример распределения, не являющегося безгранично делимым.

Теорема 1

Х. ф. безгранично делимых распределений не могут обращаться в нуль.

Доказательство. Если $h_\xi(u)$ –х. ф. б. д. с. в. ξ, то $\overline{h}_\xi(u)$ х. ф. $-\xi$ и $h_\eta(u) = |h_\xi(u)|^2 \geq 0$ х. ф. б. д. разности $\eta = \xi - \xi'$, причем из $h_\xi(u) \neq 0$ в фиксированной точке u следует $h_\eta(u) > 0$. Поскольку $h_\eta(0) = 1$ и h_η –непрерывная -х. ф. б. д. распределения, то, во-первых, существует U_ε-окрестность точки 0, где $h_\eta(u) > 0$ и, во-вторых, $h_\eta^{1/n}(u)$ также являются х. ф. Для любой х. ф. $h_\eta(u) \geq 0$ выполнены оценки

$$\begin{aligned} 1 - h_\eta(2u) &= \int (1 - \cos(2ux))\, P_\eta(dx) = 2\int \sin^2 ux\, P_\eta(dx) \\ &= 2\int (1 - \cos ux)(1 + \cos ux)\, P_\eta(dx) \\ &\leq 4\int (1 - \cos ux)\, P_\eta(dx) = 4(1 - h_\eta(u)). \end{aligned}$$

так как $\mathrm{Im}\, h_\eta(u) = 0$. Поэтому

$$h_\eta(2u) \geq 1 - 4(1 - h_\eta(u)). \tag{1}$$

Отсюда следует, что для неотрицательной б. д. х. ф. $h_\eta(u)$ и любого натурального n выполнено $1 - h_n(2u) \le 4(1 - h_n(u))$, где $h_n(u) \stackrel{\text{def}}{=} h_\eta(u)^{1/n} \ge 0$. Если $h_\eta(u) \ge \epsilon$ при всех $u: \ |u| \le a = a(\epsilon)$, то $h_n(u) \ge \epsilon^{1/n}$ на этом множестве и из $h_n(2u) \ge 1 - 4(1 - h_n(u))$ получаем

$$h_\eta(2u) = (h_n(2u))^n \ge \lim_n \big(1 - 4(1 - h_n(u))\big)^n \ge \lim_n \big(1 - 4(1 - \epsilon^{1/n})\big)^n$$

$$= \lim_n \big(1 - 4(1 - e^{-\frac{1}{n}\ln\frac{1}{\epsilon}})\big)^n = \lim_n \left(1 - \frac{4}{n}\ln\frac{1}{\epsilon}\right)^n = e^{-4\ln\frac{1}{\epsilon}} = \epsilon^4.$$

Из оценки $\min_{|u|\le a} h_\eta(u) \ge \epsilon$ следует $\min_{|u|\le 2a} h_\eta(u) \ge \epsilon^4$ и

$$\min_{|u|\le 2^2 a} h_\eta(u) \ge \left(\epsilon^4\right)^4 > 0, \dots, \min_{|u|\le 2^m a} h_\eta(u) \ge \epsilon^{4^m} > 0$$

для любого $m > 0$. Поэтому безгранично делимые х. ф. не обращаются в нуль ни в какой точке $\mathbb{R}$. $\square$

Функция $h_P(u) = e^{a(e^{iub}-1)}$ является х. ф. для пуассоновского распределения $p(\xi = bn) = \frac{a^k}{k!}e^{-a}$. Нетрудно видеть, что функции

$$h_N(u) = (h(u))^{1/N} = e^{aN^{-1}(e^{iub}-1)}$$

также являются х. ф. пуассоновских $\frac{(a/N)^k}{k!}e^{-a/N}$.

Функция $h_G(u) = e^{-\frac{(u\sigma)^2}{2}+iuc}$ является х. ф. нормального распределения с плотностью $p(x) = (2\pi\sigma^2)^{-1/2}e^{-\frac{(x-c)^2}{2\sigma^2}}$, а функциям

$$h_N(u) = (h(u))^{1/N} = e^{-\frac{(u\sigma)^2}{2N}+iuc/N}$$

соответствуют распределения $(N/2\pi\sigma^2)^{-1/2}e^{-\frac{(x-c/N)^2}{2N\sigma^2}}$. Оба семейства состоят из б. д. распределений.

Поэтому произведения любого числа х. ф.

$$h(u) = h_G(u)h_P(u) = e^{\psi(u)},$$

$$\psi(u|a,b,c) = a(e^{iub} - 1) - \frac{(u\sigma)^2}{2} + iuc, \qquad (2)$$

и все функции вида

$$h(u) = e^{\Phi(u)}, \quad \Phi(u) = \int \psi(u|a,b,c)\mu(da,db,dc)$$

с положительными конечными мерами μ, имеющими два первых абсолютных момента, являются х. ф. безгранично делимых распределений. Случайные величины $z_n(N)$ в определении безграничной делимости имеют в этом случае характеристические функции $h_N(u) = e^{\frac{1}{N}\Phi(u)}$.

Это наблюдение обобщается теоремой де Финетти–Леви–Хинчина, которую мы приводим без доказательства.

Теорема 2

Случайная величина является безгранично делимой тогда и только тогда, когда существуют постоянные $\sigma^2 \in \mathbb{R}_+$, $c \in \mathbb{R}$ *и конечная положительная мера* $m: \; m(\{0\}) = 0$, *такие, что* $h(u) = e^{\Phi(u)}$, *где функция* $\Phi(u)$ *имеет вид*

$$\Phi(u) = iuc - \frac{(u\sigma)^2}{2} + \int_{\mathbb{R}} \left(e^{iux} - 1 - \frac{iux}{1+x^2} \right) \frac{1+x^2}{x^2} m(dx). \quad (3)$$

Интеграл сконструирован так, чтобы формула имела вид (2) и не требовала существования моментов меры m. Достаточность условия (3) является очевидным следствием формулы (2).

В то же время нетрудно видеть, что

$$h_1(u) = e^{e^{iu\mu}-1}, \qquad h_2(u) = e^{\left(e^{iu\mu}-1-\frac{iu\mu}{1+\mu^2}\right)\frac{1+\mu^2}{\mu^2}} \tag{4}$$

являются положительно определенными х.ф. случайных величин со средним значением μ:

$$\mu_1 = \frac{1}{i}\frac{\partial}{\partial u} h_{1,2}(u)|_{u=0} = \mu. \tag{5}$$

Поэтому из (3) следует простое выражение для момента случайной величины, х. ф. которой описана в теореме 2:

$$\mu_\Phi = \frac{1}{i}\frac{\partial}{\partial u}\, e^{\Phi(u)}|_{u=0} = c + \int xm(dx). \tag{6}$$

Упражнение 1

Докажите формулы (5)-(6).

Рассмотрим важный пример характеристических функций б. д. распределений. Заметим, что справедливы равенства

$$h_+(u,a) = \int_0^\infty (e^{iut}-1)t^{-(1+a)}\,dt = e^{-\frac{i\pi a}{2}\mathrm{sign}\,u}|u|^a\Gamma(-a), \qquad (7)$$

$$h_-(u,a) = \int_{-\infty}^0 (e^{iut}-1)|t|^{-(1+a)}\,dt = e^{\frac{i\pi a}{2}\mathrm{sign}\,u}|u|^a\Gamma(-a)$$

при $0 < a < 1$, причем $e^{h_\pm(u,a)}$ -положительно определенные х. ф.характеристические функции. Аналогично,

$$\int_0^\infty (e^{iut}-1-iut)t^{-(1+a)}\,dt = h_+(u,a), \qquad (8)$$

$$\int_{-\infty}^0 (e^{iut}-1-iut)|t|^{-(1+a)}\,dt = h_-(u,a)$$

справедливы при $a \in (0,1)\cup(1,2)$. Поскольку $\Gamma(-a) < 0$ при $a \in (0,1)$ и $\Gamma(-a) > 0$ при $a \in (1,2)$ (см. рис.), то в обоих случаях $\mathrm{Re}\,h_\pm(u,a) < 0$ и $|e^{h_\pm(u,a)}| \le 1$.

Упражнение 2

Проверьте равенства (7)-(8) в системе Mathematica аналитически или численно в MatLab.

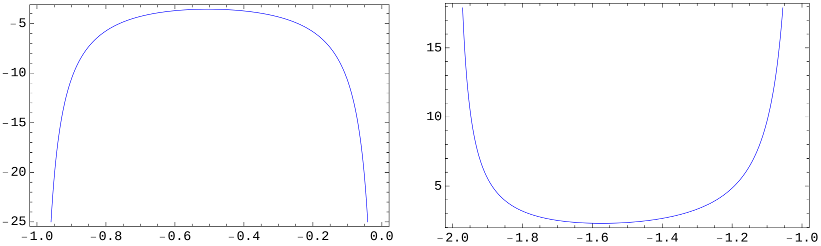

Рис. 1: Гамма функция на интервалах $(-1, 0)$, $(-2, -1)$

Запишем вычисленные выше интегралы в виде (3):

$$e^{-\frac{i\pi a}{2}\operatorname{sign} u}|u|^a\,\Gamma(-a) = \int_0^\infty (e^{iux}-1)x^{-(1+a)}\,dx$$

$$= ic_1 u + \int_0^\infty \left(e^{iux} - 1 - \frac{iux}{1+x^2}\right)\frac{1+x^2}{x^2}\,m(dx) \text{ при } 0<a<1,$$

$$e^{-\frac{i\pi a}{2}\operatorname{sign} u}|u|^a\,\Gamma(-a) = \int_0^\infty (e^{iux}-1-iux)x^{-(1+a)}\,dx$$

$$= ic_2 u + \int_0^\infty \left(e^{iux} - 1 - \frac{iux}{1+x^2}\right)\frac{1+x^2}{x^2}\,m(dx) \text{ при } 1<a<2,$$

где мера $m(dx) = \frac{x^{1-a}}{1+x^2}dx$ - конечна при $0<a<2$ и $c_1 = \int_0^\infty \frac{x^{-a}}{1+x^2}dx$ при $a\in(0,1)$, $c_2 = -\int_0^\infty \frac{x^{2-a}}{1+x^2}dx$ при $a\in(1,2)$ и д. с. в. имеющие х. ф. вида $h(u)=e^{\Phi(u)}$

$$\Phi(u) = \sum \lambda_k e^{-\frac{i\pi a_k}{2}\operatorname{sign} u}|u|^{a_k}\,\Gamma(-a_k), \quad \lambda_k>0,\ a_k\in(0,1)\cup(0,2)$$

безгранично делимы.

Как необходимое условие (т.е. как утверждение о том, что других безгранично делимых законов не существует), теорема 2 является неожиданным и не очевидным утверждением[1].
Как достаточное условие, эта теорема не описывает принципиально новых компонент по сравнению с рассмотренными выше примерами, более того, в приведенной формулировке эта теорема не гарантирует единственность параметров σ, c, μ.

[1] Доказательство можно найти в книге В.Б. Гнеденко и А.Н. Колмогорова "Предельные распределения сумм независимых случайных величин" (Гос. издательство технико–теоретической литературы. М.–Л., 1949).

Действительно, разлагая в ряд Тейлора по степеням u логарифм характеристической функции и интеграл (3), получим уравнения для коэффициентов для линейных и квадратичных по u членов

$$c_0 = 0, \quad c + \mu_1 = c_1, \quad \sigma^2 + \mu_0 + \mu_2 = c_2,$$

где $\{c_n\}_0^\infty$ —известные постоянные, определяющие разложение логарифма х. ф. в ряд Тейлора:
$\Phi(u) = \sum_{n>0} c_n u^n$, а μ_n – неизвестные моменты меры μ.
В общем виде система зацепляющихся уравнений

$$\mu_1 = c_1 - c, \; \mu_0 + \mu_2 = c_2 - \sigma^2, \; \dots, \; \mu_n + \mu_{n+2} = c_{n+2}, \; \dots.$$

содержит 4 неизвестных $(c, \{\mu_n\}_0^2)$ в первых двух уравнениях. Неоднозначность их решения распространяется на всю систему.

Устойчивые законы

Более узкий класс образуют *устойчивые законы*. Д. с. в. ξ имеет устойчивый закон распределения, если для *любых* $c_n > 0$, $n = 1, \dots, N$ и N независимых выборочных значений $\{\xi_n\}_1^N$ д. с. в. ξ существуют $A_N \in \mathbb{R}_+$, $B_N \in \mathbb{R}$, такие что

$$\xi = \frac{1}{A_N}\left(\sum_{n=1}^{N} c_n\xi_n - B_N\right) = \sum_{n=1}^{N} z_n(N), \quad z_n(N) = \frac{c_n\xi_n - N^{-1}B_N}{A_N}.$$

Напомним, что в определении бесконечно делимых случайных величин $c_n = 1$, $A_N = 1$, $B_N = 0$.
Линейное преобразование случайной величины индуцирует очевидное преобразование х. ф., так что в терминах х. ф. условие устойчивости имеет вид

$$\exists A_N,\ B_N: \quad h_\xi(A_N u) = \prod_{n=1}^{N} h_\xi(c_n u))e^{-iB_N u}. \tag{9}$$

Эмпирические данные с такими распределениями встречаются в статистиках физического и экономического происхождения и в общем случае имеют х. ф., зависящую от нескольких параметров, описываемую с помощью *представления Леви–Хинчина*.

Теорема 3

Д. с. в. x имеет устойчивое распределение тогда и только тогда, когда ее х. ф. имеет вид

$$h_{a,\beta,b,c}(u) = \exp\{iub - |cu|^a(1 - i\beta\theta\,\mathrm{sign}\,u)\}, \tag{10}$$

где $\theta = \{\tan\frac{\pi a}{2}$ *при* $a \in (0,1)\cup(1,2)$; $\frac{2}{\pi}\ln|u|$ *при* $a = 1\}$, $\beta \in [-1,1]$, $b \in \mathbb{R}$, $c > 0$; $\beta = 0$ *при* $a = 2$.

При этом $A_N = \left(\sum_{n=1}^N c_n^a\right)^{\frac{1}{a}}$, $B_N = \left(\sum_{n=1}^N c_n - A_N\right)b$ в формуле (9).

Достаточность условий теоремы 3 вытекает из комбинации формул (7)-(8) и сдвига. Доказательство достаточно провести для $N = 2$. При $a \in (0,1) \cup (1,2)$ получаем формулу (10):

$$h(u, c_1, c_2) = h_+(u, c_1) h_-(u, c_2) e^{iub}$$
$$= e^{iub - (c_1 e^{-\frac{i\pi a}{2}\operatorname{sign} u} + c_2 e^{\frac{i\pi a}{2}\operatorname{sign} u})\,|u|^a\,\Gamma(-a)}$$
$$= e^{iub - |u|^a\,\Gamma(-a)\,(c_1 + c_2)\left(\cos\frac{\pi a}{2} - i(c_1 - c_2)\,(\operatorname{sign} u \sin\frac{\pi a}{2})\right)}$$
$$= e^{iub - |u|^a \Gamma(-a)\,(c_1 + c_2)\,\left(\cos\frac{\pi a}{2}\left(1 - i\frac{c_1 - c_2}{c_1 + c_2}\operatorname{sign} u\,\tan(\frac{\pi a}{2})\right)\right)},$$

имеет вид $\exp\{iub - |cu|^a(1 - i\beta\theta \operatorname{sign} u)\}$, где
$c: \ (c_1 + c_2)\Gamma(-a)\cos\frac{\pi a}{2} = c^a > 0$ при $a \in (0,1) \cup (1,2)$,
$\frac{c_2 - c_1}{c_1 + c_2} = \beta \in [-1, 1]$.
Для произведения $h_\pm(u, c_1) h_\pm(u, c_2) e^{iub}$ доказательство аналогично: $\beta = 1$ для произведения $h_+(u, c_1) h_+(u, c_2) e^{iub}$ и $\beta = -1$ для $h_-(u, c_1) h_-(u, c_2) e^{iub}$.

В общем случае явный вид плотности устойчивых распределений не выражается в элементарных функциях, но при больших $|x|$ и $a \in (0,2)$ известна его асимптотика

$$p(x) = p_{a,\beta,b,c}(x) \approx \frac{c^a(1+\beta)\sin(\pi a/2)\Gamma(1+a)}{\pi\,|x-b/c|^{1+a}} \tag{11}$$

и представление в виде ряда

$$p_{a,\beta,b,c}(x) = \frac{1}{\pi}\mathrm{Re}\sum_{n=1}^{\infty}\frac{(-q)^n}{n!}\left(\frac{i}{x-b/c}\right)^{an+1}\Gamma(an+1), \tag{12}$$

где $q = c^a(1-i\beta\theta)$. Параметры c и b/c задают масштаб и сдвиги распределения по оси x, $\beta \in [-1,1]$ – характеризует его асимметрию. Для симметричных распределений $\beta = 0$. Для значений $a < 2$ такие распределения называются *распределениями с тяжелыми хвостами*.

Если z_n – независимые выборочные значения фиксированной случайной величины z, причем существуют $a_N > 0$ и $b_N \in \mathbb{R}$ такие, что $\frac{1}{a_N}(\sum z_n - b_N) \xrightarrow{d} \xi$ (сходимость по распределению), то говорят, что z *принадлежит области притяжения* случайной величины ξ. Область притяжения $\mathcal{D}(\xi)$ состоит из всех случайных величин z, удовлетворяющих этому условию. Следующая теорема описывает множество случайных величин с устойчивыми законами распределения как область притяжения распределений с тяжелыми хвостами.

Теорема 4

Пусть $\{z_n\}$ *– последовательность независимых одинаково распределенных действительных случайных величин с полиномиально убывающими хвостами при* $|x| \to \infty$*:*

$$P(z_1 > x) \sim a_+ x^{-a}, \quad P(z_1 < -x) \sim a_- |x|^{-a}, \quad a \in (0, 2),$$

$a_\pm > 0$. *Тогда существуют такие* $\{a_N\} > 0$, $\{b_N\} \in \mathbb{R}$, $\beta = \frac{a_+ - a_-}{a_+ + a_-}$ *и* $c = a_+ + a_-$, *что*

$$\frac{1}{a_N}\left(\sum_{n=1}^{N} z_n - b_N\right) \xrightarrow{d} \xi_{a,\beta,0,c}\ ,$$

где $\xi_{a,\beta,0,c}$ *– случайная величина, имеющая устойчивое распределение с характеристической функцией вида* (10)

Здесь

$$a_N = \left(\frac{\pi c N}{2\Gamma(a)\sin(a\pi/2)} \right)^{1/a}$$

и возможны следующие варианты правильного выбора центрирующих коэффициентов b_N в зависимости от значений a:

Таблица 1:

$\mathrm{dom}\, a$	b_N
$(1,2)$	Na
$\{1\}$	$\beta c N \log N$
$(0,1)$	0