Записки научных
семинаров ПОМИ
Том 501, 2021 г.

Ф. Гётце, А. Ю. Зайцев

СХОДИМОСТЬ К БЕСКОНЕЧНОМЕРНЫМ ОБОБЩЕННЫМ РАСПРЕДЕЛЕНИЯМ ПУАССОНА НА ВЫПУКЛЫХ МНОГОГРАННИКАХ

Цель данной работы – предоставить дополнение к статье авторов [4]. Показано, что наши результаты о приближении распределений сумм независимых слагаемых сопровождающими обобщенными законами Пуассона и оценки близости последовательных сверток многомерных распределений на выпуклых многогранниках могут быть почти автоматически перенесены на бесконечномерный случай. Ясно, что это делает результаты существеннно более общими.

Введем сначала некоторые обозначения. Пусть $1 \leqslant d \leqslant \infty$ и $\mathfrak{F}_d$ обозначает множество вероятностных распределений, определенных на борелевской σ-алгебре подмножеств евклидова пространства $\mathbf{R}^d$, а $\mathcal{L}(\xi) \in \mathfrak{F}_d$ – распределение d-мерного случайного вектора ξ. Пусть $\mathfrak{F}_d^s \subset \mathfrak{F}_d$ – множество симметричных распределений. Для $F \in \mathfrak{F}_d$ мы обозначим соответствующие характеристические функции через $\widehat{F}(t)$, $t \in \mathbf{R}^d$, а функции распределения – через

$$F(x) = F\{(-\infty, x_1] \times \cdots \times (-\infty, x_d]\},$$

$x = (x_1, \dots, x_d) \in \mathbf{R}^d$. Равномерное расстояние Колмогорова определяется формулой

$$\rho(G, H) = \sup_{x \in \mathbf{R}^d} \big| G(x) - H(x) \big|, \quad G, H \in \mathfrak{F}_d.$$

Символами c и $c(\,\cdot\,)$ мы обозначаем вообще говоря различные положительные абсолютные постоянные и величины, зависящие только от аргумента в скобках. При $0 \leqslant \alpha \leqslant 2$ мы обозначим

$$\mathfrak{F}_d^{(\alpha)} = \Big\{ F \in \mathfrak{F}_d^s : \widehat{F}(t) \geqslant -1 + \alpha, \text{ при всех } t \in \mathbf{R}^d \Big\}, \quad \mathfrak{F}_d^+ = \mathfrak{F}_d^{(1)}.$$

Заметим, что наши обозначения применимы даже для $d = \infty$, то есть для распределений в гильбертовом пространстве $\mathbf{R}^\infty = \mathbf{H}$.

Ключевые слова: суммы независимых случайных величин, близость последовательных сверток, выпуклые многогранники, аппроксимация, неравенства.

Работа поддержана SFB 1283, Bielefeld, и грантом РФФИ–ННИО 20-51-12004. Работа второго автора поддержана грантом РФФИ 19-01-00356.

Произведения и степени мер будут пониматься в смысле свертки: $GH = G * H,\ H^m = H^{m*},\ H^0 = E = E_0$, где E_x – распределение, сосредоточенное в точке $x \in \mathbf{R}^d$. Естественным аппроксимирующим безгранично делимым распределением для $\prod_{i=1}^n F_i$ является сопровождающее обобщенное распределение Пуассона $\prod_{i=1}^n e(F_i)$, где

$$e(H) = e^{-1} \sum_{k=0}^{\infty} \frac{H^k}{k!}, \quad H \in \mathfrak{F}_d,$$

и, в более общем случае,

$$e(\alpha H) = e^{-\alpha} \sum_{k=0}^{\infty} \frac{\alpha^k H^k}{k!}, \quad \alpha > 0. \tag{1}$$

Хорошо известно, что распределение $e(\alpha H)$ безгранично делимо и $e(nH) = (e(H))^n$, $n \in \mathbf{N}$.

Т. Арак [1] показал, что если $F \in \mathfrak{F}_1^+$ является симметричным одномерным распределением с характеристической функцией, неотрицательной при всех $t \in \mathbf{R}$, то

$$\rho(F^n, e(nF)) \leqslant c\, n^{-1}, \tag{2}$$

Он разработал и использовал так называемый метод треугольных функций (см. [2, глава 3, параграфы 2–4]).

А. Ю. Зайцев [8] применил методы, использованные Араком при доказательстве неравенства (2) (см. также [2, глава 5, параграфы 2, 5–7]). Позднее ему удалось видоизменить эти методы, приспособив их к многомерному случаю (см. [9]– [13]). В частности, в работе [12] получен многомерный аналог неравенства (2).

С помощью метода треугольных функций и его обобщений было получено несколько неравенств типа

$$\rho(G, H) \leqslant c(d)\, \varepsilon \tag{3}$$

когда $0 < \varepsilon < 1$ мало, $G, H \in \mathfrak{F}_d$.

Неравенство (3) эквивалентно справедливости неравенства

$$\bigl|G\{P\} - H\{P\}\bigr| \leqslant c(d)\, \varepsilon \tag{4}$$

для всех множеств P вида

$$P = \bigl\{x \in \mathbf{R}^d : \langle x, e_j \rangle \leqslant b_j,\ j = 1, \ldots, d\bigr\}, \tag{5}$$

где $e_j \in \mathbf{R}^d$ – векторы стандартного евклидова базиса, $-\infty < b_j \leqslant \infty$, $j = 1, \ldots, d$.

При $m \in \mathbf{N}$ мы обозначим через $\mathcal{P}_m$ совокупность выпуклых многогранников $P \subset \mathbf{R}^d$, представимых в виде

$$P = \bigl\{x \in \mathbf{R}^d : \langle x, t_j\rangle \leqslant b_j,\ j = 1, \dots, m\bigr\},$$

где $t_j \in \mathbf{R}^d$, $-\infty < b_j \leqslant \infty$, $j = 1, \dots, m$. Определим также

$$\rho_m(G, H) = \sup_{P \in \mathcal{P}_m} \bigl|G\{P\} - H\{P\}\bigr|.$$

Наши результаты показывают, что ρ_m является естественным многомерным вариантом расстояния Колмогорова. Аналогичное определение многомерного расстояния Леви дано в работе [5].

В конечномерном случае $d < \infty$ утверждения наших теорем 1–5 были доказаны в статье авторов [4], см. также [6]– [13]. В следующих ниже формулировках $d = \infty$.

Теорема 1. *Пусть* $d = \infty$ *и* $F \in \mathfrak{F}_d^{(\alpha)}$, $0 \leqslant \alpha \leqslant 2$, $m, n \in \mathbf{N}$. *Тогда*

$$\max\bigl\{\rho_m(F^n, e(nF)), \rho_m(F^n, F^{n+1})\bigr\} \leqslant c(m)\Bigl(n^{-1} + \exp(-n\alpha + c\, m\, \log^3 n)\Bigr). \tag{6}$$

Ясно, что из теоремы 1 следует неравенство (2).

Теорема 2. *Предположим, что* $d = \infty$ *и распределения* $G_i \in \mathfrak{F}_d$ *представлены в виде*

$$G_i = (1 - p_i)\, E + p_i\, V_i, \quad i = 1, \dots, n, \tag{7}$$

где $V_i \in \mathfrak{F}_d$ *– произвольные распределения,* $0 \leqslant p_i \leqslant p = \max\limits_{1 \leqslant i \leqslant n} p_i$,

$$m \in \mathbf{N}, \quad G = \prod_{i=1}^{n} G_i, \quad D = \prod_{i=1}^{n} e(G_i).$$

Тогда

$$\rho_m(G, D) \leqslant c(m)\, p. \tag{8}$$

Ситуацию, рассмотренную в теореме 2, можно интерпретировать как сравнение выборки, содержащей независимые наблюдения нежелательных редких событий (таких как катастрофы, эпидемии, аварии, банкротства и т. п.) с пуассоновским точечным процессом, который получается после пуассонизации исходной выборки (см. [3, 14]). Пусть $Y_1, Y_2, \dots, Y_n$ – независимые неодинаково распределенные наблюдения, принимающие значения в измеримом пространстве $(\mathfrak{X}, \mathcal{S})$. Одновременно мы наблюдаем некоторые редкие события. Предположим, что множество $\mathfrak{X}$ представлено в виде $\mathfrak{X} = \mathfrak{X}_1 \cup \mathfrak{X}_2$, причем $\mathfrak{X}_1, \mathfrak{X}_2 \in \mathcal{S}$, $\mathfrak{X}_1 \cap \mathfrak{X}_2 = \varnothing$. Мы говорим, что i-ое редкое событие произошло, если $Y_i \in \mathfrak{X}_2$. Соответственно, оно не произошло, если $Y_i \in \mathfrak{X}_1$.

Нас интересуют суммарные потери

$$S = f(Y_1) + \cdots + f(Y_n), \tag{9}$$

где $f : \mathfrak{X} \to \mathbf{R}^d$ — функция потерь, такая что $f(x) = 0$ при $x \in \mathfrak{X}_1$. Пусть $G_i = \mathcal{L}(f(Y_i))$, $i = 1, 2, \ldots, n$. Тогда распределения $G_i \in \mathfrak{F}_d$ могут быть представлены в виде смесей

$$G_i = (1 - p_i)\, E + p_i\, V_i, \tag{10}$$

где $E, V_i \in \mathfrak{F}_d$ – условные распределения векторов $f(Y_i)$ при условиях $Y_i \in \mathfrak{X}_1$ и $Y_i \in \mathfrak{X}_2$ соответственно,

$$0 \leqslant p_i = \mathbf{P}\big\{Y_i \in \mathfrak{X}_2\big\} = 1 - \mathbf{P}\big\{Y_i \in \mathfrak{X}_1\big\} \leqslant 1. \tag{11}$$

Мы имеем дело с редкими событиями, когда величина

$$p = \max_{1 \leqslant i \leqslant n} p_i \tag{12}$$

мала. Другими словами, это модель равномерно редких событий.

Пусть случайное множество $\mathbf{P} = \{Z_k\}$ – реализация пуассоновского точечного процесса на пространстве $\mathfrak{X}$ с мерой интенсивности $\sum\limits_{i=1}^{n} \mathcal{L}(Y_i)$. Сумма S, определенная в (9), имеет распределение $G = \prod\limits_{i=1}^{n} G_i$. Легко в то же время видеть, что

$$D = \prod_{i=1}^{n} e(G_i). \tag{13}$$

есть распределение вектора

$$T = \sum_{Z_k \in \mathbf{P}} f(Z_k). \tag{14}$$

Из теоремы 2 следует, что

$$\rho_m\bigg(\mathcal{L}\Big(\sum_i f(Y_i)\Big), \mathcal{L}\Big(\sum_{Z_k \in \mathbf{P}} f(Z_k)\Big)\bigg) \leqslant c(m)\, p. \tag{15}$$

Таким образом, оценивая совокупный убыток, можно заменить наблюдения $Y_1, Y_2, \ldots, Y_n$ пуассоновским точечным процессом $\mathbf{P}$, если наши редкие события равномерно редки. Тогда мы можем использовать известное свойство пространственной независимости пуассоновского точечного процесса.

В оставшейся части статьи мы изучим, насколько мала разница между F^{n+k} и F^n, то есть, насколько может измениться распределение суммы n независимых одинаково распределенных слагаемых после добавления к ней следующего слагаемого или группы слагаемых. Частный случай этой проблемы рассматривается в неравенстве (6) теоремы 1.

Теорема 3. *Пусть* $F \in \mathfrak{F}_d^s$, $d = \infty$ *и* $k, m, n \in \mathbf{N}$. *Тогда*

$$\rho_m(F^n, e(nF)) \leqslant c(m)\, n^{-1/2}, \tag{16}$$

$$\rho_m(F^n, F^{n+2k}) \leqslant c(m)\, k\, n^{-1}, \tag{17}$$

$$\rho_m(F^n, F^{n+2k+1}) \leqslant c\, m\, n^{-1/2} + c(m)\, k\, n^{-1}. \tag{18}$$

В частности,

$$\sup_{k \leqslant \sqrt{n}} \rho_m(F^n, F^{n+k}) \leqslant c(m)\, n^{-1/2}. \tag{19}$$

Для $m \in \mathbf{N}$, $t_1, \ldots, t_m \in \mathbf{R}^d$ мы обозначим через $\mathcal{P}(t_1, \ldots, t_m)$ совокупность выпуклых многогранников $P \subset \mathbf{R}^d$, представимых в виде

$$P = \bigl\{x \in \mathbf{R}^d : \langle x, t_j \rangle \leqslant b_j,\ j = 1, \ldots, m\bigr\},$$

где $-\infty < b_j \leqslant \infty$, $j = 1, \ldots, m$. Очевидно, что

$$\mathcal{P}_m = \bigcup_{t_1, \ldots, t_m} \mathcal{P}(t_1, \ldots, t_m).$$

Теорема 4. *Пусть* $F = \mathcal{L}(\xi) \in \mathfrak{F}_d$, $d = \infty$, $m \in \mathbf{N}$, $t_1, \ldots, t_m \in \mathbf{R}^d$, *и все распределения случайных величин* $\langle \xi, t_j \rangle$, $j = 1, \ldots, m$, *либо невырождены, либо равны* $E \in \mathfrak{F}_1$. *Тогда при всех* $P \in \mathcal{P}(t_1, \ldots, t_m)$,

$$\bigl|(F^n)\{P\} - (F^{n+1})\{P\}\bigr| \leqslant c(F, t_1, \ldots, t_m)\, n^{-1/2}. \tag{20}$$

Очевидно, что если распределение $F \in \mathfrak{F}_d$ сосредоточено на гиперплоскости, не содержащей нуля, то $\rho_m(F^n, F^{n+k}) = 1$ для любых $m, n, k \in \mathbf{N}$. Таким образом, имеет место альтернатива: левая часть (20) либо равна единице, либо убывает, по крайней мере как $O(n^{-1/2})$.

В заключение сформулируем вытекающий из теоремы 3 результат о близости распределений сумм случайного числа независимых одинаково распределенных случайных векторов. Для расстояния $\rho(\,\cdot\,,\,\cdot\,)$ этот результат содержится в работе [10, теорема 1.3].

Пусть $\xi_1, \xi_2, \ldots$ – независимые одинаково распределенные случайные векторы с общим распределением $F \in \mathfrak{F}_d$ и пусть $(\mu, \nu) \in \mathbf{Z}^2$ – двумерный случайный вектор с целочисленными неотрицательными координатами, не зависящий от последовательности $\{\xi_j\}_{j=1}^{\infty}$. Обозначим

$$U = \mathcal{L}(\mu), \;\; V = \mathcal{L}(\nu), \;\; G = \mathcal{L}(\xi_1 + \cdots + \xi_\mu), \;\; H = \mathcal{L}(\xi_1 + \cdots + \xi_\nu). \tag{21}$$

Хорошо известно, что тогда

$$G = \sum_{k=0}^{\infty} \mathbf{P}\{\mu = k\}\, F^k, \;\; H = \sum_{k=0}^{\infty} \mathbf{P}\{\nu = k\}\, F^k. \tag{22}$$

Теорема 5. *Пусть $d = \infty$. Если $F \in \mathfrak{F}_d^s$, то*

$$\rho_m(G, H) \leqslant \inf \mathbf{E} \,\min\Big\{\frac{cm}{\sqrt{\nu+1}} + c(m)\, \frac{|\mu - \nu|}{\nu + 1}, 1\Big\}, \tag{23}$$

а если $F \in \mathfrak{F}_d^+$, то

$$\rho_m(G, H) \leqslant \inf \mathbf{E} \,\min\Big\{c(m)\, \frac{|\mu - \nu|}{\nu + 1}, 1\Big\}. \tag{24}$$

Здесь инфимум берется по всевозможным двумерным распределениям $\mathcal{L}((\mu, \nu)) \in \mathfrak{F}_2$, таким что $\mathcal{L}(\mu) = U$, $\mathcal{L}(\nu) = V$.

Доказательство теорем 1–5. Прежде всего заметим, что теоремы 1–5 при $d < \infty$ были доказаны в [4].

Зафиксируем некоторый многогранник $P \in \mathcal{P}_m$:

$$P = \big\{x \in \mathbf{H} : \langle x, t_j\rangle \leqslant b_j, \; j = 1, \ldots, m\big\},$$

где $t_j \in \mathbf{H}$, $\|t_j\| = 1$, $b_j \in \mathbf{R}$, $j = 1, \ldots, m$. Пусть $\mathbf{L}_t \subset \mathbf{H}$ – линейная оболочка векторов

$$\big\{t_j, j = 1, \ldots, m\big\}, \quad k = \dim \mathbf{L}_t \leqslant m,$$

и пусть $\mathbb{P}_t : \mathbf{H} \to \mathbf{L}_t$ – оператор ортогонального проецирования на подпространство $\mathbf{L}_t$. Рассмотрим многогранник $\overline{P} \subset \mathbf{L}_t$, определяемый равенством

$$\overline{P} = \big\{x \in \mathbf{L}_t : \langle x, t_j\rangle \leqslant b_j, \; j = 1, \ldots, m\big\}.$$

Легко видеть, что для любого случайного вектора $\zeta \in \mathbf{H}$ мы имеем

$$\langle \mathbb{P}_t \zeta, t_j\rangle = \langle \zeta, t_j\rangle, \quad j = 1, \ldots, m.$$

Поэтому

$$\mathbf{P}\{\zeta \in P\} = \mathbf{P}\{\mathbb{P}_t \zeta \in \overline{P}\}. \tag{25}$$

Распределения k-мерных векторов $\mathbb{P}_t\xi$, $\mathbb{P}_t\eta$ удовлетворяют таким же k-мерным условиям, каким удовлетворяют сравниваемые распределения векторов $\xi, \eta \in \mathbf{H}$. Например, если $\mathcal{L}(\xi) \in \mathfrak{F}_\infty^{(\alpha)}$ для некоторого α, удовлетворяющего $0 \leqslant \alpha \leqslant 2$, то $\mathcal{L}(\mathbb{P}_t\xi) \in \mathfrak{F}_k^{(\alpha)}$. Аналогично, если $\mathcal{L}(\xi) \in \mathfrak{F}_\infty^s$, то $\mathcal{L}(\mathbb{P}_t\xi) \in \mathfrak{F}_k^s$ и так далее. Остается применить k -мерные варианты теорем 1–5. $\square$

Интересной проблемой является распространение наших неравенств на произвольные выпуклые множества X. Это нельзя вывести из теорем 1–5, так как константы $c(m)$ зависят от m.

Список литературы

1. Т. В. Арак, *О сближении n-кратных сверток распределений, имеющих неотрицательную характеристическую функцию, с сопровождающими законами.* — Теория вероятн. и ее примен. **25**, No. 2 (1980), 225–246.
2. Т. В. Арак, А. Ю. Зайцев, *Равномерные предельные теоремы для сумм независимых случайных величин.* — Тр. МИАН СССР **174** (1986), 214 с.
3. Ф. Гётце, А. Ю. Зайцев, *Редкие события и пуассоновские точечные процессы.* — Зап. научн. семин. ПОМИ **466** (2017), 109–119.
4. Ф. Гётце, А. Ю. Зайцев, *Оценки близости сверток вероятностных распределений на выпуклых многогранниках.* — Зап. научн. семин. ПОМИ **474** (2018), 108–117.
5. Ф. Гётце, А. Ю. Зайцев, *Об альтернативных аппроксимирующих распределениях в многомерном варианте второй равномерной предельной теоремы Колмогорова* . — Теория вероятн. и ее примен. **66** (2022) (в печати).
6. А. Ю. Зайцев, *Оценка близости распределений последовательных сумм независимых одинаково распределенных случайных векторов.* — Зап. научн. семин. ЛОМИ **97** (1980), 83–87.
7. А. Ю. Зайцев, *Некоторые свойства n-кратных сверток распределений.* — Теория вероятн. и ее примен. **26**, No. 1 (1981), 152–156.
8. А. Ю. Зайцев, *О точности аппроксимации распределений сумм независимых случайных величин, отличных от нуля с малой вероятностью, с помощью сопровождающих законов.* — Теория вероятн. и ее примен. **28**, No. 2 (1983), 625–636.
9. А. Ю. Зайцев, *К многомерному обобщению метода треугольных функций.* — Зап. научн. семин. ЛОМИ **158** (1987), 81–104.
10. A. Yu. Zaitsev, *Estimates for the closeness of successive convolutions of multidimensional symmetric distributions.* — Probab. Theory Relat. Fields **79**, No. 2 (1988), 175–200.
11. А. Ю. Зайцев, *Многомерный вариант второй равномерной предельной теоремы Колмогорова.* — Теория вероятн. и ее примен. **34**, No. 1 (1989), 128–151.
12. А. Ю. Зайцев, *Об аппроксимации сверток многомерных симметричных распределений сопровождающими законами.* — Зап. научн. семин. ЛОМИ **177** (1989), 55–72.

13. А. Ю. Зайцев, *Об одном классе неравномерных оценок в многомерных предельных теоремах.* — Зап. научн. семин. ПОМИ **184** (1990), 92–105.
14. А. Ю. Зайцев, *Об аппроксимации выборки пуассоновским точечным процессом.* — Зап. научн. семин. ПОМИ **298** (2003), 111–125.

Götze F., Zaitsev A. Yu. Convergence to infinite-dimensional compound Poisson distributions on convex polyhedra.

The aim of the present work is to provide a supplement to the authors' paper [4]. It is shown that our results on the approximation of distributions of sums of independent summands by the accompanying compound Poisson laws and the estimates of the proximity of sequential convolutions of multidimensional distributions on convex polyhedra may be almost automatically transferred to the infinite-dimensional case.

Fakultät für Mathematik,
Universität Bielefeld, Postfach 100131,
D-33501 Bielefeld, Germany
E-mail: goetze@math.uni-bielefeld.de

Поступило 27 июля 2021 г.

Санкт-Петербургское отделение
Математического института им. В. А. Стеклова
Фонтанка 27
Санкт-Петербург 191023, Россия
и Санкт-Петербургский государственный университет,
Университетская наб. 7/9
Санкт-Петербург, 199034 Россия
E-mail: zaitsev@pdmi.ras.ru