Записки научных
семинаров ПОМИ
Том 474, 2018 г.

Я. С. Голикова

ОБ УЛУЧШЕНИИ ОЦЕНКИ РАССТОЯНИЯ МЕЖДУ РАСПРЕДЕЛЕНИЯМИ ПОСЛЕДОВАТЕЛЬНЫХ СУММ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

В настоящей работе приведена верхняя оценка абсолютной константы для частного случая неравенства Колмогорова–Рогозина, на основе которой улучшена абсолютная постоянная в оценке близости n и $(n+1)$-кратных сверток одномерных вероятностных распределений.

Пусть $\xi_1,\ \xi_2,\ \ldots,\ \xi_n,\ \ldots$ – независимые случайные величины, F_k – функции распределения случайных величин ξ_k, а $\tilde{F}_k$ – функция распределения симметризованной случайной величины $\tilde{\xi}_k$.

В работе [1] С. В. Нагаевым и С. С. Ходжабагяном были получены верхние оценки функции концентрации для свертки распределений $Q\Big(\prod_{k=1}^n F_k,\ \tau\Big)$, где $Q(F,\ \tau) = \sup\limits_{x\in\mathbb{R}} F\{[x, x+\tau]\}$. Здесь и далее произведения и степени распределений понимаются в смысле свертки.

Пусть l_k, m_k $(k = 1, \ldots, n)$ – любые положительные числа. Положим

$$B_n^2 = \sum_{k=1}^n \sigma_k^2(l_k) = \sum_{k=1}^n \int\limits_{|x|\leqslant l_k} x^2\, d\tilde{F}_k(x), \tag{1}$$

$$\gamma_n = \sum_{k=1}^n \gamma_k(l_k) = \sum_{k=1}^n \int\limits_{|x|\leqslant l_k} |x|^3\, d\tilde{F}_k(x). \tag{2}$$

Теорема 1. *Для любого неотрицательного τ справедливы оценки*

$$Q\Big(\prod_{k=1}^n F_k,\ \tau\Big) \leqslant \min\Big\{c_1\,\frac{\tau}{B_n} + c_2\,\frac{\gamma_n}{B_n^3},\quad c_3\,\Big(\sum_{k=1}^n \Big(1+\frac{\tau\sqrt{2\pi}}{m_k}\Big)^{-2} \mathbf{P}(|\tilde{\xi}| > m_k)\Big)^{-1/2}\Big\}. \tag{3}$$

Ключевые слова: равномерное расстояние, суммы независимых случайных величин, функции концентрации, неравенства, оценка абсолютной постоянной.

Работа выполнена при частичной финансовой поддержке гранта СПбГУ–ННИО 6.37.65.2017.

$$Q\Big(\prod_{k=1}^{n} F_k,\, \tau\Big) \leqslant \Big(\sum_{k=1}^{n} \Big(c_1\,\frac{\tau}{\sigma_k(l_k)} + c_2\,\frac{\gamma_k(l_k)}{\sigma_n^3(l_k)}\Big)^{-2} + \Big(c_3\,\Big(1+\frac{\tau\sqrt{2\pi}}{l_k}\Big)\Big)^{-2}\,\mathbf{P}(|\tilde{\xi}| > l_k)\Big)^{-1/2}, \tag{4}$$

где $c_4 \leqslant 2.128815$, $c_2 \leqslant 6.6383125$, $c_3 \leqslant e^{-3/8} \approx 0.687289$.

Теорема 2. *Для любого неотрицательного* τ *справедливы оценки*

$$Q\Big(\prod_{k=1}^{n} F_k,\, \tau\Big) \leqslant \Big(\sum_{k=1}^{n} (c_4\tau + c_5 l_k)^{-2}\,\sigma_k^2(l_k) + \Big(c_3\,\Big(1+\frac{\tau\sqrt{2\pi}}{l_k}\Big)\Big)^{-2}\,\mathbf{P}(|\tilde{\xi}| > l_k)\Big)^{-1/2}, \tag{5}$$

где $c_4 \leqslant 2.34584$, $c_5 \leqslant 4.86119$.

В контексте настоящей статьи основное значение имеет то, что все константы c_i в теоремах 1, 2 лучше констант, которые получил Г. Зигель в своей работе [2].

В случае одинаково распределенных случайных величин ξ_k с распределением F из теоремы 2 путем несложных преобразований неравенства (5) в качестве следствия можно получить частный случай неравенства Колмогорова–Рогозина. Заметим, что при этом в неравенстве значительно улучшится константа.

Следствие 1. *Пусть* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *– независимые одинаково распределенные случайные величины с распределением* F*. Тогда*

$$Q(F^n,\, \tau) \leqslant \frac{c_0}{\sqrt{n\,(1 - Q(F,\, \tau))}}, \tag{6}$$

где $c_0 = \frac{1+2\sqrt{2\pi}}{e^{3/8}} \approx 4.132847$.

Для доказательства этого следствия достаточно применить теорему 2 для независимых одинаково распределенных случайных величин и положить $l = \frac{\tau}{2}$. В таком случае, B_n и γ_n из (1), (2) можно представить в виде:

$$B_n^2 = n\,\sigma_k^2(l) = n \int\limits_{|x|\leqslant l} x^2\,d\tilde{F}(x),$$

$$\gamma_n = n\,\gamma(l) = n \int\limits_{|x|\leqslant l} |x|^3\,d\tilde{F}(x).$$

Подставив получившиеся выражения в (5), получим:

$$\begin{aligned}
Q(F^n,\,\tau) &\leqslant \Big(n\Big(\big(c_4\tau+\frac{c_5\tau}{2}\big)^{-2}\sigma^2\big(\frac{\tau}{2}\big)+(c_3(1+2\sqrt{2\pi}))^{-2}\mathbf{P}\big(|\tilde{\xi}|>\frac{\tau}{2}\big)\Big)\Big)^{-1/2}\\
&= \frac{1}{\sqrt{n}}\Big(\frac{\sigma^2(\frac{\tau}{2})}{(c_4+\frac{c_5}{2})^2\tau^2}+\frac{\mathbf{P}(|\tilde{\xi}|>\frac{\tau}{2})}{(c_3(1+2\sqrt{2\pi}))^2}\Big)^{-1/2}\\
&= \frac{1}{\sqrt{n}}\Big(\frac{(c_3(1+2\sqrt{2\pi}))^2\sigma^2(\frac{\tau}{2})+(c_4+\frac{c_5}{2})^2\tau^2\mathbf{P}(|\tilde{\xi}|>\frac{\tau}{2})}{(c_4+\frac{c_5}{2})^2\tau^2\,(c_3(1+2\sqrt{2\pi}))^2}\Big)^{-1/2}\\
&= \frac{(c_4+\frac{c_5}{2})\,c_3\,(1+2\sqrt{2\pi})\tau}{\sqrt{n((c_3(1+2\sqrt{2\pi}))^2\sigma^2(\frac{\tau}{2})+(c_4+\frac{c_5}{2})^2\tau^2\mathbf{P}(|\tilde{\xi}|>\frac{\tau}{2}))}}\\
&\leqslant \frac{(c_4+\frac{c_5}{2})\,c_3\,(1+2\sqrt{2\pi})\tau}{\sqrt{n(c_4+\frac{c_5}{2})^2\tau^2(1-Q(F,\,\tau))}} = \frac{c_3\,(1+2\sqrt{2\pi})}{\sqrt{n(1-Q(F,\,\tau))}}.
\end{aligned}$$

Последнее неравенство верно, так как слагаемое $(c_3(1+2\sqrt{2\pi}))^2\sigma^2(\tau)$ в знаменателе неотрицательно, а $\mathbf{P}\big(|\tilde{\xi}|>\frac{\tau}{2}\big)\geqslant 1-Q(F,\,\tau)$. Аналогичным образом можно получить следствие 1 и из неравенств (3), (4).

Получившийся результат позволяет существенно улучшить оценку абсолютной постоянной в оценке близости распределений последовательности сумм независимых одинаково распределенных случайных величин из статьи Е. Л. Майстренко [4].

Пусть $\xi_1,\,\xi_2,\,\ldots,\,\xi_n,\,\ldots$ – независимые одинаково распределенные случайные величины с распределением F, 0 – q-квантиль распределения F, то есть выполнены неравенства: $F\{(-\infty,0)\}\leqslant q$ и $F\{(0,\infty)\}\leqslant 1-q$, где $0\leqslant q\leqslant 1$. В работе [3] приведена следующая оценка равномерного расстояния:

$$\rho(F^n,F^{n+1})\leqslant\frac{c}{\sqrt{n\min\{q,1-q\}}}\leqslant\frac{c}{\sqrt{nq(1-q)}},\tag{7}$$

где $\rho(F,G) = \sup_{x\in\mathbb{R}} |F(x) - G(x)|$. Зависимость от q в данном неравенстве является правильной, так как справедлива аналогичная нижняя оценка, то есть оценка (7) является оптимальной [3, с. 111]. Указанная оценка основана на частном случае неравенства Колмогорова–Рогозина (6). Изначально неравенство (7) было получено для многомерного случая [5], также частный случай указанного неравенства для $q = \frac{1}{2}$ представлен в работе [6].

Утверждение 1. *В неравенстве* (7) *абсолютную постоянную с можно взять равной* $c_0 = \frac{1+2\sqrt{2\pi}}{e^{3/8}} \approx 4.132847$.

Для доказательства утверждения обратимся к доказательству неравенства (7). В нем рассматривается разность $|F^n(x) - F^{n+1}(x)|$ в предположении (с последующим избавлением от него), что $F = F_1\Phi_\varepsilon$, где Φ_ε – нормальное распределение с математическим ожиданием 0 и дисперсией $\varepsilon > 0$, F_1 – некоторое распределение, и, применяя формулу свертки к $F^{n+1}(x)$ и неравенство (6), получаем следующее соотношение:

$$\begin{aligned}
&\big|F^n(x) - F^{n+1}(x)\big| = \Big|F^n(x) - \int\limits_{\mathbb{R}} F^n(x-y)F\{dy\}\Big| \\
&< \max\Big\{ \int\limits_{y\leqslant 0} \big|F^n(x-y) - F^n(x)\big| F\{dy\}, \int\limits_{y>0} \big|F^n(x-y) - F^n(x)\big| F\{dy\}\Big\} \\
&< \max\Big\{ \int\limits_{y\leqslant 0} Q(F^n, |y|)F\{dy\}, \int\limits_{y>0} Q(F^n, y)F\{dy\}\Big\} \\
&< c_0 \max\Big\{ \int\limits_{y\leqslant 0} \frac{F\{dy\}}{\sqrt{n(1-Q(F,|y|))}}, \int\limits_{y>0} \frac{F\{dy\}}{\sqrt{n(1-Q(F,y))}}\Big\} \\
&< \frac{c_0}{\sqrt{n}} \max\Big\{ \int\limits_{y\leqslant 0} \frac{F\{dy\}}{\sqrt{1-Q(F,|y|)}}, \int\limits_{y>0} \frac{F\{dy\}}{\sqrt{1-Q(F,y)}}\Big\}. \qquad (8)
\end{aligned}$$

Затем предполагается, что $q \leqslant \frac{1}{2}$. Тогда:

$$\int\limits_{y\leqslant 0} \frac{F\{dy\}}{\sqrt{1-Q(F,|y|)}} \leqslant \int\limits_{y\leqslant 0} \frac{dF(y)}{\sqrt{F(y)}},$$

$$\int\limits_{y>0} \frac{F\{dy\}}{\sqrt{1-Q(F,\,y)}} \leqslant \int\limits_{q<F(y)<1-q} \frac{dF(y)}{\sqrt{q}} + \int\limits_{F(y)>1-q} \frac{dF(y)}{\sqrt{1-F(y)}}.$$

Произведем замену $x = F(y)$, тогда (8) можно представить в виде:

$$\begin{aligned} &\frac{c_0}{\sqrt{n}} \max\Bigl\{ \int\limits_{y\leqslant 0} \frac{F\{dy\}}{\sqrt{1-Q(F,\,|y|)}}, \int\limits_{y>0} \frac{F\{dy\}}{\sqrt{1-Q(F,\,y)}} \Bigr\} \\ &\leqslant \frac{c_0}{\sqrt{n}} \max\Bigl\{ \int\limits_0^q \frac{dx}{\sqrt{x}}, \int\limits_q^{1-q} \frac{dx}{\sqrt{q}} + \int\limits_{1-q}^1 \frac{dx}{\sqrt{1-x}} \Bigr\} \\ &= \frac{c_0}{\sqrt{n}} \max\Bigl\{ 2\sqrt{q}, \frac{1}{\sqrt{q}} \Bigr\} = \frac{c_0}{\sqrt{nq}}. \end{aligned}$$

В случае $q > \frac{1}{2}$ можно рассмотреть распределение случайной величины $-\xi$. Для него 0 будет являться $(1-q)$-квантилью, тогда:

$$\frac{c_0}{\sqrt{n}} \max\Bigl\{ \int\limits_{y\leqslant 0} \frac{F\{dy\}}{\sqrt{1-Q(F,\,|y|)}}, \int\limits_{y>0} \frac{F\{dy\}}{\sqrt{1-Q(F,\,y)}} \Bigr\} \leqslant \frac{c_0}{\sqrt{n(1-q)}}.$$

Если подставить получившиеся неравенства в (7), то получим утверждение 1.

В работе [4] в качестве абсолютной постоянной бралась $c \approx 8.1320$. Таким образом, в настоящей работе удалось значительно улучшить оценку постоянной, что позволит эффективнее использовать неравенство (7). В случае симметричных распределений неравенство (7) начинает быть информативным уже при $n = 35$ против $n = 133$. При этом нельзя не отметить, что в монографии [3] для симметричных распределений с отделенной от -1 характеристической функцией для $\rho(F^n, F^{n+1})$ приведена принципиально лучшая оценка порядка $O\bigl(\frac{1}{n}\bigr)$. Однако вычисление постоянной в этом случае достаточно трудоемко, и, по всей видимости, она достаточно велика. Поэтому может оказаться, что оценка (7) позволяет получить лучший результат вплоть до очень больших n.

Список литературы

1. С. В. Нагаев, С. С. Ходжабагян, *Об оценке функции концентрации сумм независимых случайных величин.* — Теория вероятн. и ее примен. **41**, No. 3 (1996), 655–665.

2. Г. Зигель, *Верхние оценки для функции концентрации в гильбертовом пространстве.* — Теория вероятн. и ее примен. **26**, No. 2 (1981), 335–349.
3. Т. В. Арак, А. Ю. Зайцев, *Равномерные предельные теоремы для сумм независимых случайных величин.* — Тр. МИАН СССР **174** (1986), 3–214.
4. Е. Л. Майстренко, *Оценка абсолютной постоянной в неравенстве для равномерного расстояния между распределениями последовательных сумм независимых случайных величин.* — Зап. научн. сем. ПОМИ **454** (2016), 216–219.
5. А. Ю. Зайцев, *Оценка близости распределений последовательных сумм независимых одинаково распределенных случайных векторов.* — Зап. научн. семин. ЛОМИ **97** (1980), 83–87.
6. А. Ю. Зайцев, *Некоторые свойства n-кратных сверток распределений.* — Теория вероятн. и ее примен. **26**, No. 1 (1981), 152–156.

Golikova Ia. S. On improvement of the estimate of the distance between sequential sums of independent random variables.

The aim of the present paper is to improve the previously obtained estimate of the constant in the inequality for the uniform distance between n and $(n+1)$-fold convolution of one-dimensional probability distributions in the case where distribution F has 0 as q-quantile.

С.-Петербургский
государственный университет,
Университетская наб. 7/9
С.-Петербург, 199034 Россия;
Балтийский государственный
технический университет
"ВОЕНМЕХ" им. Д. Ф. Устинова,
1-я Красноармейская, д.1
С.-Петербург, 190005 Россия
E-mail: `laviniaspb@gmail.com`

Поступило 28 ноября 2018 г.