ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

УДК 519.2

УЛУЧШЕНИЕ НЕРАВЕНСТВА МАССАРТА ДЛЯ СТАТИСТИКИ СМИРНОВА

Ташков И.А.
МГУ им. М.В. Ломоносова, г. Москва

Поступила в редакцию 10.12.2017, после переработки 25.12.2017.

Обозначим через F_n эмпирическую функцию распределения по выборке из n независимых одинаково распределенных случайных величин с функцией распределения F. В данной работе доказано неравенство

$$\mathbb{P}\{\sqrt{n} \sup_{|x|<\infty} (F_n(x) - F(x)) > \lambda\} \leq \exp\{-2\lambda^2 - \lambda^4/36n\}$$

для $n \geq 39, \min\{\gamma n^{-1/6}, \sqrt{\ln 2/2}\} \leq \lambda \leq \sqrt{n}/2, \gamma = 1.0841$. Кроме того, доказано для тех же n и $\lambda \leq \sqrt{n}/2$, что

$$\mathbb{P}\{\sqrt{n} \sup_{|x|<\infty} (F_n(x) - F(x)) > \lambda\} \leq 2\exp^{(\ln 2)^2/(144n)} \exp\{-2\lambda^2 - \lambda^4/36n\}.$$

Для частных случаев $n = 2, 3, 4$ доказана более сильная оценка

$$\mathbb{P}\{\sqrt{n} \sup_{|x|<\infty} (F_n(x) - F(x)) > \lambda\} \leq \exp\{-2\lambda^2 - 4\lambda^4/9n\}.$$

Ключевые слова: распределение статистики Смирнова, экспоненциальные неравенства.

Вестник ТвГУ. Серия: Прикладная математика. 2018. № 1. С. 5–20.
https://doi.org/10.26456/vtpmk489

Введение

Пусть независимые одинаково распределенные случайные величины $X_n, n \in \mathbb{N} = \{1, 2, \dots\}$, определены на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Предположим, что их функция распределения $F(x) = \mathbb{P}\{X_1 \leq x\}, x \in \mathbb{R} = (-\infty, \infty)$ непрерывна. Напомним определение эмпирической функции распределения

$$F_n(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{I}_{\{X_k \leq x\}}, x \in \mathbb{R} = (-\infty, \infty),$$

где $\mathbb{I}_A$ – индикаторная функция множества A. Смирнов [1] показал, что распределение статистики $D_n^+ = \sqrt{n}\sup_{x\in\mathbb{R}}(F_n(x) - F(x))$ имеет вид

$$\begin{gathered}\mathbb{P}\{D_n^+ > \lambda\} = 1, \lambda < 0; \mathbb{P}\{D_n^+ > \lambda\} = 0, \lambda \geq \sqrt{n},\\ \mathbb{P}\{D_n^+ > \lambda\} = q_{\lambda,n}(0) + \lambda\sqrt{n}\sum_{k=[\lambda\sqrt{n}]+1}^{n-1} q_{\lambda,n}(k), 0 \leq \lambda < \sqrt{n},\\ q_{\lambda,n}(0) = (1 - \frac{\lambda}{\sqrt{n}})^n, q_{\lambda,n}(k) = \frac{1}{n^n}\binom{n}{k}(k - \lambda\sqrt{n})^k(n - k + \lambda\sqrt{n})^{n-k-1}.\end{gathered} \tag{1}$$

Символ $[\lambda\sqrt{n}]$ обозначает целую часть числа $\lambda\sqrt{n}$.

В статье Дворецкого, Кифера и Волфовица [2] доказано неравенство

$$\mathbb{P}\{D_n^+ \geq \lambda\} \leq ce^{-2\lambda^2}$$

с неопределенной постоянной c. Краткую историю вычислений данной константы можно найти в статье Массарта [3]. Там же приведено доказательство неравенства

$$\mathbb{P}\{D_n^+ \geq \lambda\} \leq e^{-2\lambda^2} \text{ для } \lambda \geq \min\{\gamma n^{-1/6}, \sqrt{\ln 2/2}\},$$

где $\gamma = 1.0841$. В данной статье доказано, что

$$\mathbb{P}\{D_n^+ \geq \lambda\} \leq \exp\Big\{-2\lambda^2 - \frac{1}{36n}\lambda^4\Big\}, \min\{\gamma n^{-1/6}, \sqrt{\ln 2/2}\} \leq \lambda \leq \sqrt{n}/2, n \geq 39. \tag{2}$$

Заметим, что $e^{-2\lambda^2-\lambda^4/36n} > 2^{-1}e^{-(\ln 2)^2/(144n)}$ для $0 \leq \lambda \leq \sqrt{\ln 2/2}$. Отсюда и из (2) следует неравенство

$$\mathbb{P}\{D_n^+ > \lambda\} \leq 1 \leq 2e^{(\ln 2)^2/(144n)}e^{-2\lambda^2-\lambda^4/36n}, 0 \leq \lambda \leq \sqrt{n}/2, n \geq 39.$$

Для частных случаев $n = 2, 3, 4$ доказано более строгое неравенство

$$\mathbb{P}\{D_n^+ \geq \lambda\} \leq e^{-2\lambda^2-4\lambda^4/9n},\ \lambda \geq 0. \tag{3}$$

Доказательство неравенства (2) базируется на следующей теореме.

Теорема. *Пусть ξ — случайная величина, имеющая биномиальное распределение*

$$\mathbb{P}\{\xi = k\} = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k},\ 0 \leq p \leq 1,\ k = 0, \dots, n,\ n \in \mathbb{N}.$$

Следующее неравенство справедливо для любых $\varepsilon, 0 \leq \varepsilon \leq q = 1 - p$:

$$\mathbb{P}\{\xi - np > n\varepsilon\} \leq \exp\Big\{-\frac{\varepsilon^2}{2(p+\varepsilon/3)(q-\varepsilon/3)} - \frac{1}{36}\varepsilon^4\Big\}. \tag{4}$$

Доказательство (3) для $n = 2, 3, 4$ основано на использовании теоремы Будана-Фурье о количестве вещественных корней многочлена

$$P(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0, a_n \neq 0, x \in \mathbb{R} = (-\infty, \infty), \tag{5}$$

степени $n \in \mathbb{N}$ с вещественными коэффициентами.

Пусть даны вещественные числа $c_1, \dots, c_n$, среди которых некоторые, скажем, $c_{k_1}, \dots, c_{k_m}$, не равны нулю. Обозначим через $W(c_1, \dots, c_n)$ количество изменений знака в последовательности $c_{k_1}, \dots, c_{k_m}$, если $m > 1$, и $W(c_1, \dots, c_n) = 0$, если $m = 1$. Число $W(c_1, \dots, c_n)$ называется количеством изменений знака в последовательности чисел $c_1, \dots, c_n$.

Теорема Будана-Фурье. *Пусть даны многочлен (5) и его производные $P^{(k)}$, $k = 1, \dots, n$. Обозначим через $W(x)$ количество изменений знака в последовательности $P(x), P^{(1)}(x), \dots, P^{(n)}(x)$, $x \in \mathbb{R}$. Тогда для любых $a, b \in \mathbb{R}, a < b$ количество корней P на полуинтервале $(a, b]$ с учетом их кратностей равно $W(a) - W(b) - 2m$, где m — некое неотрицательное целое число.*

1. Основные результаты

Начнем с доказательства неравенства (4). Непосредственно видно, что неравенство справедливо для $\varepsilon = 0$. Если неравенство справедливо для $0 < \varepsilon < q$, то оно справедливо для $\varepsilon = q$. Далее предполагается, что $0 < \varepsilon < q$. Используя неравенство Маркова, получим

$$\mathbb{P}\{\xi - np > n\varepsilon\} = \mathbb{P}\{e^{x(\xi - np)} > e^{xn\varepsilon}\} \le e^{-xn\varepsilon}\mathbb{E}e^{x(\xi-np)} = e^{-xn\varepsilon}(pe^{xq} + qe^{-xp})^n, x \ge 0.$$

Положим $x = n\ln((p+\varepsilon)q/(q-\varepsilon)p)$. Нетрудно показать, что

$$\mathbb{P}\{\xi - np > n\varepsilon\} \le e^{-h(p,\varepsilon)},\ h(p,\epsilon) = (p+\varepsilon)\ln\left(\frac{p+\varepsilon}{p}\right) + (q-\varepsilon)\ln\left(\frac{q-\varepsilon}{q}\right). \quad (6)$$

Рассмотрим функции $h(p,\epsilon), 0 < \varepsilon < q < 1$, и

$$\phi(t) = t - \frac{t^2}{2(1+2t/3)} - \ln(1+t), t \ge 0.$$

Так как $\phi(0) = 0$ и

$$\phi'(t) = \frac{t^3}{9(1+2t/3)^2(1+t)} > 0$$

для любого положительного t, то функция $\phi(t), t \ge 0$, положительна, непрерывна, не убывает. Легко видеть, что

$$\lim_{t\to\infty}\frac{\phi(t)}{t} = \frac{1}{4}. \quad (7)$$

Далее докажем следующее неравенство

$$h(p,\varepsilon) \ge \frac{\varepsilon^2}{2(p+\varepsilon/3)(q-\varepsilon/3)} + \frac{1}{36}\varepsilon^4 + \frac{\varepsilon\phi(t)}{t},\ \text{где } t = \frac{\varepsilon}{q-\varepsilon}.$$

Для этого положим $x = p + \varepsilon$ и рассмотрим разность

$$\begin{aligned} h(p,\varepsilon) &- \frac{\varepsilon^2}{2(p+\varepsilon/3)(q-\varepsilon/3)} - \frac{1}{36}\varepsilon^4 - \frac{\varepsilon\phi(t)}{t} = \\ &= x\ln x - x\ln(x-\varepsilon) - \frac{1}{36}\varepsilon^4 - \frac{\varepsilon^2}{2(x-2\varepsilon/3)} - \varepsilon, 0 < \varepsilon < x < 1. \end{aligned} \quad (8)$$

Производная функции справа по переменной ε равна

$$\frac{\varepsilon^3}{9(x-\varepsilon)(x-2\varepsilon/3)^2}-\frac{1}{9}\varepsilon^3>\frac{\varepsilon^3}{9x^3}-\frac{1}{9}\varepsilon^3>0.$$

Следовательно функция слева (8) возрастает. Так как $\phi(t)\geq 0$ и $h(p,0)=0$, то

$$h(p,\varepsilon)\geq\frac{\varepsilon^2}{2(p+\varepsilon/3)(q-\varepsilon/3)}+\frac{1}{36}\varepsilon^4+\frac{\varepsilon\phi(t)}{t}. \tag{9}$$

Неравенство (4) следует (6) из (9). В действительности было доказано более строгое неравенство

$$\mathbb{P}\{\xi-np>n\varepsilon\}\leq\exp\Big(-\frac{\varepsilon^2}{2(p+\varepsilon/3)(q-\varepsilon/3)}-\frac{\varepsilon^4}{36}-\frac{\varepsilon\phi(t)}{t}\Big), t=\varepsilon/(q-\varepsilon).$$

Для доказательства неравенства (2) нам понадобится следующее равенство

$$e^{-2\lambda^2}=\int_0^1 f_\lambda(s)\,ds, f_\lambda(s)=\frac{\lambda}{\sqrt{2\pi}}\frac{1}{\sqrt{s^3(1-s)}}e^{-\lambda^2/2s(1-s)}. \tag{10}$$

Это равенство может быть получено в качестве следствия глубокого результата в теории стохастических процессов. Здесь предлагается его непосредственное доказательство. Для начала покажем, что функция

$$g(\lambda)=\frac{|\lambda|}{\pi}\int_0^1\frac{1}{\sqrt{s^3(1-s)}}e^{-\lambda^2/(2s(1-s))}\,ds, \lambda\in\mathbb{R},$$

является плотностью нормальной случайной величины ξ, с $\mathbb{E}\xi=0$ и $\mathbb{E}\xi^2=1/4$. По теореме Фуббини для любого целого $k\geq 0$ выполняется равенство

$$\int_{-\infty}^{\infty}|\lambda|^{2k}g(\lambda)\,d\lambda=\int_0^1\frac{2}{s\sqrt{2\pi}}\int_{-\infty}^{\infty}|\lambda|^{2k+1}\frac{1}{\sqrt{2\pi s(1-s)}}e^{-\lambda^2/(2s(1-s))}\,d\lambda\,ds.$$

Внутренний интеграл равен $\pi^{-1/2}2^{k+1/2}(s(1-s))^{k+1/2}k!$. Это следует из того, что он равен математическому ожиданию $\mathbb{E}|\eta|^{2k+1}$ нормально распределенной случайной величины η с $\mathbb{E}\eta=0$ и $\mathbb{E}\eta^2=s(1-s)$. Предыдущее равенство принимает следующий вид

$$\int_{-\infty}^{\infty}|\lambda|^{2k}g(\lambda)\,d\lambda=\frac{2^{k+1}k!}{\pi}\int_0^1 s^{k-1/2}(1-s)^{k+1/2}\,ds.$$

Интеграл в правой части есть B-функция $B(k+1/2,k+3/2)$. Он может быть записан с помощью Γ-функций:

$$\int_0^1 s^{k-1/2}(1-s)^{k+1/2}\,ds=B(k+1/2,k+3/2)=\frac{\Gamma(k+1/2)\Gamma(k+3/2)}{\Gamma(2k+2)}.$$

Так как $\Gamma(1/2)\Gamma(1/2)=\pi$ и $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$ для любого $\alpha>0$, то

$$\int_{-\infty}^{\infty}|\lambda|^{2k}g(\lambda)\,d\lambda=\frac{2^{k+1}}{\pi}\frac{\Gamma(k+1/2)\Gamma(k+3/2)}{\Gamma(2k+2)}=\begin{cases}1, & \text{при } k=0,\\ \frac{1}{4^k}(2k-1)!!, & \text{при } k\geq 1,\end{cases}$$

где $(2k-1)!! = 1 \cdot 3 \cdots (2k-1)$ является произведением нечетных натуральных чисел от 1 до $2k-1$. Из равенства для $k=0$ следует, что функция $g(\lambda), \lambda \in \mathbb{R}$, является симметричной функцией плотности. Равенства для $k \geq 1$ показывают, что все четные моменты случайной величины ξ с плотностью g совпадают с четными моментами нормальной случайной величины ξ' с $\mathbb{E}\xi' = 0$ и $\mathbb{E}|\xi'|^2 = 1/4$. Любая нормальная случайная величина определяется своими моментами. Поэтому справедливо равенство $g(\lambda) = \sqrt{2/\pi}\exp\{-2\lambda^2\}$ для любых $\lambda \in \mathbb{R}$ и, следовательно,

$$\int_0^1 f_\lambda(s)\,ds = \sqrt{\frac{\pi}{2}} g(\lambda) = \sqrt{\frac{\pi}{2}}\sqrt{\frac{2}{\pi}} e^{-2\lambda^2} = e^{-2\lambda^2},\ \lambda \geq 0.$$

Неравенство (10) доказано.

Доказательство (2) основано на идеях Массарта [3]. Обозначив $j = n-k$, выражение (1) для $0 < \lambda < \sqrt{n}$ может быть переписано в следующем виде

$$\mathbb{P}\{D_n^+ > \lambda\} = \sum_{0 \leq j < n - \lambda\sqrt{n}} p_{\lambda,n}(j),$$

$$p_{\lambda,n}(j) = \lambda\sqrt{n}(j + \lambda\sqrt{n})^{j-1}(n - j - \lambda\sqrt{n})^{n-j} n^{-n}\binom{n}{j}.$$

Лемма 1. *Пусть $v_n(s') = 1/(s'(s'^2 - 1/(4n)))$. Тогда при любых положительных $\varepsilon, s, s' = 1-s$ таких, что $n\varepsilon \geq 2$ и $ns' \geq 1$, справедливо следующее неравенство*

$$\Big(1 + \frac{2\varepsilon}{3s'}\Big)^{-\frac{1}{2}} \leq \Big(1 - \frac{\varepsilon}{3s'} + \frac{\varepsilon^2}{6s'^2}\Big)\exp\Big(-\frac{\varepsilon^2 v_n(s')}{24n}\Big). \tag{11}$$

Доказательство. Рассмотрим функцию $\rho(x) = 1 + \alpha x - \exp(2\beta x)$, где $\alpha = 135/32, \beta = 0.826$. Заметим, что $\rho(x)$ вогнута и неотрицательна на сегменте $[0,1]$, так как $\rho(0) = 0$ и $\rho(1) \geq 1.3 \cdot 10^{-3}$. Следовательно

$$\exp(2\beta x^3) \leq 1 + \alpha x^3 \leq 1 + \alpha x^3 + \frac{x^3(5-12x)^2}{32} = \Big(1 - x + \frac{3x^2}{2}\Big)^2(1+2x).$$

Используем неравенство выше с $x = \varepsilon/(3s')$. Тогда для доказательства (11) остается показать, что

$$\frac{\beta\varepsilon^3}{27s'^3} \geq \frac{\varepsilon^2 v_n(s')}{24n},$$

что равносильно

$$\frac{8\beta}{9} \geq (n\varepsilon(1 - (2ns')^{-2}))^{-1}.$$

Так как $n\varepsilon \geq 2$ и $ns' \geq 1$, то выполняются неравенства

$$(n\varepsilon(1 - (2ns')^{-2}))^{-1} \leq \frac{2}{3} < \frac{8\beta}{9}.$$

Это завершает доказательство леммы. □

Лемма 2. *Пусть* $0 \le j < n - \lambda\sqrt{n}, \varepsilon = \lambda/\sqrt{n}, s = 2\varepsilon/3 + j/n, s' = 1 - s$. *Если* $n\varepsilon \ge 2$, *то*

$$p_{\lambda,n}(j) \le \frac{1}{n}\Big(1 - \frac{\varepsilon}{3s'} + \frac{\varepsilon^2}{6s'^2}\Big) \exp\Big(\frac{0.4}{ns} - \frac{\varepsilon^2}{24n}(v_n(s) + v_n(s'))\Big) f_\lambda(s) e^{-\lambda^4/36n}, \qquad (12)$$

где $v_n(s) = 1/(s(s^2 - 1/(4n)))$.

Доказательство. Рассмотрим случай $j = 0$. В этом случае $s = 2\varepsilon/3$. Напомним, что $\varepsilon = \lambda/\sqrt{n}$ и $q = 1 - p$. Из (7) и (9) следует, что

$$-\ln(1-\varepsilon) = \lim_{q\to\varepsilon} h(p,\varepsilon) \ge \frac{\varepsilon^2}{2(1-2\varepsilon/3)2\varepsilon/3} + \frac{\varepsilon}{4} + \frac{\varepsilon^4}{36} = \frac{\varepsilon}{4} + \frac{\varepsilon^2}{2ss'} + \frac{\varepsilon^4}{36}.$$

Обозначим

$$H(\nu) = \frac{3\ln(3/2)}{2} - \frac{\ln(2\pi)}{2} + \frac{\nu}{4} + \frac{0.4}{\nu} - \frac{\ln\nu}{2}, \nu > 1.$$

Величину $p_{\lambda,n}(0)$ можно оценить следующим образом

$$p_{\lambda,n}(0) = (1-\varepsilon)^n = e^{n\ln(1-\varepsilon)} \le \exp\Big(-\frac{n\varepsilon}{4} - \frac{\lambda^2}{2ss'} - \frac{\lambda^4}{36n}\Big) \le$$
$$\le \frac{\lambda}{n\sqrt{2\pi}} s^{-3/2} \exp\Big(\frac{0.4}{n\varepsilon}\Big) \exp\Big(-\frac{\lambda^2}{2ss'}\Big) \exp\Big(-H(n\varepsilon)\Big) e^{-\lambda^4/36n}.$$

Функция $H(\nu), \nu > 1$ положительна. Она принимает минимальное значение в точке $\nu_0 = 1 + \sqrt{2.6}$ и $H(\nu_0) \ge 1.5 \cdot 10^{-2}$. Поэтому

$$p_{\lambda,n}(0) \le \frac{\lambda}{n\sqrt{2\pi}} s^{-3/2} \exp\Big(\frac{0.4}{n\varepsilon}\Big) \exp\Big(-\frac{\lambda^2}{2ss'}\Big) e^{-\lambda^4/36n} \le$$
$$\le \frac{1}{n}\Big(1 + \frac{2\varepsilon}{3s'}\Big)^{-1/2} \exp\Big(\frac{0.4}{n\varepsilon}\Big) f_\lambda(s) e^{-\lambda^4/36n}.$$

Так как $\varepsilon n \ge 2$, то справедливы неравенства

$$\frac{\varepsilon^2 v_n(s)}{24n} \le \frac{1}{16n\varepsilon(4/9 - 1/16)} \le \frac{9}{55n\varepsilon},$$
$$\frac{0.4}{n\varepsilon} + \frac{\varepsilon^2 v_n(s)}{24n} \le \frac{0.4 + 9/55}{n\varepsilon} \le \frac{0.4}{ns}.$$

Из этих неравенств и предыдущей оценки для $p_{\lambda,n}(0)$ следует, что

$$p_{\lambda,n}(0) \le \frac{1}{n}\Big(1 + \frac{2\varepsilon}{3s'}\Big)^{-1/2} \exp\Big(\frac{0.4}{ns} - \frac{\varepsilon^2 v_n(s)}{24n}\Big) f_\lambda(s) e^{-\lambda^4/36n}.$$

Это неравенство вместе с (11) влекут неравенство (12) для $j = 0$.

Предположим теперь, что $1 \le j < n - \lambda\sqrt{n}$. По формуле Стирлинга (см. [4], стр. 54) выполняются неравенства

$$\sqrt{2\pi} n^{n+1/2} e^{-n} e^{-1/(12n)} < n! < \sqrt{2\pi} n^{n+1/2} e^{-n} e^{-1/(12n+1)}.$$

В силу правого неравенства справедлива следующая оценка

$$\binom{n}{j} \le \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{j(n-j)}}\frac{n^n}{j^j(n-j)^{n-j}}C_j, C_j = e^{-1/(12j+1)}, j \ge 1.$$

Это, в свою очередь, ведет к оценке

$$p_{\lambda,n}(j) \le \frac{\lambda}{\sqrt{2\pi}}\frac{n}{\sqrt{j(n-j)}(j+\lambda\sqrt{n})}\Big(\frac{j+\lambda\sqrt{n}}{j}\Big)^j\Big(\frac{n-j-\lambda\sqrt{n}}{n-j}\Big)^{n-j}C_j.$$

Положим $\varepsilon = \lambda/\sqrt{n}, s_j = s = 2\varepsilon/3 + j/n$ и $s'_j = s' = 1 - s$. Предыдущее неравенство с помощью функции $h(\cdot,\cdot)$ из (6) преобразуется к следующему виду

$$p_{\lambda,n}(j) \le \frac{\lambda}{n\sqrt{2\pi}}\Big(s - \frac{2\varepsilon}{3}\Big)^{-1/2}\Big(s + \frac{\varepsilon}{3}\Big)^{-1}\Big(s' + \frac{2\varepsilon}{3}\Big)^{-1/2}e^{-nh(s'-\varepsilon/3,\varepsilon)}C_j.$$

Определим функцию

$$\psi(t) = -\ln(1+t) + \frac{3}{2}\ln(1+\frac{2t}{3}), t \ge 0.$$

Из неравенства (9) с $p = s' - \varepsilon/2$ следует, что

$$\begin{aligned} p_{\lambda,n}(j) &\le \frac{1}{n}\Big(1+\frac{2\varepsilon}{3s'}\Big)^{-1/2}\frac{s^{3/2}}{\sqrt{s-2\varepsilon/3}(s+\varepsilon/3)}C_j e^{-n\varepsilon\phi(t)/t}f_\lambda(s)e^{-\lambda^4/36n} = \\ &= \frac{1}{n}\Big(1+\frac{2\varepsilon}{3s'}\Big)^{-1/2}\frac{s^{3/2}}{\sqrt{s-2\varepsilon/3}(s+\varepsilon/3)}C_j e^{-n\varepsilon\phi(t)/t+\psi(t)}f_\lambda(s)e^{-\lambda^4/36n}. \end{aligned}$$

Это неравенство является более строгой версией неравенства (2.7) из [3]. Неравенство (2.7) из статьи Массарта получается путем замены $e^{-\lambda^4/36n}$ на единицу. Теперь для доказательства (12) при $1 \le j < n - \lambda\sqrt{n}$ можно пошагово повторить действия Массарта [3]. □

Докажем, что (2) следует из (12) для $n \ge 39, \gamma n^{-1/6} \le \lambda \le \sqrt{n}/2$. По аналогии с [3] получим

$$e^{2\lambda^2+\lambda^4/36n}\mathbb{P}\{D_m^+ > \lambda\} \le J_{0,1}(\lambda) - \frac{\varepsilon}{3}J_{1,1}(\lambda) + \frac{\varepsilon^2}{6}J_{2,1}(\lambda) + \frac{\mu}{n}J_{2,0}(\lambda) - \frac{2\mu}{3n}J_{2,1} + \frac{\varepsilon^2\mu}{6n}J_{2,2}(\lambda),$$

где $\mu = 0.4345$,

$$J_{a,b}(\lambda) = \frac{\lambda}{\sqrt{2\pi}}e^{2\lambda^2}\int_0^1 u^{-1/2-a}(1-u)^{-1/2-b}e^{-\lambda^2/(2u(1-u))}\,du.$$

Из (10) получим $1 = J_{1,0}(\lambda) = 1/2\cdot(J_{1,0}(\lambda) + J_{0,1}(\lambda)) = 1/2\cdot J_{1,1}(\lambda)$. С помощью (10) нетрудно доказать, что $J_{2,2}(\lambda)/2 = J_{2,1}(\lambda) = 4 + \lambda^{-2}, J_{2,0}(\lambda) = 2 + \lambda^{-2}$. Из предыдущего неравенства следует, что

$$\begin{aligned} &\frac{3\sqrt{n}}{2\lambda}\Big(e^{2\lambda^2+\lambda^4/36n}\mathbb{P}(D_n^+ > \lambda) - 1\Big) \le \\ &\le -1 + \Big(\lambda + \frac{1}{4\lambda} + \frac{3\mu}{\lambda} + \frac{3\mu}{2\lambda^3}\Big)n^{-1/2} - \frac{\mu}{2}\Big(4 + \frac{1}{\lambda^2}\Big)n^{-1} + \frac{\mu}{2}\Big(4\lambda + \frac{1}{\lambda}\Big)n^{-3/2}. \end{aligned}$$

Массарт [3] доказал, что правая часть выражения выше отрицательна. Это завершает доказательство (2).

2. Случай малых объемов выборок

Доказать (3) нам поможет следующая лемма.

Лемма 3. *Справедливо следующее неравенство*

$$-\ln(1-\varepsilon) \geq 2\varepsilon^2 + \frac{4}{9}\varepsilon^4, 0 \leq \varepsilon < 1. \tag{13}$$

Доказательство. Достаточно доказать, что функция $f(\varepsilon) = \ln(1-\varepsilon) + 2\varepsilon^2 + 4\varepsilon^4/9$ принимает неположительные значения для всех $\varepsilon \in [0,1)$. Производная данной функции равна

$$f^{(1)}(\varepsilon) = \frac{P(\varepsilon)}{1-\varepsilon}, P(\varepsilon) = -\frac{16}{9}\varepsilon^4 + \frac{16}{9}\varepsilon^3 - 4\varepsilon^2 + 4\varepsilon - 1.$$

Первая и вторая производная функции $P(\varepsilon), 0 \leq \varepsilon \leq 1$, имеют следующий вид

$$P^{(1)}(\varepsilon) = -\frac{64}{9}\varepsilon^3 + \frac{48}{9}\varepsilon^2 - 8\varepsilon + 4, P^{(2)}(\varepsilon) = -\frac{64}{3}\varepsilon^2 + \frac{32}{3}\varepsilon - 8.$$

Вторая производная $P^{(2)}(\varepsilon) = -(8\varepsilon-3)^2/3 - 5$ отрицательна. Следовательно $P^{(1)}$ убывает, и функция P вогнута. Т.к. $P(0) = P(1) = -1$ и $P(0.6) = 0.1136$, существуют $0 < \varepsilon_1 < \varepsilon_2 < 1$, такие, что функция $P(\varepsilon)$ отрицательна при $\varepsilon \in [0, \varepsilon_1) \cup (\varepsilon_2, 1]$ и положительна для $\varepsilon \in (\varepsilon_1, \varepsilon_2)$. Получим, что $f(\varepsilon)$ убывает на $[0, \varepsilon_1) \cup (\varepsilon_2, 1]$ и возрастает на $(\varepsilon_1, \varepsilon_2)$. Поэтому функция f принимает максимальное значение при $\varepsilon = 0$ или $\varepsilon = \varepsilon_2$. Заметим, что $f(0) = 0$. Покажем, что $f(\varepsilon_2) \leq 0$. Так как $P(14/20) = 43/1875 > 0$ и $P(15/20) = -0.0625 < 0$, получим, что $14/20 < \varepsilon_2 < 15/20$. По формуле Тэйлора существует $\varepsilon' \in [14/20, \varepsilon_2]$, что

$$f(\varepsilon_2) = f(14/20) + \frac{P(\varepsilon')}{1-\varepsilon'}(\varepsilon_2 - 14/20) \leq f(14/20) + \frac{P(14/20)}{1-15/20}(15/20 - 14/20) < 0.$$

Это доказывает, что f отрицательна. Неравенство (13) доказано. □

Утверждение 1. *Неравенство (3) справедливо при* $n = 2$.

Доказательство. Из (1) и (13) для $(n-1)/\sqrt{n} \leq \lambda < \sqrt{n}, n \geq 1$, следует, что

$$\mathbb{P}\{D_n^+ > \lambda\} = \left(1 - \frac{\lambda}{\sqrt{n}}\right)^n \leq e^{-2\lambda^2 - 4\lambda^4/9}. \tag{14}$$

Осталось рассмотреть промежуток $0 < \lambda < 1/\sqrt{2}$. Обозначим $f_2(\lambda) = \mathbb{P}\{D_2^+ > \lambda\}$. Достаточно доказать, что

$$\left(\ln(f_2(\lambda)e^{2\lambda^2 + 2\lambda^4/9})\right)' = \frac{f_2'(\lambda) + (4\lambda + 8\lambda^3/9)f_2(\lambda)}{f_2(\lambda)} \leq 0. \tag{15}$$

При $0 < \lambda < 1/\sqrt{2}$ получим

$$f_2(\lambda) = -\frac{1}{2}\lambda^2 - \frac{1}{\sqrt{2}}\lambda + 1; f_2'(\lambda) = -\lambda - \frac{1}{\sqrt{2}};$$

$$f_2'(\lambda) + (4\lambda + \frac{8}{9}\lambda^3)f_2(\lambda) = -\frac{4}{9}\lambda^5 - \frac{4\sqrt{2}}{9}\lambda^4 - \frac{10}{9}\lambda^3 - 2\sqrt{2}\lambda^2 + 3\lambda - \frac{1}{\sqrt{2}} \leq$$
$$\leq -\frac{10}{9}\lambda^3 - 2\sqrt{2}\lambda^2 + 3\lambda - \frac{1}{\sqrt{2}}.$$

Рассмотрим многочлен $P_2(\lambda) = -10\lambda^3/9 - 2\sqrt{2}\lambda^2 + 3\lambda - 1/\sqrt{2}, \lambda \in \mathbb{R}$, и его производную

$$P_2^{(1)}(\lambda) = -\frac{10}{3}\lambda^2 - 4\sqrt{2}\lambda + 3 = -\frac{10}{3}(\lambda + \frac{3}{\sqrt{2}})(\lambda - \frac{3}{5\sqrt{2}}).$$

Функция $P_2(\lambda), 0 \leq \lambda \leq 1/\sqrt{2}$, принимает максимальное значение в точке $\lambda = 3/5\sqrt{2}$. Так как $P_2(3/5\sqrt{2}) = -1/25\sqrt{2} < 0$, то $P_2(\lambda) \leq 0$ для всех $0 \leq \lambda \leq 1/\sqrt{2}$. Следовательно, неравенство (14) справедливо. Утверждение доказано. □

Утверждение 2. *Неравенство (3) справедливо при* $n = 3$.

Доказательство. В силу (14) достаточно доказать неравенство только для $0 < \lambda < 2/\sqrt{3}$. Обозначим $f_3(\lambda) = \mathbb{P}\{D_3^+ > \lambda\}$ и рассмотрим следующую функцию

$$\Big(\ln(f_3(\lambda)e^{2\lambda^2+4\lambda^4/27})\Big)' = \frac{f_3'(\lambda) + (4\lambda + 16\lambda^3/27)f_3(\lambda)}{f_3(\lambda)}. \tag{16}$$

При $0 < \lambda < 1/\sqrt{3}$ имеем

$$f_3(\lambda) = -\frac{1}{3\sqrt{3}}\lambda^3 - \frac{2}{3}\lambda^2 - \frac{1}{\sqrt{3}}\lambda + 1; f_3'(\lambda) = -\frac{1}{\sqrt{3}}\lambda^2 - \frac{4}{3}\lambda - \frac{1}{\sqrt{3}};$$

$$f_3'(\lambda) + (4\lambda + \frac{16\lambda^3}{27})f_3(\lambda) = -\frac{16}{81\sqrt{3}}\lambda^6 - \frac{32}{81}\lambda^5 - \frac{52}{27\sqrt{3}}\lambda^4 - \frac{56}{27}\lambda^3 - \frac{5}{\sqrt{3}}\lambda^2 + \frac{8}{3}\lambda - \frac{1}{\sqrt{3}} \leq$$
$$\leq -\frac{5}{\sqrt{3}}\lambda^2 + \frac{8}{3}\lambda - \frac{1}{\sqrt{3}} = -\frac{5}{\sqrt{3}}(\lambda - \frac{1}{\sqrt{3}})(\lambda - \frac{\sqrt{3}}{5}) < 0.$$

Функция (16) отрицательна, следовательно

$$f_3(\lambda)e^{2\lambda^2+4\lambda^4/27} \leq f_3(0) = 1, \lambda \in [0, \frac{1}{\sqrt{3}}).$$

Для $\lambda \in [1/\sqrt{3}, 2/\sqrt{3})$ имеем

$$f_3(\lambda) = \frac{2}{3\sqrt{3}}\lambda^3 - \frac{1}{3}\lambda^2 - \frac{5}{3\sqrt{3}}\lambda + 1; f_3'(\lambda) = \frac{2}{\sqrt{3}}\lambda^2 - \frac{2}{3}\lambda - \frac{5}{3\sqrt{3}};$$

$$f_3'(\lambda) + (4\lambda + \frac{16\lambda^3}{9n})f_3(\lambda) = \frac{32}{81\sqrt{3}}\lambda^6 - \frac{16}{81}\lambda^5 + \frac{136}{81\sqrt{3}}\lambda^4 - \frac{20}{27}\lambda^3 - \frac{14}{3\sqrt{3}}\lambda^2 + \frac{10}{3}\lambda - \frac{5}{3\sqrt{3}}.$$

Рассмотрим многочлен $P_3(\lambda) = f_3'(\lambda) + (4\lambda + 16\lambda^3/(27))f_3(\lambda)$, $\lambda \in \mathbb{R}$. Для применения теоремы Будана-Фурье найдем его производные

$$\begin{aligned}
P_3^{(1)}(\lambda) &= \frac{64}{27\sqrt{3}}\lambda^5 - \frac{80}{81}\lambda^4 + \frac{544}{81\sqrt{3}}\lambda^3 - \frac{20}{9}\lambda^2 - \frac{28}{3\sqrt{3}}\lambda + \frac{10}{3};\\
P_3^{(2)}(\lambda) &= \frac{320}{27\sqrt{3}}\lambda^4 - \frac{320}{81}\lambda^3 + \frac{544}{27\sqrt{3}}\lambda^2 - \frac{40}{9}\lambda - \frac{28}{3\sqrt{3}};\\
P_3^{(3)}(\lambda) &= \frac{1280}{27\sqrt{3}}\lambda^3 - \frac{320}{27}\lambda^2 + \frac{1088}{27\sqrt{3}}\lambda - \frac{40}{9};\\
P_3^{(4)}(\lambda) &= \frac{1280}{9\sqrt{3}}\lambda^2 - \frac{640}{27}\lambda + \frac{1088}{27\sqrt{3}};\\
P_3^{(5)}(\lambda) &= \frac{2560}{9\sqrt{3}}\lambda - \frac{640}{27}.
\end{aligned}$$

Непосредственные вычисления показывают, что

$$\begin{array}{ll}
+,\ P_3(1/\sqrt{3}) = \dfrac{95}{2187\sqrt{3}} > 0; & +,\ P_3(2/\sqrt{3}) = \dfrac{47}{2187\sqrt{3}} > 0;\\
+,\ P_3^{(1)}(1/\sqrt{3}) = \dfrac{50}{243} > 0; & +,\ P_3^{(1)}(2\sqrt{3}) = \dfrac{854}{729} > 0;\\
-,\ P_3^{(2)}(1/\sqrt{3}) = -\dfrac{572}{81\sqrt{3}} < 0; & +,\ P_3^{(2)}(2/\sqrt{3}) = \dfrac{4660}{243\sqrt{3}} > 0;\\
+,\ P_3^{(3)}(1\sqrt{3}) = \dfrac{2504}{243} > 0; & +,\ P_3^{(3)}(2/\sqrt{3}) = \dfrac{11848}{243} > 0;\\
+,\ P_3^{(4)}(1\sqrt{3}) = \dfrac{64}{\sqrt{3}} > 0; & +,\ P_3^{(4)}(2\sqrt{3}) = \dfrac{4928}{27\sqrt{3}} > 0;\\
+,\ P_3^{(5)}(1/\sqrt{3}) = \dfrac{640}{9} > 0; & +P_3^{(5)}(2/\sqrt{3}) = \dfrac{4480}{27} > 0;\\
+,\ P_3^{(6)}(1\sqrt{3}) = \dfrac{2560}{9\sqrt{3}} > 0; & +,\ P_3^{(6)}(2\sqrt{3}) = \dfrac{2560}{9\sqrt{3}} > 0.
\end{array}$$

По теореме Будана-Фурье многочлен $P_3^{(3)}(\lambda)$ положителен на сегменте $\lambda \in [1/\sqrt{3}, 2/\sqrt{3}]$. Следовательно функция $P_3^{(1)}(\lambda)$ выпукла на сегменте $[1/\sqrt{3}, 2\sqrt{3}]$. Так как

$$P_3^{(1)}(1/\sqrt{3}) = 95/2187\sqrt{3} > 0, P_3^{(1)}(\sqrt{3}/2) = -10/27 < 0,$$

$$P_3^{(1)}(2/\sqrt{3}) = 47/2187\sqrt{3} > 0,$$

то существуют $1/\sqrt{3} < \lambda_1 < \lambda_2 < 2/\sqrt{3}, \lambda_1 < \sqrt{3}/2$, такие что $P_3^{(1)}(\lambda_1) = = P_3^{(1)}(\lambda_2) = 0$. Это значит, что функция $P_3(\lambda)$ возрастает при $[1/\sqrt{3}, \lambda_1] \cup [\lambda_2, 2/\sqrt{3}]$ и убывает при $[\lambda_1, \lambda_2]$. Так как $P_3^{(1)}(2/\sqrt{3}) = 854/729 > 0$, то из (16) следует, что максимальное значение функции $f_3(\lambda)e^{2\lambda^2+4\lambda^4/27}, \lambda \in [\lambda_2, 2/\sqrt{3}]$, достигается в точке $\lambda = 2/\sqrt{3}$ и равно $f_3(2/\sqrt{3})e^{2(2/\sqrt{3})^2+4(2/\sqrt{3})^4/27} = 0.693644 < 1$. Функция $f_3(\lambda)e^{2\lambda^2+4\lambda^4/27}, \lambda \in [1/\sqrt{3}, \lambda_2]$ принимает свое максимальное значение в точке λ_1.

Заметим, что $1/\sqrt{3} < 39/50 < 40/50 < \sqrt{3}/2$ и

$$P(39/50) = \frac{21400873}{9765625} - \frac{5559684907}{1464843750\sqrt{3}} > 0.00016 > 0;$$
$$P(40/50) = \frac{562616}{253125} - \frac{1629301}{421875\sqrt{3}} < -0.007 < 0.$$

Следовательно $39/50 < \lambda_1 < 40/50$. Поэтому $f_3(\lambda)e^{2\lambda^2+4\lambda^4/27}$ возрастает на $[39/50, \lambda_1]$ и убывает на $[\lambda_1, 40/50]$. Так как функция $f(\lambda)$ убывает, получим

$$f(\lambda_1)e^{2\lambda_1^2+4\lambda_1^4/27} \le f(39/50)e^{2(40/50)^2+4(40/50)^2/27} = 0.488788 < 1.$$

Это завершает доказательство утверждения. □

Утверждение 3. *Неравенство (3) справедливо при* $n = 4$.

Доказательство. В силу (14) достаточно доказать неравенство только для $0 \le \lambda < 3/2$. Рассмотрим функцию $f_4(\lambda) = \mathbb{P}\{D_4^+ > \lambda\}$ и производную

$$\Big(\ln(f_4(\lambda)e^{2\lambda^2+\lambda^4/9}\Big)' = \frac{f_4'(\lambda) + (4\lambda + 4\lambda^3/9)f_4(\lambda)}{f_4(\lambda)}. \tag{17}$$

При $0 < \lambda < 1/2$ имеем

$$f_4(\lambda) = -\frac{1}{16}\lambda^4 - \frac{3}{8}\lambda^3 - \frac{3}{4}\lambda^2 - \frac{1}{2}\lambda + 1; f_4'(\lambda) = -\frac{1}{4}\lambda^3 - \frac{9}{8}\lambda^2 - \frac{3}{2}\lambda - \frac{1}{2};$$
$$f_4'(\lambda) + (4\lambda + \frac{4}{9}\lambda^3)f_4(\lambda) = -\frac{1}{36}\lambda^7 - \frac{1}{6}\lambda^6 - \frac{7}{12}\lambda^5 - \frac{31}{18}\lambda^4 - \frac{101}{36}\lambda^3 - \frac{25}{8}\lambda^2 + \frac{5}{2}\lambda - \frac{1}{2} \le$$
$$\le -\frac{25}{8}\lambda^2 + \frac{5}{2}\lambda - \frac{1}{2} = -\frac{1}{8}\Big(5\lambda - 2\Big)^2 \le 0.$$

Из (17) следует, что

$$f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \le f_4(0) = 1, \text{ при } 0 \le \lambda < 1/2.$$

Если $1/2 \le \lambda < 1$, то

$$f_4(\lambda) = \frac{3}{16}\lambda^4 + \frac{1}{4}\lambda^3 - \frac{9}{16}\lambda^2 - \frac{25}{32}\lambda + 1; f_4'(\lambda) = \frac{3}{4}\lambda^3 + \frac{3}{4}\lambda^2 - \frac{9}{8}\lambda - \frac{25}{32};$$
$$f_4'(\lambda) + (4\lambda + \frac{4}{9}\lambda^3)f_4(\lambda) = \frac{\lambda^7}{12} + \frac{\lambda^6}{9} + \frac{\lambda^5}{2} + \frac{47\lambda^4}{72} - \frac{19\lambda^3}{18} - \frac{19\lambda^2}{8} + \frac{23\lambda}{8} - \frac{25}{32}.$$

Если $1/2 \le \lambda \le 14/25$, то

$$f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \le f_4(1/2)e^{2(14/25)^2+(14/25)^4/9} = 0.968649 < 1 \text{ для } 1/2 \le \lambda < 14/25.$$

Если $0.99 \le \lambda < 1$, то

$$f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \le f_4(99/100)e^{2(1)^2+(1)^4/9} < 0.81 < 1, \ \frac{99}{100} \le \lambda < 1.$$

Рассмотрим $P_4(\lambda) = f_4'(\lambda) + (2\lambda^2 + \lambda^4/9)f_4(\lambda)$ на сегменте $[14/25, 99/100]$. Его производные равны

$$\begin{aligned}
P_4^{(1)}(\lambda) &= \frac{7\lambda^6}{12} + \frac{2\lambda^5}{3} + \frac{5\lambda^4}{2} + \frac{47\lambda^3}{18} - \frac{19\lambda^2}{6} - \frac{19\lambda}{4} + \frac{23}{8};\\
P_4^{(2)}(\lambda) &= \frac{7\lambda^5}{2} + \frac{10\lambda^4}{3} + 10\lambda^3 + \frac{47\lambda^2}{6} - \frac{19\lambda}{3} - \frac{19}{4};\\
P_4^{(3)}(\lambda) &= \frac{35\lambda^4}{2} + \frac{40\lambda^3}{3} + 30\lambda^2 + \frac{47\lambda}{3} - \frac{19}{3};\\
P_4^{(4)}(\lambda) &= 70\lambda^3 + 40\lambda^2 + 60\lambda + \frac{47}{3};\\
P_4^{(5)}(\lambda) &= 210\lambda^2 + 80\lambda + 60;\\
P_4^{(6)}(\lambda) &= 420\lambda + 80.
\end{aligned}$$

Непосредственные вычисления показывают, что

$$\begin{array}{ll}
-,\ P_4(14/25) = -\frac{8476927729}{1757812500000} < 0; & -,\ P_4(99/100) = -\frac{844091394167}{400000000000000} < 0;\\
-,\ P_4^{(1)}(14/25) = -\frac{333023813}{17578125000} < 0; & +,\ P_4^{(1)}(99/100) = \frac{4748354581669}{4000000000000} > 0;\\
-,\ P_4^{(2)}(14/25) = -\frac{139195339}{39062500} < 0; & +,\ P_4^{(2)}(99/100) = \frac{257817837493}{20000000000} > 0;\\
+,\ P_4^{(3)}(14/25) = \frac{3729043}{234375} > 0; & +,\ P_4^{(3)}(99/100) = \frac{8199290021}{120000000} > 0;\\
+,\ P_4^{(4)}(14/25) = \frac{694723}{9375} > 0; & +,\ P_4^{(4)}(99/100) = \frac{54657479}{300000} > 0;\\
+,\ P_4^{(5)}(14/25) = \frac{21332}{125} > 0; & +,\ P_4^{(5)}(99/100) = \frac{345021}{1000} > 0;\\
+,\ P_4^{(6)}(14/25) = \frac{1576}{5} > 0; & +,\ P_4^{(6)}(99/100) = \frac{2479}{5} > 0;\\
+,\ P_4^{(7)}(14/25) = 420 > 0; & +,\ P_4^{(7)}(99/100) = 420 > 0.
\end{array}$$

По теореме Будана-Фурье многочлен $P_4(\lambda)$ не имеет вещественных корней на отрезке $[14/25, 99/100]$. Так как $P_4(1/2) < 0$, то производная (17) отрицательна и, следовательно,

$$\begin{aligned}
&f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \le f_4(14/25)e^{2(14/25)^2+(14/25)^4/9} \le\\
&\le f_4(1/2)e^{2(14/25)^2+(14/25)^4/9} = 0.968649 < 1,\ \lambda \in [\frac{14}{25}, \frac{99}{100}).
\end{aligned}$$

Нам осталось доказать неравенство (3) на отрезке $[1, 3/2)$. В этом случае имеем

$$\begin{aligned}
&f_4(\lambda) = -\frac{3}{16}\lambda^4 + \frac{5}{8}\lambda^3 - \frac{3}{16}\lambda^2 - \frac{37}{32}\lambda + 1;\ f_4'(\lambda) = -\frac{3}{4}\lambda^3 + \frac{15}{8}\lambda^2 - \frac{3}{8}\lambda - \frac{37}{32};\\
&f_4'(\lambda) + (4\lambda + \frac{16}{4n}\lambda^3)f_4(\lambda) = -\frac{1}{12}\lambda^7 + \frac{5}{18}\lambda^6 - \frac{5}{6}\lambda^5 + \frac{143}{72}\lambda^4 - \frac{19}{18}\lambda^3 - \frac{11}{4}\lambda^2 + \frac{29}{8}\lambda - \frac{37}{32}.
\end{aligned}$$

Так как функция $f_4(\lambda), \lambda \ge 0$, убывает, то

$$\begin{aligned}
&f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \le f_4(53/50)e^{2(2(54/50)^2+(54/50)^4/9} = 0.855641 < 1 \text{ для } 53/50 \le \lambda \le 54/50;\\
&f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \le f_4(7/5)e^{2(29/20)^2+(29/20)^4/9} = 0.9255 < 1 \text{ для } 7/5 \le \lambda \le 29/20;\\
&f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \le f_4(29/20)e^{2(3/2)^2+(3/2)^4/9} = 0.910697 < 1 \text{ для } 29/20 \le \lambda \le 3/2.
\end{aligned}$$

Заметим, что $[1,3/2) = [1,53/50) \cup [53/50,54/50) \cup [54/50,7/5) \cup [7/5,29/20) \cup$ $\cup[29/20,3/2)$. Докажем неравенство (3) на отрезке $[1,53/50)$. Рассмотрим многочлен $P_4(\lambda) = f_4'(\lambda) + (4\lambda + 4\lambda^3/9), \lambda \in \mathbb{R}$, и его производные:

$$P_4^{(1)}(\lambda) = -\frac{7}{12}\lambda^6 + \frac{5}{3}\lambda^5 - \frac{25}{6}\lambda^4 + \frac{143}{18}\lambda^3 - \frac{19}{6}\lambda^2 - \frac{11}{2}\lambda + \frac{29}{8};$$
$$P_4^{(2)}(\lambda) = -\frac{7}{2}\lambda^5 + \frac{25}{3}\lambda^4 - \frac{50}{3}\lambda^3 + \frac{143}{6}\lambda^2 - \frac{19}{3}\lambda - \frac{11}{2};$$
$$P_4^{(3)}(\lambda) = -\frac{35}{2}\lambda^4 + \frac{100}{3}\lambda^3 - 50\lambda^2 + \frac{143}{3}\lambda - \frac{19}{3};$$
$$P_4^{(4)}(\lambda) = -70\lambda^3 + 100\lambda^2 - 100\lambda + \frac{143}{3};$$
$$P_4^{(5)}(\lambda) = -210\lambda^2 + 200\lambda - 100;$$
$$P_4^{(6)}(\lambda) = -420\lambda + 200;$$
$$P_4^{(7)}(\lambda) = -420.$$

Непосредственные вычисления показывают, что

$$\begin{array}{ll}
+,\ P_4(1) = \frac{1}{96} > 0; & +,\ P_4(53/50) = \frac{1205119163}{9375000000000} > 0;\\
-,\ P_4^{(1)}(1) = -\frac{13}{72} < 0; & -,\ P_4^{(1)}(53/50) = -\frac{89167729709}{562500000000} < 0;\\
+,\ P_4^{(2)}(1) = \frac{1}{6} > 0; & +,\ P_4^{(2)}(53/50) = \frac{1035722147}{1875000000} > 0;\\
+,\ P_4^{(3)}(1) = \frac{43}{6} > 0; & +,\ P_4^{(3)}(53/50) = \frac{42153899}{7500000} > 0;\\
-,\ P_4^{(4)}(1) = -\frac{67}{3} < 0; & -,\ P_4^{(4)}(53/50) = -\frac{1100417}{37500} < 0;\\
-,\ P_4^{(5)}(1) = -110 < 0; & -,\ P_4^{(5)}(53/50) = -\frac{30989}{250} < 0;\\
-,\ P_4^{(6)}(1) = -220 < 0; & -,\ P_4^{(6)}(53/50) = -\frac{1226}{5} < 0;\\
-,\ P_4^{(7)}(1) = -420 < 0; & -,\ P_4^{(7)}(53/50) = -420 < 0.
\end{array}$$

По теореме Будана-Фурье $P_4(\lambda)$ не имеет вещественных корней на сегменте $[1,53/50]$. Так как $P_4(1) = 1/96 > 0$, многочлен $P(\lambda)$ положителен на сегменте $[1,53/50]$. Из (17) следует, что

$$f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \le f_4(53/50)e^{2(53/50)^2+(53/50)^4/9} = 0.776938 < 1 \text{ для } 1 \le \lambda \le 53/50.$$

Применим теорему Будана-Фурье для доказательства неравенства $P_4(\lambda) < 0$ для $27/25 = 54/50 \le \lambda \le 7/5$. Для этого нам потребуются следующие вычисления:

$$\begin{array}{ll}
-,\ P_4(27/25) = -\frac{571111333}{195312500000} < 0; & -,\ P_4(7/5) = -\frac{149057}{22500000} < 0;\\
-,\ P_4^{(1)}(27/25) = -\frac{285915157}{1953125000} < 0; & +,\ P_4^{(1)}(7/5) = \frac{93067}{1125000} > 0;\\
+,\ P_4^{(2)}(27/25) = \frac{12866401}{19531250} > 0; & -,\ P_4^{(2)}(7/5) = -\frac{3697}{18750} < 0;
\end{array}$$

$$
\begin{array}{ll}
+,\ P_4^{(3)}(27/25) = \dfrac{2347739}{468750} > 0; & -,\ P_4^{(3)}(7/5) = -\dfrac{10021}{750} < 0; \\
-,\ P_4^{(4)}(27/25) = -\dfrac{298811}{9375} < 0; & -,\ P_4^{(4)}(7/5) = -\dfrac{6631}{75} < 0; \\
-,\ P_4^{(5)}(27/25) = -\dfrac{16118}{125} < 0; & -,\ P_4^{(5)}(7/5) = -\dfrac{1158}{5} < 0; \\
-,\ P_4^{(6)}(27/25) = -\dfrac{1268}{5} < 0; & -,\ P_4^{(6)}(7/5) = -388 < 0; \\
-,\ P_4^{(7)}(27/25) = -420 < 0; & -,\ P_4^{(7)}(7/5) = -420 < 0.
\end{array}
$$

По теореме Будана-Фурье $P_4(\lambda)$ не имеет вещественных корней на сегменте $[27/25, 7/5]$. Т.к. $P_4(27/25) < 0$, многочлен $P_4(\lambda)$ отрицателен при $\lambda \in [27/25, 7/5]$. Из (17) следует, что

$$f_4(\lambda)e^{2\lambda^2+\lambda^4/9} \leq f_4(27/25)e^{2(27/25)^2+(27/25)^4/9} = 0.776609 < 1,\ 27/25 \leq \lambda \leq 7/5.$$

Это завершает доказательство утверждения. □

Благодарности. Благодарю своего научного руководителя профессора В.М. Круглова за постановку задачи и полезные советы при работе над статьей.

Заключение

В статье доказано неравенство

$$\mathbb{P}\{\sqrt{n} \sup_{|x|<\infty} (F_n(x) - F(x)) > \lambda\} \leq \exp\{-2\lambda^2 - \lambda^4/36n\}$$

для $n \geq 39, \min\{\gamma n^{-1/6}, \sqrt{\ln 2/2}\} \leq \lambda \leq \sqrt{n}/2, \gamma = 1.0841$.
Для частных случаев $n = 2, 3, 4$ доказана более сильная оценка

$$\mathbb{P}\{\sqrt{n} \sup_{|x|<\infty} (F_n(x) - F(x)) > \lambda\} \leq \exp\{-2\lambda^2 - 4\lambda^4/9n\}.$$

Список литературы

[1] Смирнов Н.В. Приближение законов распределения случайных величин по эмпирическим данным // Успехи математических наук. 1944. № 10. С. 179–206.

[2] Dvoretzky A., Kiefer J., Wolfowitz J. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator // The Annals of Mathematical Statistics. 1956. Vol. 27, № 3. Pp. 642–669.

[3] Massart P. The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality // The Annals of Probability. 1990. Vol. 18, № 3. Pp. 1269–1283.

[4] Feller W. An Introduction to Probability Theory and Its Applications. New York: Wiley, 1968. 509 p.

Образец цитирования

Ташков И.А. Улучшение неравенства Массарта для статистики Смирнова // Вестник ТвГУ. Серия: Прикладная математика. 2018. № 1. С. 5–20. https://doi.org/10.26456/vtpmk489

Сведения об авторах

1. **Ташков Иван Андреевич**
   студент факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова.

   Россия, 119992, г. Москва, ГСП-1, Воробьевы горы, МГУ им. М.В. Ломоносова.
   E-mail: tashkov95@ya.ru

AN IMPROVEMENT OF MASSART'S INEQUALITY FOR THE DISTRIBUTION OF SMIRNOV'S STATISTIC

Tashkov Ivan Andreevich
Student of Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University
Russia, 119991, Moscow, GSP-1, 1-52, Leninskiye Gory, MSU.
E-mail: tashkov95@ya.ru

Received 10.12.2017, revised 25.12.2017.

Let F_n be the empirical distribution function for a sample of independent identically distributed random variables with distribution function F. The main result is the inequality

$$\mathbb{P}\{\sqrt{n} \sup_{|x|<\infty} (F_n(x) - F(x)) > \lambda\} \leq \exp\{-2\lambda^2 - \lambda^4/36n\}$$

for $n \geq 39, \min\{\gamma n^{-1/6}, \sqrt{\ln 2/2}\} \leq \lambda \leq \sqrt{n}/2, \gamma = 1.0841$. It is also proved for the same n and $\lambda \leq \sqrt{n}/2$ that

$$\mathbb{P}\{\sqrt{n} \sup_{|x|<\infty} (F_n(x) - F(x)) > \lambda\} \leq 2\exp^{(\ln 2)^2/(144n)} \exp\{-2\lambda^2 - \lambda^4/36n\}.$$

In particular cases $n = 2, 3, 4$ it is proved that

$$\mathbb{P}\{\sqrt{n} \sup_{|x|<\infty} (F_n(x) - F(x)) > \lambda\} \leq \exp\{-2\lambda^2 - 4\lambda^4/9n\}.$$

Keywords: distribution of Smirnov's statistics, exponential inequalities.

Citation

Tashkov I.A. An improvement of Massart's inequality for the distribution of Smirnov's statistic. *Vestnik TvGU. Seriya: Prikladnaya Matematika* [Herald of Tver State University. Series: Applied Mathematics], 2018, no. 1, pp. 5–20. (in Russian). https://doi.org/10.26456/vtpmk489

References

[1] Smirnov N.V. Approximation of the laws of distribution of random variables from empirical data. *Uspekhi Matematicheskikh Nauk* [Successes of Mathematical Sciences], 1944, no. 10, pp. 179–206. (in Russian)

[2] Dvoretzky A., Kiefer J., Wolfowitz J. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, 1956, vol. 27(3), pp. 642–669.

[3] Massart P. The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality. *The Annals of Probability*, 1990, vol. 18(3), pp. 1269–1283.

[4] Feller W. *An Introduction to Probability Theory and Its Applications*. Wiley, New York, 1968. 509 p.