

От редакционной коллегии

Настоящий выпуск тематического сборника "Статистические методы оценивания и проверки гипотез издаваемого Пермским государственным университетом, является 28-м. Предшествующие выпуски издавались в следующие годы: вып. 1 — 1978, вып. 2 — 1980, вып. 3 — 1982, вып. 4 — 1984, вып. 5 — 1986, вып. 6 — 1988, вып. 7 — 1990, вып. 8 — 1991, вып. 9 — 1993, вып. 10 — 1995, вып. 11 — 1996, вып. 12 — 1998, вып. 13 — 1999, вып. 14 — 2000, вып. 15 — 2001, вып. 16 — 2002, вып. 17 — 2003, вып. 18 — 2005, вып. 19 — 2006, вып. 20 — 2007, вып. 21 — 2008, вып. 22 — 2010, вып. 23 — 2011, вып. 24 — 2012, вып. 25 — 2013, вып. 26 — 2015, вып. 27 — 2016.

Сборники 1978–1988 гг. переведены в журнале "Journal of Soviet Mathematics" (1987: Vol. 39, № 2; Vol. 39, № 4; 1988: Vol. 40, № 2; Vol. 41, № 1; 1991: Vol. 53, № 6; Vol. 56, № 3); сборники 1990–2008 гг. — в журнале "Journal of Mathematical Sciences" (1990: Vol. 75, № 1; 1991: Vol. 75, № 2; 1993: Vol. 81, № 4; 1995: Vol. 88, № 6; 1996: Vol. 103, № 4; 1998: Vol. 103, № 5; 1999: Vol. 126, № 1; 2000: Vol. 119, № 3; 2001: Vol. 127, № 4; 2002, 2003: Vol. 189, № 6; 2005: Vol. 205, № 1; 2006: Vol. 220, № 6; Vol. 228, № 5; 2007: Vol. 221, № 4; 2008: Vol. 227, № 2).

Это издание включает в себя статьи преподавателей и научных сотрудников Московского и Пермского государственных университетов, Национального университета Узбекистана, НИУ Высшая школа экономики (г. Москва), Московского государственного технического университета им. Н.Э. Баумана, и ряда других научных и учебных заведений.

Большинство статей посвящено решению задач теории оценивания и статистической проверки гипотез. В ряде публикаций рассмотрены асимптотические задачи теории вероятностей и математической статистики, а также вероятностно-статистические модели, имеющие важное прикладное значение.

УДК 519.24

Асимптотические свойства последовательного интервального оценивания функционалов от функции распределения

А. А. Абдушукуров¹, Г. Г. Рахимова²

¹Филиал Московского Государственного Университета имени М.В.Ломоносова, факультет прикладной математики и информатики, Узбекистан, Ташкент; a_abdushukurov@rambler.ru; (+99871) 232-28-22; Узбекистан, 100060, г. Ташкент, проспект Амира Темура 22, Филиал Московского Государственного Университета имени М.В.Ломоносова, кафедра "Прикладная математика и информатика"

²Национальный Университет Узбекистана имени Мирзо Улугбека, математический факультет, Узбекистан, Ташкент; rakhimova_gulnoza@mail.ru; (+99871)246-39-25; Узбекистан, 100174, г. Ташкент, ВУЗ городок, Национальный Университет Узбекистана, кафедра "Теория вероятностей и математическая статистика"

Аннотация. Рассмотрено последовательное интервальное оценивание достаточно широкого класса функционалов от неизвестной функции распределения. Получены условия асимптотической состоятельности доверительного интервала фиксированной ширины и асимптотической эффективности момента останова.

Ключевые слова: случайная величина, момент останова, доверительный интервал, фиксированная ширина, асимптотическая состоятельность, асимптотическая эффективность.

1. Введение

В отличие от классических методов математической статистики, в которых число производимых наблюдений фиксируется заранее, методы последовательного анализа характеризуются тем, что момент прекращения наблюдений (момент останова) является случайным и определяется в зависимости от значений наблюдаемых данных и от принятой меры оптимальности построенной статистической оценки. Изложим основную асимптотическую задачу последовательного оценивания интервалами фиксированной ширины.

Рассмотрим на вероятностном пространстве (Ω, F, P) последовательность $\xi_1, \xi_2, \dots, \xi_n, \dots$ независимых одинаково нормально распределенных случайных величин со средним θ и дисперсией σ^2 . Используя эмпирическое среднее $\theta_n = \frac{1}{n} \sum_{i=1}^n \xi_i$, построим для неизвестного среднего θ интервальную оценку $I(n) = [\theta_n - \varepsilon, \theta_n + \varepsilon]$, где $\varepsilon > 0$ – малое число.

Пусть $0 < \gamma < 1$, $a = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$, где $\Phi(x)$ – функция распределения стандартной нормальной случайной величины. Тогда для $n_\varepsilon = \inf\left(n \geq 1 : n \geq \frac{a^2 \sigma^2}{\varepsilon^2}\right)$

$$\begin{aligned} P\{\theta \in I(n_\varepsilon)\} &= P\left\{\left|\frac{\sum_{i=1}^{n_\varepsilon}(\xi_i - \theta)}{\sigma\sqrt{n_\varepsilon}}\right| \leq \frac{\varepsilon\sqrt{n_\varepsilon}}{a\sigma} \times a\right\} \geq \\ &\geq P\left\{\frac{1}{\sqrt{n_\varepsilon}} \left|\sum_{i=1}^{n_\varepsilon} \frac{\xi_i - \theta}{\sigma}\right| \leq a\right\} = 2\Phi(a) - 1 = \gamma. \end{aligned}$$

Следовательно, $I(n_\varepsilon)$ является интервальной оценкой для θ с доверительным уровнем γ .

Если дисперсия нормального распределения σ^2 неизвестна, то как показано в работе [1], по выборке фиксированного неслучайного объема невозможно построить доверительный интервал фиксированной ширины для среднего θ . В этом случае задача последовательного оценивания доверительными интервалами фиксированной ширины сводится к задаче выбора правила или момента остановки $N_\varepsilon = \inf\left(n \geq 1 : n \geq a^2 \sigma_n^2 / \varepsilon^2\right)$, где σ_n^2 – состоятельная оценка для неизвестной дисперсии σ^2 , и нахождению условий, при которых $\lim_{\varepsilon \rightarrow 0} P\{\theta \in I(N_\varepsilon)\} \geq \gamma$ (асимптотическая состоятельность доверительного интервала $I(N_\varepsilon)$).

Чоу и Роббинс [2] разработали основную асимптотическую теорию последовательного доверительного оценивания интервалами фиксированной ширины в общем случае. Они рассмотрели задачу оценивания среднего неизвестного распределения при единственном предположении существования второго момента распределения. Задача доверительного оценивания среднего значения интервалами фиксированной ширины для распределения, притягивающемуся к устойчивому распределению, у которого не существует момент второго порядка, рассмотрена в работе [3].

Содержательный обзор работ по последовательному непараметрическому оцениванию интервалами фиксированной ширины приведен в моно-

графии [4]. Из последующих работ следует отметить статью [5], где изучено точечное и интервальное последовательное оценивание интегрального функционала от неизвестной спектральной плотности гауссовского стационарного процесса. В работе [6] приведен обзор приложений методов последовательного оценивания, включая метод непараметрического оценивания интервалами фиксированной ширины.

Начиная с работы [2], многие авторы в конкретных статистических моделях для доказательства состоятельности доверительных интервалов фиксированной ширины применяли классическую теорему Анскомба [7], которая предполагала выполнение условия "равномерной непрерывности по вероятности". Обзор таких работ приведен в [4]. Альтернативный подход к доказательству асимптотической состоятельности доверительных интервалов фиксированной ширины, основанный на слабом принципе инвариантности, детально рассмотрен в [8].

В данной работе для доказательства асимптотической состоятельности доверительных интервалов фиксированной ширины использован подход, основанный на применении предельных теорем для случайно остановленных стохастических процессов. В первой части работы приведены общие условия асимптотической состоятельности доверительного интервала фиксированной ширины и асимптотической эффективности момента остановки в последовательном интервальном оценивании широкого класса функционалов $\theta(F)$ от неизвестной функции распределения $F(x)$, $x \in R$ из Ψ , где Ψ – семейство функций распределений, удовлетворяющих определенным условиям регулярности. Во второй части работы рассмотрены задачи последовательного интервального оценивания неизвестной плотности вероятности линейного процесса.

Конкретные статистические модели, в которых для доказательства асимптотической состоятельности доверительных интервалов фиксированной ширины использован подход этой работы, основанный на применении предельных теорем для случайно остановленных стохастических процессов, рассмотрены в работах [9-10].

2. Последовательное интервальное оценивание функционала от неизвестной функции распределения

Рассмотрим на вероятностном пространстве (Ω, F, P) случайные величины $\xi_1, \xi_2, \dots, \xi_n$ с неизвестной функцией распределения $F(x)$, $x \in R$ из

Ψ , где Ψ – семейство функций распределений, удовлетворяющих определенным условиям регулярности. Для оценки функционала $\theta(F)$ от функции распределения F будем использовать состоятельную оценку $\theta_n = \theta_n(\xi_1, \xi_2, \dots, \xi_n)$, допускающую разложение

$$\begin{aligned} \theta_n &= \theta(F) + [\theta_n - E(\theta_n)] + [E(\theta_n) - \theta(F)] = \\ &= \theta(F) + \sum_{k=1}^n Y_n(\xi_k) + Z_n. \end{aligned} \quad (1)$$

Здесь величины $Y_n(\xi_k), 1 \leq k \leq n$ и Z_n таковы, что существуют числа $\alpha > 0, \sigma^2(F) > 0$, и при $n \rightarrow \infty$

$$n^\alpha \sum_{k=1}^n Y_n(\xi_k) \xrightarrow{D} N(0, \sigma^2(F)) \quad \text{и} \quad n^\alpha Z_n \rightarrow 0,$$

где $N(0, \sigma^2(F))$ – нормально распределенная случайная величина со средним 0, дисперсией $\sigma^2(F)$, а символ \xrightarrow{D} означает слабую сходимость случайных величин.

Пусть $0 < \gamma < 1, a = \Phi^{-1}\left(\frac{1+\gamma}{2}\right), I(n_\varepsilon) = [\theta_{n_\varepsilon} - \varepsilon, \theta_{n_\varepsilon} + \varepsilon]$ – интервал длины $2\varepsilon, \varepsilon > 0$, и

$$n_\varepsilon = \inf \left(n \geq 1 : n \geq \left(\frac{a^2 \sigma^2(F)}{\varepsilon^2} \right)^{\frac{1}{2\alpha}} \right). \quad (2)$$

Тогда из сходимости $\sigma^{-1}(F)n^\alpha(\theta_n - \theta(F)) \xrightarrow{D} N(0, 1)$ при $n \rightarrow \infty$ и $\varepsilon \rightarrow 0$ получаем

$$\begin{aligned} P \{ \theta(F) \in I(n_\varepsilon) \} &= P \{ |\theta_{n_\varepsilon} - \theta(F)| \leq \varepsilon \} = \\ &= P \left\{ \sigma^{-1}(F)n_\varepsilon^\alpha |\theta_{n_\varepsilon} - \theta(F)| \leq \frac{\sigma^{-1}(F)n_\varepsilon^\alpha \varepsilon}{a} a \right\} \geq \\ &\geq P \{ \sigma^{-1}(F)n_\varepsilon^\alpha |\theta_{n_\varepsilon} - \theta(F)| \leq a \} \rightarrow 2\Phi(a) - 1 = \gamma. \end{aligned}$$

Эта последовательная процедура обладает тем недостатком, что n_ε зависит от неизвестного функционала $\sigma^2(F)$, который необходимо оценить. Пусть $V_n^2 = V_n^2(\xi_1, \xi_2, \dots, \xi_n)$ – состоятельная оценка для $\sigma^2(F)$. Исходя из (2), определим момент остановки N_ε следующим образом:

$$N_\varepsilon = \inf \left(n \geq 1 : n \geq \left(\frac{a^2 V_n^2}{\varepsilon^2} \right)^{\frac{1}{2\alpha}} \right), \quad \varepsilon > 0. \quad (3)$$

Тогда последовательный доверительный интервал для $\theta(F)$, основанный на моменте остановки (3) имеет вид:

$$I(N_\varepsilon) = [\theta_{N_\varepsilon} - \varepsilon, \theta_{N_\varepsilon} + \varepsilon], \quad \varepsilon > 0.$$

Следуя Чоу и Роббинс [2], введем следующие понятия.

Определение 1.1. Доверительный интервал фиксированной ширины $I(N_\varepsilon)$ называется состоятельным, если $\lim_{\varepsilon \rightarrow 0} P\{\theta(F) \in I(N_\varepsilon)\} \geq \gamma$ для некоторого $0 < \gamma < 1$ и для всех $F \in \Psi$.

Определение 1.2. Момент остановки N_ε называется эффективным, если $I(N_\varepsilon)$ состоятелен и $\lim_{\varepsilon \rightarrow 0} \frac{N_\varepsilon}{n_\varepsilon} = 1$.

Заметим, что n_ε равен минимальному количеству наблюдений, необходимых для того, чтобы асимптотическая при $\varepsilon \rightarrow 0$ вероятность накрытия интервалом $I(n_\varepsilon) = [\theta_{n_\varepsilon} - \varepsilon, \theta_{n_\varepsilon} + \varepsilon]$ известного функционала $\theta(F)$ была не меньше γ . Поэтому асимптотическая эффективность в смысле Чоу и Роббинса момента остановки N_ε означает, что N_ε при малых ε в среднем ведет себя также, как и в случае известного функционала $\theta(F)$.

Теорема 2.1. Если $V_n^2 \rightarrow \sigma^2(F)$ с вероятностью 1 при $n \rightarrow \infty$, то
 1) для любого $\varepsilon > 0$ $P\{N_\varepsilon < \infty\} = 1$; 2) $N_\varepsilon \rightarrow \infty$ с вероятностью 1 при $\varepsilon \rightarrow 0$; 3) $\lim_{\varepsilon \rightarrow 0} E(N_\varepsilon) = \infty$; 4) $\frac{N_\varepsilon}{n_\varepsilon} \rightarrow 1$ с вероятностью 1 при $\varepsilon \rightarrow 0$.

Доказательство теоремы 2.1. 1) Для любого $\varepsilon > 0$

$$P\{N_\varepsilon = \infty\} = \lim_{n \rightarrow \infty} P\{N_\varepsilon \geq n\} \leq \lim_{n \rightarrow \infty} P\left\{n \leq \left(\frac{a^2 V_n^2}{\varepsilon^2}\right)^{\frac{1}{2\alpha}}\right\} = 0.$$

2) Из (3) видно, что N_ε возрастает при уменьшении ε и $N_\varepsilon \rightarrow \infty$ с вероятностью 1 при $\varepsilon \rightarrow 0$.

3) Из теоремы о монотонной сходимости $\lim_{\varepsilon \rightarrow 0} E(N_\varepsilon) = \infty$.

4) Из (2) и (3) для любого $\varepsilon > 0$ справедливы неравенства $\frac{(\varepsilon^2)^{\frac{1}{2\alpha}}}{(\varepsilon^2)^{\frac{1}{2\alpha}} + (a^2 \sigma^2(F))^{\frac{1}{2\alpha}}} \leq \frac{1}{n_\varepsilon} \leq \left(\frac{\varepsilon^2}{a^2 \sigma^2(F)}\right)^{\frac{1}{2\alpha}}$ и $\left(\frac{a^2 V_{N_\varepsilon}^2}{\varepsilon^2}\right)^{\frac{1}{2\alpha}} \leq N_\varepsilon \leq 1 + \left(\frac{a^2 V_{N_\varepsilon-1}^2}{\varepsilon^2}\right)^{\frac{1}{2\alpha}}$. Поскольку $N_\varepsilon \rightarrow \infty$ при $\varepsilon \rightarrow 0$ и $V_n^2 \rightarrow \sigma^2(F)$ при $n \rightarrow \infty$ с вероятностью 1, то $\left(\frac{V_{N_\varepsilon}^2}{\sigma^2(F)}\right)^{\frac{1}{2\alpha}} \rightarrow 1$ и $\left(\frac{V_{N_\varepsilon-1}^2}{\sigma^2(F)}\right)^{\frac{1}{2\alpha}} \rightarrow 1$ с вероятностью 1 при $\varepsilon \rightarrow 0$. Отсюда и из соотношения

$$\frac{(V_{N_\varepsilon}^2)^{\frac{1}{2\alpha}}}{(\varepsilon^2 a^{-2})^{\frac{1}{2\alpha}} + (\sigma^2(F))^{\frac{1}{2\alpha}}} \leq \frac{N_\varepsilon}{n_\varepsilon} \leq \frac{1}{n_\varepsilon} + \left(\frac{V_{N_\varepsilon-1}^2}{\sigma^2(F)}\right)^{\frac{1}{2\alpha}}$$

получается последнее утверждение теоремы.

Теорема 2.2. Если для некоторого $\varepsilon_0 > 0$

$$\sum_{m=1}^{\infty} \sup_{0 \leq \varepsilon \leq \varepsilon_0} P \left\{ \frac{N_\varepsilon}{n_\varepsilon} \geq m \right\} < \infty,$$

$$\text{то } \lim_{\varepsilon \rightarrow 0} \frac{E(N_\varepsilon)}{n_\varepsilon} = 1.$$

Для доказательства утверждения теоремы 2.2 достаточно показать (в силу теоремы 2.1 [11]) равномерную интегрируемость множества $\left\{ \frac{N_\varepsilon}{n_\varepsilon}, \varepsilon \geq 0 \right\}$. Как показано в работе [12] (лемма 3.2), для этого достаточно выполнение условия теоремы 2.2.

Пусть символ \xrightarrow{J} означает слабую сходимость в J -топологии Скорохода.

Теорема 2.3. Пусть выполнено условие теоремы 2.1 и (I) $n^\alpha (E(\theta_n) - \theta(F)) \rightarrow 0$ при $n \rightarrow \infty$; (II) $\{\zeta_\varepsilon(t), t \in [0, 1]\} \xrightarrow{J} \{W(t), t \in [0, 1]\}$ при $\varepsilon \rightarrow 0$. Здесь $\zeta_\varepsilon(t) = n_\varepsilon^\alpha \sigma^{-1}(F) \sum_{k=1}^{[n_\varepsilon t]} Y_{n_\varepsilon}(\xi_k), t \in [0, 1]$, а $W(t), t \in [0, 1]$ – винеровский процесс с $E(W(t)) = 0, E(W^2(t)) = 1$. Тогда
1) $\sigma^{-1}(F) N_\varepsilon^\alpha (\theta_{N_\varepsilon} - \theta(F)) \xrightarrow{D} N(0, 1)$ при $\varepsilon \rightarrow 0$;
2) $\lim_{\varepsilon \rightarrow 0} P(\theta(F) \in I(N_\varepsilon)) \geq \gamma$, то есть доверительный интервал $I(N_\varepsilon)$ фиксированной ширины является состоятельным.

Для доказательства утверждений теоремы 2.3 заметим, что из условия (II) теоремы 2.3 и утверждения 4) теоремы 2.1 получаем, что для случайного процесса $\{\zeta_\varepsilon(t), t \in [0, 1]\}$ и для случайной величины $N_\varepsilon/n_\varepsilon$ выполняются условия теоремы 2.8.1 в [13] о предельном распределении случайно остановленных стохастических процессов. Из этой теоремы, условия (I) теоремы 2.3 и из разложения (1) получаем утверждения теоремы 2.3.

Замечание 2.1. Если выполнены условия теорем 2.2. и 2.3., то момент остановки (3) является асимптотически эффективным в смысле Чоу и Роббинса.

3. Последовательное интервальное оценивание плотности вероятности линейного случайного процесса

Рассмотрим на вероятностном пространстве (Ω, F, P) строго стационарный линейный случайный процесс $\{X_t, t \geq 1\}$, представимый следую-

щим образом:

$$X_t = \sum_{n=0}^n g_n Y_{t-n}, \quad (4)$$

где $Y_t, t \in Z = \{\dots, -1, 0, 1, \dots\}$ – последовательность независимых одинаково распределенных случайных величин, а $\{g_n, n \geq 0\}$ – последовательность чисел такая, что $g_n \rightarrow 0$ при $n \rightarrow \infty$. Случайными процессами вида (7) описываются, например, схемы авторегрессии и смешанной авторегрессии скользящего среднего.

Через $f_{\vec{X}}(x_1, x_2, \dots, x_k)$ обозначим совместную плотность вероятности случайного вектора $\vec{X} = (X_1, \dots, X_k)$. Предположим, что маргинальная плотность вероятности $f(x), x \in (-\infty, \infty)$ неизвестна. Для оценки плотности вероятности $f(x)$ в точке x рассмотрим ядерную оценку

$$f_n(x) = \frac{1}{nh(n)} \sum_{i=1}^n K\left(\frac{x - X_i}{h(n)}\right), \quad (5)$$

где $\{K(x), x \in (-\infty, \infty)\}$ – неотрицательная и ограниченная функция, $\{h(n), n \geq 1\}$ – последовательность действительных чисел такая, что $h(n) \rightarrow 0$ при $n \rightarrow \infty$.

Замечание 3.1. Оценка (5) зависит от выбора убывающей последовательности чисел $\{h(n), n \geq 1\}$, неизвестной плотности вероятности, а также ее производных. При оптимальном выборе последовательности среднеквадратическая ошибка оценки (5) имеет наименьший порядок малости. Способы выбора оптимальной последовательности $\{h(n), n \geq 1\}$ на основе наблюдений, балансирования разброса и смещения оценки (5) приведены, например, в монографии [14].

Оценка (5) допускает разложение

$$f_n(x) = f(x) + \sum_{i=1}^n Y_n(K, X_i) + Z_n,$$

где

$$Y_n(K, X_i) = \frac{1}{nh(n)} K\left(\frac{x - X_i}{h(n)}\right) - E\left(\frac{1}{nh(n)} K\left(\frac{x - X_i}{h(n)}\right)\right), 1 \leq i \leq n,$$

$$Z_n = E(f_n(x) - f(x)).$$

Обозначим

$$k(u) = \int_{-\infty}^{\infty} \exp(iuy) K(y) dy, k_q = \lim_{u \rightarrow 0} \frac{1 - k(u)}{|u|^q},$$

и пусть $\varphi(u)$ и $\psi(u)$ являются характеристическими функциями случайных величин Y_1 и X_1 .

Введем условия:

$$(A) \int_{-\infty}^{\infty} K(y)dy < \infty, k = \int_{-\infty}^{\infty} K^2(u)du < \infty; \lim_{|y| \rightarrow \infty} |y|K(y) = 0,$$

$\int_{-\infty}^{\infty} |K(y+h) - K(y)|dy < c_1 h$, где c_1 – конечная положительная постоянная;

$$(B) \int_{-\infty}^{\infty} |u\varphi(u)|du < \infty, k_q < \infty, \left| \int_{-\infty}^{\infty} \exp(-iux)|u|^q\psi(u)du \right| < \infty \text{ для некоторого } q > 0;$$

(C) $E(|Y_1|^m) < \infty$ для некоторого $m > 0$ и если $m \geq 1$, то $E(Y_1) = 0$;

(D) $\sum_{i=n}^{\infty} i|g_i|^\alpha = O(n^{-\gamma})$ для некоторого $\gamma > 0$, где $\alpha = m/2$, если $m \leq 1$ и $\alpha = 1/2$, если $m > 1$;

(H) $h(n) = n^{-\beta}$, $0 < \beta < 1$, $C > 0$.

Замечание 3.2. Константа C в условии (H) выбирается на основе наблюдений. Для $\beta = \beta_0 = \frac{1}{2q+1}$ согласно [15] константа

$$C = C(f(x), f^{(q)}(x)) = \frac{kf(x)}{(\sqrt{2q}/k_q f^{(q)}(x))^{\beta_0}}$$

и нетрудно показать, что если $f_n(x)$ и $f_n^{(q)}(x)$ – состоятельные оценки для $f(x)$ и $f^{(q)}(x)$ соответственно, то

$$C_n = C(f_n(x), f_n^{(q)}(x)) = \frac{kf_n(x)}{(\sqrt{2q}/k_q f_n^{(q)}(x))^{\beta_0}}$$

является состоятельной оценкой для C . В случае $\beta_0 < \beta < 1$ задача выбора константы C на основе наблюдений остается открытой, по крайней мере в пределах данной работы.

Не нарушая общности, рассуждений предположим $C = 1$.

Приведем две теоремы из работ [16,17].

Теорема 3.1 [16]. Пусть выполнены условия (A), (B), (C), (D), (H) и $(2q+1)^{-1} < \beta < 1$. Тогда

$$\sigma^{-1}n^{(1-\beta)/2} \sum_{i=1}^n Y_n(K, X_i) \xrightarrow{D} N(0, 1) \text{ и } n^{(1-\beta)/2} Z_n \rightarrow 0 \text{ при } n \rightarrow \infty,$$

где $\sigma^2 = kf(x)$

Пусть $0 < \gamma < 1$, $a = \Phi^{-1}\left(\frac{1+\gamma}{2}\right)$ и $I(n_\varepsilon) = [f_{n_\varepsilon}(x) - \varepsilon, f_{n_\varepsilon}(x) + \varepsilon]$ – интервал длины 2ε для любого $\varepsilon > 0$, где

$$n_\varepsilon = \inf \left(n \geq 1 : n \geq \left(\frac{a^2 \sigma^2}{\varepsilon^2} \right)^{1/(1-\beta)} \right). \quad (6)$$

Из теоремы 2.1. получаем

$$\begin{aligned} P \{f(x) \in I(n_\varepsilon)\} &= P \{|f_{n_\varepsilon}(x) - f(x)| \leq \varepsilon\} = \\ &= P \left\{ \sigma^{-1} n_\varepsilon^{\frac{1-\beta}{2}} |f_{n_\varepsilon}(x) - f(x)| \leq \frac{\sigma^{-1} n_\varepsilon^{\frac{1-\beta}{2}} \varepsilon}{a} a \right\} \geq \\ &\geq P \left\{ \sigma^{-1} n_\varepsilon^{\frac{1-\beta}{2}} |f_{n_\varepsilon}(x) - f(x)| \leq a \right\} \rightarrow 2\Phi(a) - 1 = \gamma \end{aligned}$$

при $\varepsilon \rightarrow 0$. Эта последовательная процедура обладает тем недостатком, что n_ε зависит от неизвестной плотности вероятности $f(x)$ через σ^2 , поэтому вместо n_ε рассмотрим момент остановки

$$N_\varepsilon = \inf \left(n \geq 1 : n \geq \left(\frac{a^2 k}{\varepsilon^2} f_n(x) \right)^{\frac{1}{1-\beta}} \right), \quad \varepsilon > 0. \quad (7)$$

Тогда последовательный доверительный интервал для $f(x)$, основанный на моменте остановки (7), имеет вид $I(N_\varepsilon) = [f_{N_\varepsilon}(x) - \varepsilon, f_{N_\varepsilon}(x) + \varepsilon]$, $\varepsilon > 0$.

Введем условие

(E) $\int_{-\infty}^{\infty} |K(t) - aK(at)| dt \leq c_2 \frac{b}{1-a}$, $\int_{-\infty}^{\infty} |u|^s |\varphi(u)| du < \infty$ для $0 < b < a < 1$ и $s = 0, 1, 2$, где c_2 – положительная постоянная.

Теорема 3.2 [17]. 1) Если выполнены условия (A), (B), (C), (D), (E), (H) и $\frac{1}{2q+1} < \beta < \frac{1}{2}$, то $f_n(x) \rightarrow f(x)$ с вероятностью 1 при $n \rightarrow \infty$ и $\sum_{m=1}^{\infty} \sup_{0 \leq \varepsilon \leq \varepsilon_0} P \left\{ \frac{N_\varepsilon}{n_\varepsilon} \geq m \right\} < \infty$ для некоторого $\varepsilon_0 > 0$.

2) Если выполнены условия (A), (B), (C), (D) и $\frac{1}{2q+1} < \beta < 1$, то

$$\{\zeta_\varepsilon(t), t \in [0, 1]\} \xrightarrow{J} \{W(t), t \in [0, 1]\} \quad \text{при } \varepsilon \rightarrow 0,$$

где $\zeta_\varepsilon(t) = \sigma^{-1} n_\varepsilon^{(1-\beta)/2} \sum_{i=1}^{[n_\varepsilon t]} Y_{n_\varepsilon}(X_i)$, $t \in [0, 1]$.

Из теорем 3.1, 3.2 и 3.3 следует

Теорема 3.3. Если выполнены условия (A), (B), (C), (D), (E), (H) и $\frac{1}{2q+1} < \beta < \frac{1}{2}$, то для n_ε из (6) и момента остановки (7):

- 1) Для любого $\varepsilon > 0 : P\{N_\varepsilon < \infty\} = 1$;
- 2) $N_\varepsilon \rightarrow \infty$ с вероятностью 1 при $\varepsilon \rightarrow 0$;
- 3) $\lim_{\varepsilon \rightarrow 0} E(N_\varepsilon) = \infty$;
- 4) $\frac{N_\varepsilon}{n_\varepsilon} \rightarrow 1$ с вероятностью 1 при $\varepsilon \rightarrow 0$;
- 5) $\lim_{\varepsilon \rightarrow 0} \frac{E(N_\varepsilon)}{n_\varepsilon} = 1$.

Из теоремы 2.2 также следует

Теорема 3.4. Пусть выполнены условия (A), (B), (C), (D), (H) и $\frac{1}{2q+1} < \beta < 1$. Тогда

- 1) $\sigma^{-1} N_\varepsilon^{\frac{1-\beta}{2}} (f_{N_\varepsilon}(x) - f(x)) \xrightarrow{D} N(0, 1)$ при $\varepsilon \rightarrow 0$;
- 2) $\lim_{\varepsilon \rightarrow 0} P(f(x) \in I(N_\varepsilon)) \geq \gamma$, то есть доверительный интервал фиксированной ширины $I(N_\varepsilon)$ является состоятельным.

Из теорем 2.3 и 2.4 также следует, что если выполнены условия (A), (B), (C), (D), (E), (H) и $\frac{1}{2q+1} < \beta < \frac{1}{2}$, то момент остановки (7) является асимптотически эффективным в смысле Чоу и Роббинса.

Библиографический список

1. Dantzig G. B. On the nonexistence of tests of Students hypothesis having power functions independent of σ^2 // Ann. Math. Statist. – 1940. – V. 11. – P. 186–192.
2. Chow Y. S., Robbins H. On the asymptotic theory of fixed-width sequential confidence intervals for the mean // Ann. Math. Statist. – 1965. – V. 36. – P. 457–462.
3. Рахимова Г. Г., Турсунов Г. Т. Последовательное оценивание среднего значения интервалами фиксированной ширины // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр./ Перм. гос. нац. иссл. ун-т. – Пермь, 2015. – Вып.26. – С. 58–67.
4. Ghosh M., Mukhopadhyay N., Sen P. K. Sequential estimation. – New York: John Wiley & Sons Inc., 1997.
5. Shiohama T., Tanagushi M. Sequential estimation for a functional of the spectral density of a Gaussian stationary process // Ann.Inst.Statist. Math. – 2001. – V. 53, № 1. – P. 142–158.
6. Mukhopadhyay N., De Silva B.M. Sequential methods and their applications. – CRC Press. Chapman and Hall/CRC, 2009.
7. Anscombe F. J. Large sample theory of sequential estimation // Proc. Cambridge Philos. Soc. – 1952. – V. 48. – P. 600–607.

8. *Sen P. K.* Sequential nonparametrics: Invariance principles and statistical inference. – New York: John Wiley & Sons. Inc., 1981.
9. *Рахимова Г. Г., Турсунов Г. Т.* Последовательное оценивание интервалами фиксированной ширины коэффициента переноса в диффузионном уравнении // Материалы республиканской научно-практической конференции "Статистика и ее применения". – Ташкент, 2013. – С. 74–77.
10. *Рахимова Г. Г., Турсунов Г. Т.* Последовательное непараметрическое оценивание интервалами фиксированной ширины функции регрессии // Материалы Международной научной конференции "Теория вероятностей, случайные процессы, математическая статистика и приложения". – Минск, 2015. – С. 259–261.
11. *Geertsema J. C.* Sequential confidence intervals based on rank tests // Ann. Math. Statist. – 1970. – V. 41. – P. 1016–1026.
12. *Bickel P. J., Yahav J. A.* Asymptotically optimal Bayes and minimax procedures in sequential estimation // Ann. Math. Statist. – 1968. – V. 39. – P. 442–456.
13. *Silvestrov D. S.* Limit Theorems for Randomly Stopped Stochastic Processes. London: Springer-Verlag. 2004.
14. *Надарая Э.А.* Непараметрическое оценивание плотности вероятности и кривой регрессии. – Тбилиси: Изд-во Тбилисского Университета, 1983.
15. *Parzen E.* On estimation of a probability density function and mode // Ann. Math. Statist. – 1962. – V. 33, № 3. – P. 1065–1076.
16. *Chanda K. C.* Density estimation for linear processes // Ann. Inst. Statist. Math. – 1983. – V. 35. Part A. – P. 439–446.
17. *Mirzakhmedov M. A., Tursunov G. T.* The sequential interval estimating of linear random processes // Sixth intern. Vilnius conference on probability theory and math. statist.: Abstracts of comm. – Vilnius, 1993. – P. 35–36.

Sequential estimating of mean by fixed-width intervals

A.A.Abdushukurov¹, G. G. Rakhimova²

¹Branch of Moscow State University in Tashkent named after M. V. Lomonosov, faculty of Applied Mathematics and Informatics, department of Applied Mathematics and informatics, Uzbekistan, Tashkent; a_abdushukurov@rambler.ru

²National University of Uzbekistan named after Mirzo Ulugbek, faculty of mechanics and mathematics, department of Probability Theory and Mathematical Statistics, Uzbekistan, Tashkent; rakhimova_gulnoza@mail.ru

Abstract. *We consider sequential confidence estimation of functionals of an unknown distribution function. Conditions of consistency of fixed-width confidence intervals and asymptotic efficiency of stopping times are obtained.*

Key words: *random variable, stopping time, confidence interval, fixed width, consistency, asymptotic efficiency.*

УДК 519.2

Комбинированный сеточный метод оценивания параметров распределений Линника и Миттаг-Леффлера¹

В. Ю. Королев², Д. В. Фомина³, А. А. Щербинина⁴

²Факультет вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова; vkorolev@cs.msu.su

³Факультет вычислительной математики и кибернетики, Московский государственный университет имени М. В. Ломоносова

⁴Факультет вычислительной математики и кибернетики, Московский государственный университет имени М. В. Ломоносова

Аннотация. В данной работе описан комбинированный сеточный метод статистического оценивания параметров распределений Линника и Миттаг-Леффлера. Особенностью задачи является отсутствие явных представлений плотностей указанных распределений в терминах элементарных функций, что делает неприменимым метод максимального правдоподобия. Более того, в этой ситуации “чистый” метод моментов также неприменим из-за того, что указанные распределения имеют тяжелые хвосты. Предлагаемый в данной статье сеточный метод основан на возможности представления распределений Линника и Миттаг-Леффлера соответственно в виде масштабной смеси распределения Лапласа и смешанного показательного распределения. В обоих случаях смешивающее распределение одно и то же и соответствует отношению двух независимых односторонних строго устойчивых случайных величин с одним и тем же характеристическим показателем. Оказалось, что плотность этого распределения может быть выражена в терминах элементарных функций и имеет довольно простой явный вид. На первом этапе предлагаемого метода с помощью версии EM-алгоритма ищется дискретная аппроксимация указанного смешивающего распределения. На втором этапе путем минимизации невязки между полученной дискретной оценкой и теоретической смешивающей плотностью по параметру находится оценка параметра анализируемого распределения. Описаны разные способы выбора узлов сетки. Приведены результаты имитационного моделирования, иллюстрирующие сходимость метода. Также приведены результаты сравнения описываемого метода с оценками, полученными методом логарифмических моментов.

Ключевые слова: распределение Линника; распределение Миттаг-Леффлера; масштабная смесь; EM-алгоритм; логарифмический момент.

¹Работа выполнена при поддержке Российского научного фонда (проект 18-11-00155).

© Королев В.Ю., Фомина Д.В., Щербинина А.А. 2018

1. Введение

Распределения Линника и Миттаг-Леффлера представляют особый интерес как «тяжелохвостые» модели статистических закономерностей, при которых большие значения наблюдаемых характеристик встречаются намного чаще, чем предписывают классические нормальные или экспоненциальные модели. Эти распределения возникают в некоторых задачах, связанных с дифференциальными уравнениями дробного порядка в физике, астрономии, финансовой математике и других областях. Распределения Линника и Миттаг-Леффлера традиционно рассматривается вместе, поскольку характеристическая функция (х. ф.) распределения Линника имеет такой же аналитический вид, как преобразование Лапласа–Стильтьеса (п. Л.–С.) распределения Миттаг-Леффлера. Поэтому эти распределения обладают многими сходными свойствами. В частности, они являются геометрически устойчивыми, так как являются предельными для геометрических случайных сумм независимых одинаково распределенных случайных величин (с. в.) с бесконечными дисперсиями, и потому представимы в виде масштабных смесей устойчивых законов, в которых смешивающим является распределение Вейбулла.

В данной работе описан комбинированный сеточный метод статистического оценивания параметров распределений Линника и Миттаг-Леффлера. Особенностью задачи является отсутствие явных представлений плотностей указанных распределений в терминах элементарных функций, что делает неприменимым метод максимального правдоподобия. Более того, в этой ситуации «чистый» метод моментов также неприменим из-за того, что указанные распределения имеют тяжелые хвосты. Предлагаемый в данной статье сеточный метод основан на возможности представления распределений Линника и Миттаг-Леффлера соответственно в виде масштабной смеси распределения Лапласа и смешанного показательного распределения. В обоих случаях смешивающее распределение одно и то же и соответствует отношению двух независимых односторонних строго устойчивых с. в. с одним и тем же характеристическим показателем. Оказалось, что плотность этого распределения может быть выражена в терминах элементарных функций и имеет довольно простой явный вид. На первом этапе предлагаемого метода с помощью версии EM-алгоритма ищется дискретная аппроксимация указанного смешивающего распределения. На втором этапе путем минимизации невязки между полученной дискретной оценкой и теоретической смешивающей плотностью по параметру находится оценка параметра анализируемого распределения. Описаны разные способы выбора узлов сетки. Приведены результаты имитационного моделирования, иллю-

стрирующие сходимость метода. Также приведены результаты сравнения описываемого метода с оценками, полученными методом логарифмических моментов.

Символы $\stackrel{d}{=}$ и \implies будут соответственно обозначать совпадение распределений и сходимость распределений. С. в. со стандартным показательным распределением будет обозначаться W_1 .

2. Распределения Миттаг-Леффлера и Линника

Пусть $\alpha \in (0, 1]$ и M_α – неотрицательная с. в. с преобразованием Лапласа–Стильтьеса

$$\psi_\alpha(s) \equiv \mathbf{E} \exp\{-sM_\alpha\} = (1 + s^\alpha)^{-1}, \quad s \geq 0. \quad (1)$$

Распределения с п. Л.–С. (1) принято называть *распределениями Миттаг-Леффлера*. Происхождение этого названия связано с тем, что плотность, соответствующая п. Л.–С. (1), имеет вид

$$f_\alpha^M(x) = \frac{1}{x^{1-\alpha}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{\alpha n}}{\Gamma(\alpha n + 1)} = -\frac{d}{dx} E_\alpha(-x^\alpha), \quad x \geq 0,$$

где $E_\alpha(z)$ – функция Миттаг-Леффлера индекса α , определяемая как степенной ряд

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)}, \quad \alpha > 0, \quad z \in \mathbb{Z}.$$

Функция распределения (ф. р.), соответствующая плотности $f_\alpha^M(x)$, будет обозначаться $F_\alpha^M(x)$.

При $\alpha = 1$ распределение Миттаг-Леффлера превращается в стандартное показательное распределение: $M_1 \stackrel{d}{=} W_1$.

Для ф. р. $F_\alpha^M(x)$ при $x > 0$ справедливо интегральное представление

$$F_\alpha^M(x) = 1 - \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \frac{z^{\alpha-1} e^{-zx} dz}{1 + z^{2\alpha} + 2z^\alpha \cos(\pi\alpha)}.$$

Из этого представления вытекает, что

$$M_\alpha \stackrel{d}{=} W_1 \cdot R_\alpha,$$

где с. в. в правой части независимы, причем с. в. R_α имеет плотность

$$p_\alpha^R(z) = \frac{\sin(\pi\alpha)}{\pi} \cdot \frac{z^{\alpha-1}}{1 + z^{2\alpha} + 2z^\alpha \cos(\pi\alpha)}, \quad z \geq 0.$$

В работе [14] показано, что плотность $p_\alpha^R(z)$ соответствует отношению двух независимых с. в. с одним и тем же односторонним строго устойчивым распределением с характеристическим показателем α .

При $\alpha < 1$ плотность (10) имеет хвост, убывающий степенным образом: если $0 < \alpha < 1$, то

$$\lim_{x \rightarrow \infty} x^{\alpha+1} f_{\alpha}^M(x) = \frac{\Gamma(\alpha+1)}{\pi} \sin \pi \alpha$$

(см., например, [9]).

Моменты с. в. с распределением Миттаг-Леффлера порядков $\beta \geq \alpha$ бесконечны, но при $0 < \beta < \alpha < 1$

$$\mathbf{E}M_{\alpha}^{\beta} = \Gamma(1 + \frac{\beta}{\alpha})\Gamma(1 - \frac{\beta}{\alpha}).$$

Распределения с х. ф.

$$f_{\alpha}^L(t) = (1 + |t|^{\alpha})^{-1}, \quad t \in \mathbb{R},$$

где $0 < \alpha \leq 2$, принято называть *распределениями Линника* (в работе [24] предложено альтернативное менее употребительное название α -Laplace distribution). Они были введены Ю. В. Линником в 1953 г. [1]. При $\alpha = 2$ распределение Линника превращается в распределение Лапласа, соответствующее плотности

$$f^{\Lambda}(x) = \frac{1}{2}e^{-|x|}, \quad x \in \mathbb{R}. \quad (2)$$

Лапласовская с. в. с плотностью (2) и ее ф. р. будут соответственно обозначаться Λ и $F^{\Lambda}(x)$.

С. в., имеющая распределение Линника с параметром α , ее ф. р. и плотность будут соответственно обозначаться L_{α} , F_{α}^L и f_{α}^L . При этом $F_2^L(x) \equiv F^{\Lambda}(x)$, $x \in \mathbb{R}$.

Распределения Линника обладают многими интересными свойствами. Лин [20] доказал саморазложимость F_{α}^L . Существование моментов с. в. L_{α} обсуждается в работе [3]. Абсолютные моменты порядков $\beta < \alpha$ с. в. L_{α} имеют вид

$$\mathbf{E}|L_{\alpha}|^{\beta} = \frac{2^{\beta}}{\sqrt{\pi}} \cdot \frac{\Gamma(1 + \frac{\beta}{\alpha})\Gamma(\frac{1+\beta}{2})\Gamma(1 - \frac{\beta}{\alpha})}{\Gamma(1 - \frac{\beta}{2})}.$$

Распределения Линника безгранично делимы [7], имеют бесконечный пик плотности в нуле при $\alpha \leq 1$ [7]. В работе [10] показано, что при $0 < \alpha < 2$ хвосты распределения Линника убывают степенным образом:

$$\lim_{x \rightarrow \infty} x^{\alpha} [1 - F_{\alpha}^L(x)] = \frac{\Gamma(\alpha)}{\pi} \sin \frac{\pi \alpha}{2}.$$

В работах [13, 14, 17, 23] получены разнообразные представления распределений Линника в виде смесей. В частности, в работах [14, 17] показано, что

$$L_{\alpha} \stackrel{d}{=} \Lambda \cdot \sqrt{R_{\alpha/2}},$$

где с. в. в правой части независимы. Отсюда вытекает, что для плотности f_α^L справедливо интегральное представление

$$f_\alpha^L = \frac{\sin(\pi\alpha/2)}{\pi} \int_0^\infty \frac{z^\alpha e^{-z|x|} dz}{1 + z^{2\alpha} + 2 \cos(\pi\alpha/2)}, \quad x \in \mathbb{R},$$

полученное прямым методом в работах [15, 16].

Пожалуй, чаще всего распределения Миттаг-Леффлера и Линника упоминаются как примеры *геометрически устойчивых распределений*. Это означает, что если X_1, X_2, \dots – независимые одинаково распределенные с. в., общее распределение которых принадлежит области нормального притяжения симметричного α -строго устойчивого распределения, а $\mathcal{NB}_{1,p}$ – независимая от X_1, X_2, \dots с. в. с геометрическим распределением

$$P(\mathcal{NB}_{1,p} = n) = p(1-p)^{n-1}, \quad n = 1, 2, \dots, \quad p \in (0, 1),$$

то для каждого $p \in (0, 1)$ существует константа $a_p > 0$ такая, что $a_p(X_1 + \dots + X_{\mathcal{NB}_{1,p}}) \implies L_\alpha$ при $p \rightarrow 0$. А если общее распределение с. в. X_1, X_2, \dots принадлежит области нормального притяжения одностороннего α -строго устойчивого распределения, то для каждого $p \in (0, 1)$ существует константа $a_p > 0$ такая, что $a_p(X_1 + \dots + X_{\mathcal{NB}_{1,p}}) \implies M_\alpha$ при $p \rightarrow 0$, см., например, [4] или [12].

Распределения Миттаг-Леффлера и Линника возникают в некоторых задачах, связанных с дифференциальными уравнениями дробного порядка и аномальной диффузией в физике, а также в астрономии, финансовой математике и других областях [8, 11, 18, 21, 22].

3. Методы статистического оценивания параметров распределений Линника и Миттаг-Леффлера

В связи с тем, что распределения Миттаг-Леффлера и Линника находят применение во многих прикладных задачах, большую актуальность приобретает проблема статистического оценивания их параметров.

В 1999 году в статье [10] были предложены некоторые методы оценки параметра распределения Линника, основанные на поведении хвостов распределения. В этой статье замечено, что хвосты распределения Линника ведут себя подобно закону Парето, и предлагается оценивать параметры распределения Линника на основе этого факта. В частности, рассмотрена оценка параметра α методом характеристических функций в предположении, что другие параметры (масштаба и сдвига) известны.

В 2001 году в статье [19] было предложено оценивать параметры распределений Линника и Миттаг-Леффлера с помощью дробных моментов.

Однако предложенный в [19] подход имеет существенный недостаток – в нем никак не учитывается тот факт, что для этих распределений существуют только моменты порядка строго меньше параметра распределения, и без каких-либо априорных предположений о значении этого параметра применять этот метод, вообще говоря, нельзя.

В работах [5, 6] был предложен метод оценки параметров рассматриваемых распределений методом логарифмическим моментов. Суть метода сводится к применению формул вида:

$$\hat{\alpha} = \frac{\pi}{\sqrt{3(\hat{\sigma}_{L'}^2 - \pi^2/12)}}, \quad \hat{\delta} = \frac{2\pi}{\sqrt{2(6\hat{\sigma}_{T'}^2 + \pi^2)}},$$

где L' и T' получаются из исходных выборок распределения Линника L и Миттаг-Леффлера T путем взятия логарифма от модуля значений в выборке и логарифма значений соответственно. В данных работах также было доказано, что получаемые оценки являются асимптотически несмещенными.

В настоящей работе рассматривается альтернативный метод – двухэтапный сеточный метод. Пожалуй, главным отличием предлагаемого подхода от указанных выше является то, что фактически результатом применения описываемых далее алгоритмов является не *точечная оценка* неизвестного параметра, а *функция* плотности смешивающего закона.

4. Описание алгоритма сеточного метода оценивания

На первом этапе двухэтапного сеточного метода с помощью модификации классического EM-алгоритма осуществляется разделение смеси, результатом чего является набор известных узлов и оценок соответствующих им весов, то есть дискретная аппроксимация смешивающего распределения. На втором этапе будем оценивать параметр распределения, подгоняя теоретическую плотность к полученной дискретной аппроксимации.

Пусть $k \in \mathbb{N}$ – достаточно большое число. На первом этапе на положительной полупрямой выделим конечный отрезок, на котором должна оказаться основная часть носителя смешивающего распределения. На этом отрезке выберем k *известных* точек – узлов сетки. Обозначим полученные узлы u_1, \dots, u_k .

Распределение Линника будем приближать конечной смесью распределений Лапласа:

$$F_{\alpha}^L(x) \approx \sum_{i=1}^k p_i F^{\Lambda}(x)$$

Распределение Миттаг-Леффлера будем приближать конечной смесью экспоненциальных законов:

$$F_{\delta}^M(x) \approx \sum_{i=1}^k p_i F^{W_1}(x)$$

Пусть x_1, \dots, x_n – анализируемая выборка объема n . Зададим итерационный процесс следующим образом:

$$p_i^{(m+1)} = \frac{p_j^{(m)}}{n} \sum_{j=1}^n \frac{\phi_{ij}}{\sum_{r=1}^k p_r^{(m)} \phi_{rj}},$$

где $p_1^{(m)}, \dots, p_{k-1}^{(m)}$ – оценки параметров p_1, \dots, p_{k-1} на m -й итерации, $i = 1, \dots, k-1$, $p_k = 1 - p_1 - \dots - p_{k-1}$,

$$\phi_{ij} = \frac{1}{2u_i} e^{-\left|\frac{x_j}{u_i}\right|}$$

в случае распределения Линника и

$$\phi_{ij} = \frac{1}{u_i} e^{-\frac{x_j}{u_i}}$$

в случае распределения Миттаг-Леффлера, $i = 1, \dots, k$, $j = 1, \dots, n$.

На втором этапе алгоритма происходит оценка параметра α по данным типа гистограммы. В качестве оценки неизвестного параметра будем считать такое его значение, которое доставляет минимум расстоянию между теоретической смешивающей плотностью $p_{\alpha}^R(z)$ и полученной на первом этапе ее дискретной аппроксимацией.

В качестве расстояния в данной работе были использованы следующие метрики:

- стандартная метрика пространства l^2 ;
- метрика l^1 ;
- метрика l^{∞} ;
- метрики из l^1 и l^2 с весовыми коэффициентами.

5. Оценка верхней границы сетки

При реализации описываемого двухэтапного метода очень важным является правильный выбор сетки, а главное – оптимальный выбор верхней границы сетки. В данном разделе рассмотрим один из возможных способов

верхней оценки границы сетки, накидываемой на носитель смешивающего распределения.

Попробуем оценить носитель с помощью неравенства Маркова. Однако в связи с тем, что строго устойчивые с. в. имеют конечные моменты лишь порядков, меньших α (см. [2]), использовать их для оценки параметра не получится.

Далее используем некоторые результаты из книги [2]. В оригинальных обозначениях $Y(\alpha, \theta)$ – строго устойчивая с. в. с нулевым параметром сдвига и единичным коэффициентом масштаба – имеет логарифмические моменты любого порядка. Воспользуемся этим. Если обозначить

$$y_{kn} = E(\log(|Y(\alpha, \theta)|))^n (\text{sign } Y(\alpha, \theta))^k,$$

то окажется, что y_{kn} является полиномом степени n относительно $1/\alpha$ и полиномом степени не выше $n + 1$ относительно θ .

В [2] доказано, что для любых допустимых значений параметров (α, θ) имеют место следующие равенства (при $n = 1, 2, \dots$):

$$y_{0n} = C_n(q_1, q_2, \dots, q_n),$$

$$y_{1n} = \theta C_n(q_1, q_2, \dots, q_n),$$

где $C_n(q_1, q_2, \dots, q_n)$ – полиномы Белла, $q_1 = (1/\alpha - 1) \mathbb{C}$, \mathbb{C} – постоянная Эйлера,

$$q_k = A_k \frac{\pi^k |B_k|}{k} + (1/\alpha^k - 1) \Gamma(k) \zeta(k), k \geq 2,$$

где B_k – числа Бернулли, $\zeta(k)$ – значение ζ -функции Римана в точке k , Γ – гамма-функция Эйлера, а числа A_k определяются следующим образом: $A_k = (2^k - 1)(1 - \theta^k)$ для моментов y_{0n} , $A_k = (1 - \theta^k)$ для моментов y_{1n} .

Заметим, что нас интересует только случай $\theta = 1$, что сильно упрощает выражение, так как все A_k тождественно равны нулю. Также заметим, что при таком значении параметра с. в. $Y(\alpha/2, 1)$ неотрицательна, поэтому под логарифмом можно оставить саму с. в. без модуля.

Таким образом, получаем выражение для n -го момента:

$$y_n = C_n(q_1, q_2, \dots, q_n),$$

где $C_n(q_1, q_2, \dots, q_n)$ – полиномы Белла порядка n ,

$$q_1 = \left(\frac{1}{\alpha} - 1\right) \mathbb{C}, \quad q_k = \left(\frac{1}{\alpha^k} - 1\right) \Gamma(k) \zeta(k), k \geq 2.$$

Оценим вероятность $\mathbb{P} \left(\sqrt{\frac{Y(\alpha/2, 1)}{Y'(\alpha/2, 1)}} \geq u^* \right)$ при $u^* > 1$:

$$\begin{aligned} \mathbb{P} \left(\sqrt{\frac{Y(\alpha/2, 1)}{Y'(\alpha/2, 1)}} \geq u^* \right) &= \mathbb{P} \left(\log \sqrt{\frac{Y(\alpha/2, 1)}{Y'(\alpha/2, 1)}} \geq \log u^* \right) \leq \\ &\leq \mathbb{P} \left(\left| \log \sqrt{\frac{Y(\alpha/2, 1)}{Y'(\alpha/2, 1)}} \right| \geq \log u^* \right) \leq \frac{\mathbb{E} \left(\left| \log \sqrt{\frac{Y(\alpha/2, 1)}{Y'(\alpha/2, 1)}} \right|^2 \right)}{(\log u^*)^2}. \end{aligned}$$

Рассмотрим $\mathbb{E} \left(\left| \log \sqrt{\frac{Y(\alpha/2, 1)}{Y'(\alpha/2, 1)}} \right|^2 \right)$. Обозначив $Y = Y(\alpha/2, 1)$, $Y' = Y'(\alpha/2, 1)$, проведем выкладки:

$$\begin{aligned} \mathbb{E} \left| \log \sqrt{\frac{Y}{Y'}} \right|^2 &= \mathbb{E} \left| \frac{1}{2} (\log Y - \log Y') \right|^2 = \\ &= \frac{1}{4} \mathbb{E} \left((\log Y)^2 + (\log Y')^2 - 2 \log Y \log Y' \right) = \\ &= \frac{1}{4} \left(\mathbb{E} (\log Y)^2 + \mathbb{E} (\log Y')^2 - 2 \mathbb{E} \log Y \log Y' \right) \\ &= \frac{1}{2} \mathbb{E} (\log Y)^2 - \frac{1}{2} (\mathbb{E} \log Y)^2 = \frac{1}{2} \mathbb{D} \log Y. \end{aligned}$$

Для завершения вычислений нам требуются первый и второй моменты с.в. $\log Y$. Получим их явные выражения, используя полиномы Белла

$$C_n(q_1, \dots, q_n) = \sum_{k_1+2k_2+\dots+k_n=n, k_j \geq 0} \frac{n!}{k_1! \dots k_n!} \left(\frac{q_1}{1!} \right)^{k_1} \dots \left(\frac{q_n}{n!} \right)^{k_n}.$$

Из определения следует, что первый момент с.в. $\log Y$ совпадает с q_1 . Рассмотрим

$$C_2(q_1, q_2) = \sum_{k_1+2k_2=2, k_j \geq 0} \frac{2!}{k_1! k_2!} \left(\frac{q_1}{1!} \right)^{k_1} \left(\frac{q_2}{2!} \right)^{k_2}.$$

Для второго момента подходят две комбинации индексов: $k_1 = 2, k_2 = 0$ и $k_1 = 0, k_2 = 1$. Таким образом,

$$C_2(q_1, q_2) = q_1^2 + q_2.$$

Учитывая, что $C_1(q_1) = q_1$, а в качестве первого параметра распределения с.в Y выступает не α , а $\alpha/2$, получим

$$\begin{aligned} \mathbb{E} \left| \log \sqrt{\frac{Y}{Y'}} \right|^2 &= \frac{1}{2} \mathbb{D} \log Y = \frac{1}{2} (q_1^2 + q_2 - q_1) = \frac{1}{2} \left\{ \left[\left(\frac{1}{\alpha/2} - 1 \right) \mathbb{C} \right]^2 + \right. \\ &\quad \left. + \left(\frac{1}{(\alpha/2)^2} - 1 \right) \Gamma(2) \zeta(2) - \left(\frac{1}{\alpha/2} - 1 \right) \mathbb{C} \right\} = \\ &= \frac{1}{2} \left\{ \left[\left(\frac{1}{\alpha/2} - 1 \right) \mathbb{C} \right]^2 + \left(\frac{1}{(\alpha/2)^2} - 1 \right) \frac{\pi^2}{6} - \left(\frac{1}{\alpha/2} - 1 \right) \mathbb{C} \right\}. \end{aligned}$$

Это выражение необходимо оценить сверху равномерно по всем α . Но при α , стремящемся к нулю справа, это выражение стремится к бесконечности, и эта оценка становится абсолютно бесполезной с практической точки зрения. С правой границей подобных проблем не возникает.

Таким образом, чтобы говорить о какой-либо оценке сверху для носителя смешивающего распределения для любого параметра, придется несколько ограничить рассматриваемый интервал возможных значений параметра. Будем считать, что $\alpha \in [\alpha_0, 2)$.

Соответственно, выбирая различные α_0 , получаем оценку сверху для $\mathbb{E} \left| \log \sqrt{\frac{Y}{Y'}} \right|^2$, при этом максимум достигается в точке α_0 , и чем дальше мы “отступим” от нуля, тем меньше получится оценка сверху.

Обозначим через f выражение для второго логарифмического момента смешивающего распределения:

$$f(\alpha) = \frac{1}{2} \left\{ \left[\left(\frac{1}{\alpha} - 1 \right) \mathbb{C} \right]^2 + \left(\frac{1}{\alpha^2} - 1 \right) \frac{\pi^2}{6} - \left(\frac{1}{\alpha} - 1 \right) \mathbb{C} \right\}$$

Суммируя все вышесказанное, имеем искомую оценку сверху

$$\mathbb{P} \left(\sqrt{\frac{Y(\alpha/2, 1)}{Y'(\alpha/2, 1)}} \geq u^* \right) \leq \frac{f(\alpha)}{(\log u^*)^2} \leq \frac{f(\alpha_0)}{(\log u^*)^2} = \varepsilon.$$

Отсюда можно выразить

$$u^* = \exp \left\{ \sqrt{\frac{f(\alpha_0)}{\varepsilon}} \right\}.$$

6. Описание результатов численных экспериментов оценивания параметров распределений Линника и Миттаг-Леффлера

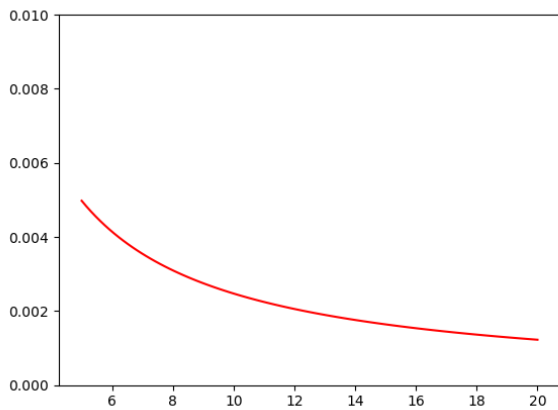
Численные результаты. Далее описаны численные эксперименты, в ходе которых генерировались выборки из распределений Линника и Миттаг-Леффлера с параметрами α_{true} и δ_{true} . После чего комбинированным сеточным методом вычислялись оценки $\hat{\alpha}$ и $\hat{\delta}$ и сравнивались с истинным значением параметра. Кроме того, полученные оценки сравниваются с оценками полученными методом логарифмическим моментом.

Для получения адекватной теоретической верхней оценки для носителя смешивающего распределения потребовалось сузить рассматриваемый отрезок возможных значений параметра до промежутка $[\alpha_0, 2)$.

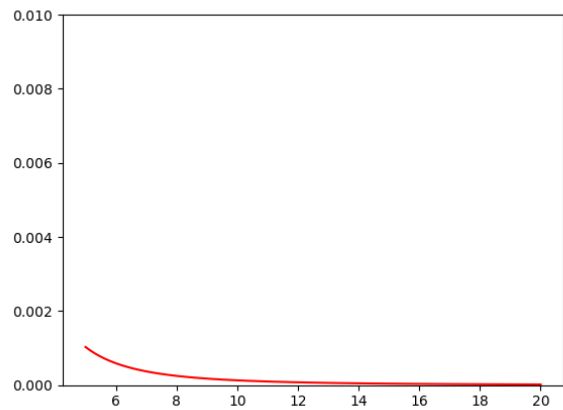
При этом в силу экспоненциальной зависимости при уменьшении α_0 верхняя граница сетки очень быстро растет, и алгоритм начинает работать заметно хуже в силу разрежения сетки. Например, если подставить значения $\alpha_0 = 0.1, \varepsilon = 0.01$, то оценка получается порядка 10^{84} , что не дает реально применимого практического результата. Первые адекватные оценки порядка 10^2 получаются лишь при сильном ухудшении точности и ограничении отрезка, например, при $\alpha_0 = 1, \varepsilon = 0.1$, что, очевидно, ограничивает рассматриваемый отрезок почти в два раза, и при этом используется меньшая вероятность.

Поэтому для практической работы алгоритма была выбрана “стартовая” длина отрезка, равная 30, и реализован механизм динамического изменения правой границы сетки в зависимости от полученных оценок весов: если веса на правой границе после остановки итерационного процесса являются слишком большими, то считается, что был взят слишком маленький интервал, и граница сдвигается вправо. И наоборот, если веса являются чересчур малыми, что характерно для больших значений α_{true} , граница сдвигается влево, и интервал становится короче.

Особенности выбора сетки. Так как заранее не делается никаких предположений о значении α_{true} , то выбор подходящих границ является очень существенным. Например, при малых значениях α в распределении Линника хвосты смешивающего распределения являются гораздо более тяжелыми, чем при значениях α , близких к двум (см. рис. 1). Поэтому для первого случая необходим более широкий интервал носителя узлов, для второго же случая такой подход отрицательно скажется на точности оценки, так как большинство узлов сетки будут содержать малые веса, и основная масса плотности окажется распределена буквально между несколькими узлами в начале сетки. Поэтому было реализовано динамическое из-



(a) $\alpha = 0,1$



(b) $\alpha = 1,9$

Рис. 1: Сравнение хвостов распределения

менение сетки, при котором ее правый конец сдвигался влево или вправо соответственно при слишком малых или слишком больших весах на крайних узлах сетки.

Еще более важным является распределение узлов по выделенному промежутку. Было опробовано несколько вариантов расположения узлов.

- Равномерная сетка. Все узлы располагаются на одинаковом расстоянии друг от друга. Наиболее простой в реализации и стабильный вариант.
- “Параболическая” сетка. Расстояние между узлами увеличивается при удалении от левого конца сетки по параболическому закону (см. рис. 2). Идея использования степенной функции, а именно параболы, возникла не только из простоты реализации, но также из поведения хвостов распределения, которые убывают с подобной скоростью, или медленнее, в зависимости от значения параметра α_{true} .
- “Интегральная” сетка. Количество узлов сетки на разных ее кусках пропорционально плотности смешивающего распределения с определенным параметром (см. рис. 3). На первой итерации параметр брался равным 1 для распределения Линника и 0.5 для распределения Миттаг-Леффлера, как середины отрезка возможного значения параметра. Далее на каждой итерации сетка перестраивалась с использованием в качестве параметра построения оценки на параметр распределения, полученной на предыдущей итерации.

Также важным вопросом является соотношение между размером выборки и количеством узлов сетки. При слишком большом числе узлов проявляет-

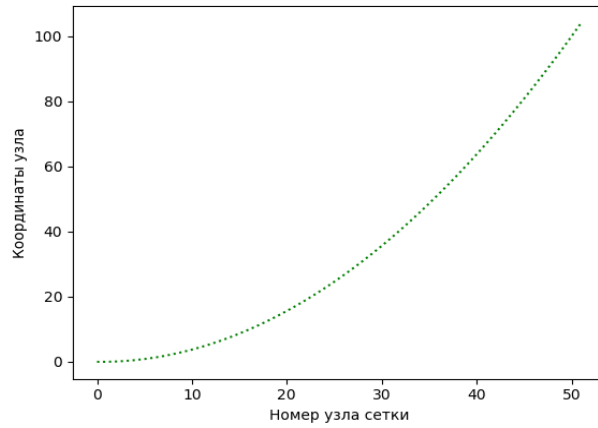


Рис. 2: Координаты узлов в "параболической" сетке

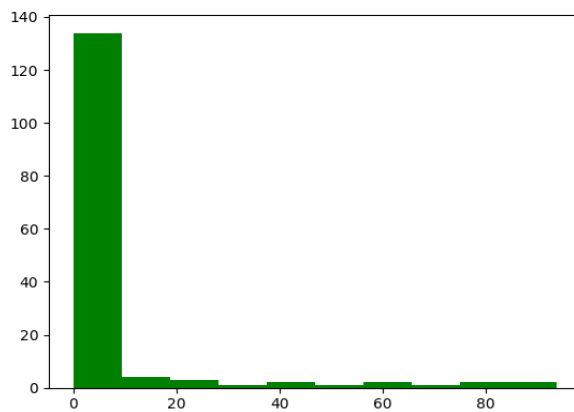


Рис. 3: Распределение количества узлов по отрезкам в "интегральной" сетке

ся эффект переобучения алгоритма в виде резких скачков весов на узлах сетки (см. рис. 4).

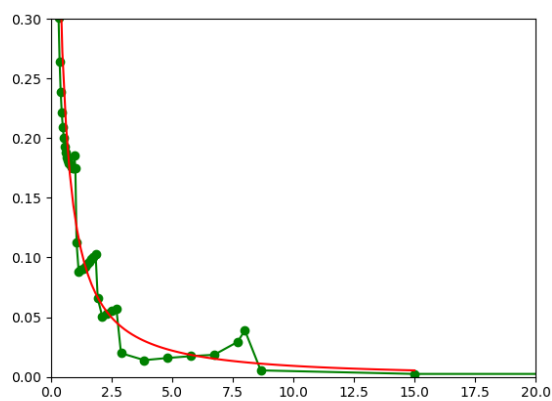


Рис. 4: Эффект переобучения алгоритма оценивания: на графике реальная плотность при заданном $\alpha_{true} = 0.5$ и полученная оценка весов

По-видимому, это можно объяснить тем, что алгоритму не хватает информации, заложенной в выборке, чтобы адекватно сформировать оценки во всех узлах. При слишком малом числе узлов же оказываются непокрытыми точки, ярко характеризующие форму кривой плотности, и точность оценки также падает.

Результаты имитационных вычислений. В табл. 1 и 2 представлены результаты работы сеточного метода для выборок из распределений Линника и Миттаг-Леффлера соответственно. Таблицы иллюстрируют работу алгоритма при различных истинных значениях параметра α_{true} на выборках разных размеров.

Значение α	Измерение	n = 100	n=500	n = 1000
0.1	Mean	0.063	0.064	0.056
	Var	0.0008048770	0.0002925522	0.0000951870
0.2	Mean	0.128	0.116	0.101
	Var	0.0063894166	0.0016142483	0.0003436977
0.3	Mean	0.223	0.177	0.199
	Var	0.0460641238	0.0025049852	0.0190297301
0.4	Mean	0.512	0.447	0.370
	Var	0.1097217963	0.0641530507	0.0584024453
0.5	Mean	0.592	0.671	0.639
	Var	0.0602995000	0.0391031986	0.0197545785
0.6	Mean	0.627	0.700	0.660
	Var	0.0794097580	0.0195715662	0.0076625768
0.7	Mean	0.746	0.725	0.713
	Var	0.0254652823	0.0135295523	0.0060995062
0.8	Mean	0.825	0.780	0.803
	Var	0.0263811490	0.0075966112	0.0025345610
0.9	Mean	0.954	0.887	0.891
	Var	0.0239822404	0.0048354797	0.0040002054
1.0	Mean	0.984	0.960	0.963
	Var	0.0154119640	0.0044957931	0.0027907073
1.1	Mean	1.127	1.043	1.068
	Var	0.0456481297	0.0050894835	0.0021984667
1.2	Mean	1.135	1.168	1.181
	Var	0.0355432149	0.0057725778	0.0019296589
1.3	Mean	1.313	1.277	1.264
	Var	0.0302404507	0.0044318573	0.0015922114
1.4	Mean	1.333	1.393	1.393
	Var	0.0276904252	0.0069074761	0.0016463854
1.5	Mean	1.478	1.483	1.479
	Var	0.0490600413	0.0049338240	0.0035612724
1.6	Mean	1.581	1.569	1.580
	Var	0.0192754862	0.0041105830	0.0036814442
1.7	Mean	1.684	1.682	1.678
	Var	0.0301132555	0.0030556651	0.0019111498
1.8	Mean	1.725	1.745	1.731
	Var	0.0129253268	0.0041883564	0.0011024229
1.9	Mean	1.763	1.812	1.820
	Var	0.0147540639	0.0015007617	0.0005750203

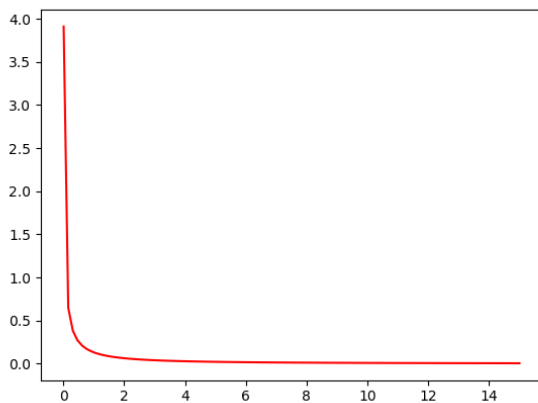
Таблица 1: Численные результаты для распределения Линника

Значение δ	Измерение	n = 100	n=500	n = 1000
0.1	Mean	0.125	0.087	0.096
	Var	0.0023854639	0.0005231545	0.0003481034
0.2	Mean	0.183	0.193	0.186
	Var	0.0070899655	0.0011037209	0.0010806788
0.3	Mean	0.300	0.316	0.299
	Var	0.0125653809	0.0024390300	0.0017165859
0.4	Mean	0.414	0.422	0.422
	Var	0.0173173798	0.0042152216	0.0009098602
0.5	Mean	0.502	0.533	0.529
	Var	0.0231448890	0.0036007410	0.0021370758
0.6	Mean	0.578	0.614	0.608
	Var	0.0131584532	0.0031056887	0.0022612102
0.7	Mean	0.657	0.664	0.681
	Var	0.0115818330	0.0021665299	0.0014961482
0.8	Mean	0.769	0.775	0.774
	Var	0.0055079200	0.0012874792	0.0007485115
0.9	Mean	0.865	0.863	0.870
	Var	0.0026004329	0.0011965831	0.0002720375

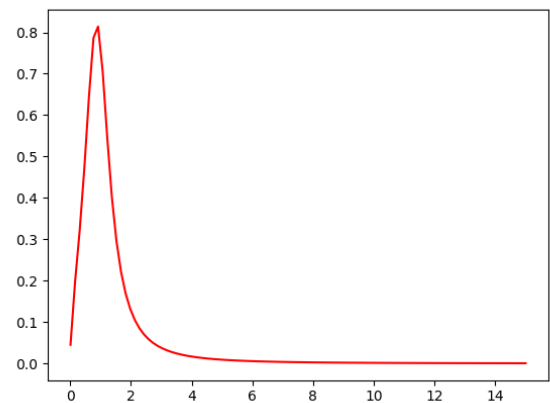
Таблица 2: Численные результаты для распределения Миттаг-Леффлера

Далее мы приведем некоторые графики, иллюстрирующие работу алгоритма. Но сначала обратим внимание на типичный вид смешивающего распределения.

Стоит заметить, что типичный вид плотности смешивающего распределения выглядит по-разному для разных промежутков значения α_{true} : при $\alpha_{true} < 1$ плотность имеет пик в нуле, при $\alpha_{true} > 1$ его нет, и плотность имеет характерный вид пика недалеко от нуля, а затем убывания с тяжелым хвостом. $\alpha_{true} = 1$ является граничным случаем (см. рис. 5). Это означает, что алгоритм должен адаптировать свое поведение, чтобы корректно обрабатывать все случаи.



(a) $\alpha = 0.5$



(b) $\alpha = 1.5$

Рис. 5: Вид плотности смешивающего распределения

Рассмотрим типичные иллюстрации работы итерационного алгоритма для разных значений α_{true} . На рис. 6 изображены истинная плотность смешивающего распределения (при значении параметра α_{true}) и плотность со значением параметра $\hat{\alpha}$, полученным в результате оценки сеточным методом.

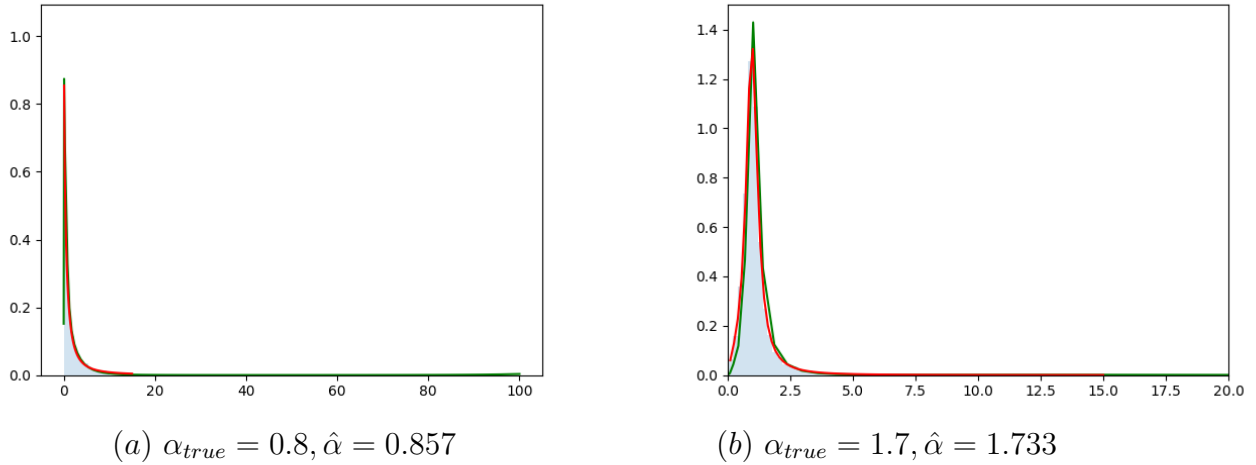


Рис. 6: Результат работы алгоритма

Значение α	Показатель	Сеточный метод с неравномерной сеткой	Метод логарифмических моментов
0.1	Mean	0.1052813	0.1002494
	StD	0.0490361	0.0041889
0.3	Mean	0.2941098	0.2982842
	StD	0.0520982	0.01106160
0.5	Mean	0.4931283	0.5008472
	StD	0.0608762	0.0197338
0.7	Mean	0.6713201	0.7004368
	StD	0.0390137	0.0235764
0.9	Mean	0.8540187	0.8978184
	StD	0.0251179	0.0266253

Таблица 3: Сравнение качества оценок параметра распределения Миттаг-Леффлера, полученных сеточным методом и методом логарифмических моментов.

В табл. 3 приведены результаты вычисления оценок параметра распределения Миттаг-Леффлера, полученных сеточным методом и методом логарифмических моментов на имитированных выборках с известными истинными значениями параметра α_{true} . Видно, что методы дают примерно одинаковые результаты, однако при α_{true} , близких к единице, сеточный метод оказывается несколько точнее метода логарифмических моментов: среднеквадратичное отклонение оценок, полученных этим методом, становится меньше.

7. Заключение

Описан двухэтапный сеточный метод оценивания параметров распределений Линника и Миттаг-Леффлера. Приведены результаты сравнения качества работы метода с разными подходами к выбору сетки, а также с методом логарифмических моментов. В результате были получены следующие выводы:

- При небольших значениях параметра α_{true} параболическая сетка с числом узлов порядка \sqrt{n} , где n – размер выборки, дает лучшие результаты, чем равномерная сетка. Для распределения Линника наблюдается смещенность оценок: это является открытой проблемой и направлением для дальнейших исследований.
- При больших значениях α_{true} лучше работает плотная равномерная сетка с большим числом узлов. На некоторых значениях удавалось получить результаты, превосходящие по точности оценки, полученные методом логарифмических моментов.
- В методе с “интегральной” сеткой требуется каждый раз решать k уравнений вида

$$\int_{u_{i-1}}^{u_i} p_{\alpha}(x) dx = \frac{1}{k}$$

для последовательного поиска узлов сетки u_1, \dots, u_k . В целях ускорения работы алгоритма плотность приближалась гистограммой, построенной по специально генерирующейся выборке размера 1000 с заданным параметром, затем желаемый отрезок для сетки разбивался на несколько частей, на каждый из которых выделялось количество узлов, пропорциональное площади гистограммы, покрывающей этот интервал. Этот способ выбора сетки при некоторых значениях параметра позволял добиваться большей точности, чем при равномерной и параболической сетках, но при таком варианте алгоритм ведет себя крайне нестабильно и склонен сходиться к границам области определения параметра, поэтому в целях достижения стабильной сходимости приоритет отдавался равномерной и параболической сетке.

В случае распределения Миттаг-Леффлера лучшие результаты были получены с использованием параболической сетки.

В случае распределения Линника (при условии отсутствия каких-либо начальных предположений о значении параметра) лучшие результаты на текущий момент были получены путем создания “гибридного” алгоритма,

сочетающего в себе стабильность равномерной сетки, и “гибкость” параболической сетки в окрестности нуля: на первом проходе алгоритма делается грубая оценка параметра, чтобы сделать вывод, в какой отрезок, скорее всего, попадает α_{true} : в $(0, 1.0]$ или $(1.0, 2]$. Далее применяется сеточный метод с использованием параболической сетки с динамическим изменением верхней границы либо плотной равномерной сетки с фиксированным отрезком соответственно.

Библиографический список

1. *Линник Ю. В.* Линейные формы и статистические критерии. I, II // Украинский математический журнал. – 1953. – Т. 5, Вып. 2. – С. 207–243; Вып. 3. – С. 247–290.
2. *Золотарев В. М.* Одномерные устойчивые распределения. – М.: Наука, 1983. – 304с.
3. *Anderson D. N.* A multivariate Linnik distribution // Statistics and Probability Letters. – 1992. – Vol. 14. – P. 333–336.
4. *Bunge J.* Compositions semigroups and random stability // Annals of Probability. – 1996. – Vol. 24. – P. 1476–1489.
5. *Cahoy D. O.* An estimation procedure for the Linnik distribution // Statistical Papers. – 2012. – Vol. 53, №3. – P. 617–628.
6. *Cahoy D. O.* Estimation of Mittag-Leffler Parameters // Communications in Statistics – Simulation and Computation. – 2013. – Vol. 42, №2. – P. 303–315.
7. *Devroye L.* A note on Linnik’s distribution // Statistics and Probability Letters. – 1990. – Vol. 9. – P. 305–306.
8. *Gorenflo R., Mainardi F.* Continuous time random walk, Mittag-Leffler waiting time and fractional diffusion: mathematical aspects / Chap. 4 in *Klages R., Radons G. and Sokolov I. M. (Editors)*. Anomalous Transport: Foundations and Applications. – Weinheim, Germany: Wiley-VCH, 2008. – P. 93–127.
Available at: <http://arxiv.org/abs/0705.0797>.
9. *Gorenflo R., Kilbas A. A., Mainardi F., Rogosin S. V.* Mittag-Leffler Functions, Related Topics and Applications. – Berlin-New York: Springer, 2014.
10. *Jacques C., Rémillard B., Theodorescu R.* Estimation of Linnik law parameters // Statistics and Decision. – 1999. – Vol. 17, №3. – P. 213–236.
11. *Jose K. K., Uma P., Lekshmi V. S., Haubold H. J.* Generalized Mittag-Leffler Distributions and Processes for Applications in Astrophysics and Time Series Modeling // Astrophysics and Space Science Proceedings. – 2010. – Iss. 202559. – P. 79–92.

12. *Klebanov L. B., Maniya G. M., Melamed I. M.* A problem of Zolotarev and analogs of infinite divisible and stable distributions in a scheme for summing a random number of random variables // *Theory Prob. Appl.* – 1985. – Vol. 29. – P. 791–794.
13. *Korolev V. Yu., Zeifman A. I.* A note on mixture representations for the Linnik and Mittag-Leffler distributions and their applications // *Journal of Mathematical Sciences.* – 2017. – Vol.218, №3. – P. 314–327.
14. *Korolev V. Yu., Zeifman A. I.* Convergence of statistics constructed from samples with random sizes to the Linnik and Mittag-Leffler distributions and their generalizations // *Journal of Korean Statistical Society.* – 2017. – Vol.46, №2. P. 161–181. Available online 25 July 2016. Also available on arXiv:1602.02480v1 [math.PR].
15. *Kotz S., Ostrovskii I. V., Hayfavi A.,* Analytic and asymptotic properties of Linnik’s probability densities, I // *Journal of Mathematical Analysis and Applications.* – 1995. – Vol. 193. – P. 353–371.
16. *Kotz S., Ostrovskii I. V., Hayfavi A.,* Analytic and asymptotic properties of Linnik’s probability densities, II // *Journal of Mathematical Analysis and Applications.* – 1995. – Vol. 193. – P. 497–521.
17. *Kotz S., Ostrovskii I.V.* A mixture representation of the Linnik distribution // *Statistics and Probability Letters.* – 1996. – Vol. 26. – P. 61–64.
18. *Kotz S., Kozubowski T.J., Podgorski K.* *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance.* – Boston: Birkhauser, 2001.
19. *Kozubowski T.J.* Fractional moment estimation of Linnik and Mittag-Leffler parameters // *Math. Comput. Model.* – 2001. – Vol. 34. – P. 1023–1035.
20. *Lin G.D.* Characterizations of the Laplace and related distributions via geometric compound // *Sankhya, A1.* – 1994. – Vol. 56. – P. 1–9.
21. *Mathai A. M., Haubold H. J.* Matrix-variate statistical distributions and fractional calculus // *Fractional Calculus & Applied Analysis.* An International Journal for Theory and Applications. – 2011. – Vol. 14, №1. – P. 138–155.
22. *Mitnik S., Rachev S.T.* Modeling asset returns with alternative stable distributions // *Econometric Rev.* – 1993. – Vol. 12. – P. 261–330.
23. *Pakes A.G.* Mixture representations for symmetric generalized Linnik laws // *Statistics and Probability Letters.* – 1998. – Vol. 37. – P. 213–221.
24. *Pillai R.N.* Semi- α -Laplace distributions // *Communications in Statistical Theory and Methods.* – 1985. – Vol.14. – P. 991-1000.

Combined grid method for estimation of the parameters of Linnik and Mittag-Leffler distributions¹

V. Yu. Korolev², D. V. Fomina³, A. A. Shcherbinina⁴

²Faculty of Computational Mathematics and Cybernetics of Moscow State University;
vkorolev@cs.msu.su

³Faculty of Computational Mathematics and Cybernetics of Moscow State University

⁴Faculty of Computational Mathematics and Cybernetics of Moscow State University

Abstract. *We propose the combined grid method of statistical estimation of parameters of Linnik and Mittag-Leffler distributions. The densities of these distributions cannot be explicitly represented in terms of elementary functions which makes maximum likelihood method inapplicable. Moreover the pure method of moments also cannot be applied here due to heavy tails of the distributions. The proposed grid method is based on the representation of Linnik and Mittag-Leffler distributions respectively in the form of a scale mixture of the Laplace distribution and mixed exponential distribution. In both cases, mixing distribution is the same and corresponds to the ratio of two independent one-sided strictly stable random variables with the same characteristic indicator. The density of this distribution can be expressed in terms of elementary functions and has a rather simple explicit form. The first step of the proposed method is to find discrete approximation of the specified mixing distribution using the version of the EM algorithm. After that we estimate the parameter of the analyzed distribution by minimizing the residual between the obtained discrete estimation and theoretical mixing density. Different methods of selecting appropriate grid are investigated. The results of simulations illustrate method convergence and the estimates are compared to the ones obtained with the method of logarithmic moments.*

Key words: *Linnik distribution; Mittag-Leffler distribution; scale mixture; EM-algorithm; logarithmic moment.*

УДК 519.2

Скорость сходимости ядерных оценок плотности в пространствах произвольной природы

А. И. Орлов

Московский государственный технический университет им. Н.Э. Баумана, Институт высоких статистических технологий и эконометрики; Московский физико-технический институт; Москва, Россия; prof-orlov@mail.ru; <http://orlovs.pp.ru>; +7(916)8305117; 123104, Москва, Сытинский пер., д.7/14, кв.14.

Аннотация. *Непараметрические оценки плотности распределения вероятностей в пространствах произвольной природы – один из основных инструментов нечисловой статистики. Рассмотрены их частные случаи – ядерные оценки плотности в пространствах произвольной природы. Цель настоящей статьи – изучение среднего квадрата ошибки ядерной оценки плотности и – с целью максимизации порядка его убывания – выбор ядерной функции и последовательности показателей размытости. Основные понятия – круговая функция распределения и круговая плотность. Порядок сходимости в общем случае тот же, что и при оценивании плотности числовой случайной величины, но основные условия наложены не на плотность случайной величины, а на круговую плотность.*

Ключевые слова: *нечисловая статистика, плотность распределения вероятностей, пространства произвольной природы, ядерные оценки плотности, предельные теоремы, скорость сходимости*

1. Введение

Согласно новой парадигме прикладной математической статистики [1], сердцевиной этой научной области является статистическая теория в пространствах произвольной природы. Эти пространства не предполагаются линейными. Подходы и результаты статистики в пространствах произвольной природы могут применяться как при анализе числовых данных, так и нечисловых (бинарных отношений, множеств и др.). Статистическая теория в пространствах произвольной природы и статистические методы анализа конкретных нечисловых данных выделены в 1979 г. как самостоятельная область прикладной математической статистики [2, 3]. Она была названа статистикой объектов нечисловой природы. Позже ее стали называть статистикой нечисловых данных или нечисловой статистикой [4, 5]. Непараметрические оценки плотности распределения вероятностей в

пространствах произвольной природы – один из основных инструментов нечисловой статистики. Систематическое изложение теории таких оценок начато в статьях [6 – 8], непосредственным продолжением которых является настоящая статья. Регулярно используются ссылки на условия и утверждения из статей [6 – 8].

Введем обозначения. Пусть (Z, \mathcal{A}) – измеримое пространство, p и q – сигма-конечные меры на (Z, \mathcal{A}) , причем p абсолютно непрерывна относительно q , т.е. из $q(\mathcal{B}) = 0$ следует $p(\mathcal{B}) = 0$ для любого множества \mathcal{B} из сигма-алгебры \mathcal{A} . В этом случае на (Z, \mathcal{A}) существует неотрицательная измеримая функция $f(x)$ такая, что

$$q(\mathcal{C}) = \int_{\mathcal{C}} f(x)p(dx) \quad (1)$$

для любого множества \mathcal{C} из сигма-алгебры измеримых множеств \mathcal{A} . Функция $f(x)$ называется производной Радона – Никодима меры q по мере p , а в случае, когда q – вероятностная мера, также плотностью вероятности q по отношению к мере p [9, с.460].

Пусть X_1, X_2, \dots, X_n – независимые одинаково распределенные случайные элементы (величины), распределение которых задается вероятностной мерой q . В статье [6] введено несколько видов непараметрических оценок плотности вероятности q по выборке X_1, X_2, \dots, X_n , подробнее изучены линейные оценки. В статье [7] рассмотрены их частные случаи – ядерные оценки плотности в пространствах произвольной природы. В статье [8] асимптотическая теория ядерных оценок плотности развита прежде всего для нужд статистики конкретных видов объектов нечисловой природы, в которой основной интерес представляют конечные пространства Z . Мера p при этом не непрерывная, а дискретная, например, считающая. Таким образом, в рамках единого подхода удастся рассмотреть оценки плотностей и оценки вероятностей.

В предположении непрерывности неизвестной плотности $f(x)$ представляется целесообразным «размазать» каждый атом эмпирической меры, т.е. рассмотреть линейные оценки, введенные в нашей первой работе по нечисловой статистике [3, с.24]:

$$f_n(x) = \frac{1}{n} \sum_{1 \leq i \leq n} g_n(x, X_i), \quad g_n : Z^2 \rightarrow R^1, \quad (2)$$

в которых действительнзначные функции g_n удовлетворяют некоторым условиям регулярности.

Пусть d – показатель различия на Z [4] (в наиболее важных частных случаях – метрика на Z). В [10] нами введены ядерные оценки плотности

– оценки вида (2) с

$$g_n(x, X_i) = \frac{1}{b(h_n, x)} K \left(\frac{d(x, X_i)}{h_n} \right), \quad K : [0, +\infty) \rightarrow R^1, \quad (3)$$

$K = K(u)$ – ядро (ядерная функция), h_n – последовательность положительных чисел (показателей размытости), $b(h_n, x)$ – нормировочный множитель. В [11] линейные оценки (2) с функциями g_n из (3) названы нами «обобщенными оценками типа Парзена–Розенблатта», т.к. в частном случае $Z = R^1$, $d(x, X_i) = |x - X_i|$, $b(h_n, x) = h_n$ они переходят в известные оценки, введенные Розенблаттом [12] и Парзеном [13].

Цель настоящей статьи – изучение среднего квадрата ошибки ядерной оценки плотности

$$\alpha_n = M(f_n(x) - f(x))^2 = (Mf_n(x) - f(x))^2 + Df_n(x) \quad (4)$$

и выбор последовательности h_n с целью максимизации порядка убывания α_n при $n \rightarrow \infty$. Здесь и далее M – символ математического ожидания, D – символ дисперсии.

2. Круговая функция распределения

Пусть справедливы следующие условия регулярности:

R1) плотность $f(x)$ непрерывна в точке x , в которой оцениваем плотность;

R2) мера p согласована с показателем различия d , т.е. мера шара радиуса t равна t ,

$$p\{y : d(x, y) < t\} = t, \quad 0 \leq t \leq t_1 = p(Z), \quad (5)$$

другими словами, показатель различия является предпочтительным [7],

R3) ядро $K(u)$ – непрерывная финитная функция, $K(u) = 0$ при $u > E$, такая, что

$$\int_0^E K(u) du = \int_0^\infty K(u) du = 1. \quad (6)$$

Введем в рассмотрение шары $L_t(x) = \{y : d(x, y) < t\}$ радиуса t и аналог функции распределения случайной величины X со значениями в Z с плотностью $f(x)$:

$$G(x, t) = P\{X \in L_t(x)\} = \int_{L_t(x)} f(y) p(dy). \quad (7)$$

Назовем $G(x, t)$ круговой функцией распределения в точке x . Все нужные в дальнейшем свойства вероятностной модели, как будет показано, выражаются с помощью круговой функции распределения.

Из непрерывности плотности в точке x и равенства (5) следует, что круговая функция распределения $G(x, t)$ дифференцируема по t при $t = 0$ и

$$G'_t(x, 0) = f(x). \quad (8)$$

Изучим смещение оценки $f_n(x)$. Имеем цепочку равенств

$$\begin{aligned} Mf_n(x) &= \frac{1}{h_n} \int_Z K\left(\frac{d(x, y)}{h_n}\right) f(y)p(dy) = \frac{1}{h_n} \int_0^\infty K\left(\frac{t}{h_n}\right) dG(x, t) = \\ &= \int_0^E K(u) \frac{dG(x, h_n u)}{h_n}. \end{aligned} \quad (9)$$

Пусть справедливо следующее условие.

R4) В некоторой окрестности точки $t = 0$ (т.е. при $0 \leq t \leq t_0$ при некотором t_0) существует производная по t круговой функцией распределения $G(x, t)$, т.е. $G'_t(x, t) = g(x, t)$.

Тогда при $h_n E < t_0$ имеем

$$Mf_n(x) = \int_0^E K(u) g(x, h_n u) du. \quad (10)$$

Основная идея дальнейших рассуждений состоит в том, чтобы для изучения скорости сходимости ядерных оценок применить разложение $g(x, t)$ в ряд по степеням t в окрестности $t = 0$ (в предположении существования указанного разложения).

R5) Пусть справедливо разложение

$$g(x, t) = g(x, 0) + tg'_t(x, 0) + \frac{t^2}{2} g''_{tt}(x, 0) + o(h_n^2), \quad 0 \leq t \leq h_n E \quad (11)$$

(согласно (8) $g(x, 0) = f(x)$). Подставим это разложение в (10), получим с учетом (6) и непрерывности функции $K(u)$:

$$Mf_n(x) = f(x) + h_n g'_t(x, 0) \int_0^E u K(u) du + h_n^2 g''_{tt}(x, 0) \int_0^E u^2 K(u) du + o(h_n^2). \quad (12)$$

3. Первые оценки скорости сходимости

Согласно теореме 7 статьи [7] при справедливости рассматриваемых условий

$$Df_n(x) = \frac{f(x)}{nh_n} \int_0^E K^2(u)du + o\left(\frac{1}{nh_n}\right). \quad (13)$$

Если

$$a = g'_t(x, 0) \int_0^E uK(u)du \neq 0, \quad (14)$$

то средний квадрат ошибки ядерной оценки плотности (4) согласно (12) и (13) равен

$$\alpha_n = h_n^2 a^2 + \frac{f(x)}{nh_n} \int_0^E K^2(u)du + o\left(h_n^2 + \frac{1}{nh_n}\right). \quad (15)$$

Следовательно, при $f(x) \neq 0$ оптимальное по порядку скорости сходимости значение h_n определяется из условия "уравнивания погрешностей" [2]

$$h_n^2 = \frac{1}{nh_n}, \quad h_n = n^{-1/3}. \quad (16)$$

Тогда, как легко видеть, средний квадрат ошибки ядерной оценки плотности имеет порядок " n в степени $(-2/3)$ ":

$$\alpha_n \cong Cn^{(-2/3)} \quad (17)$$

при некоторой константе C .

Обоснуем сказанное более подробно. С точностью до бесконечно малых более высокого порядка средний квадрат ошибки ядерной оценки плотности равен

$$B(h) = a^2 h^2 + \frac{F}{nh}, \quad F = f(x) \int_0^E K^2(u)du, \quad h = h_n. \quad (18)$$

С целью решения задачи оптимизации $B(h) \rightarrow \min$ вычислим производную $B(h)$ по h и приравняем ее 0:

$$B'(h) = 2a^2 h - \frac{F}{nh^2} = 0. \quad (19)$$

Решая уравнение (19) относительно h , получаем, что

$$h = h_n = \left(\frac{F}{2a^2n} \right)^{1/3} = \left(\frac{F}{2a^2} \right)^{1/3} n^{-1/3}. \quad (20)$$

Соотношение (20) уточняет ранее полученное соотношение (17). Для частного случая $Z = R^1$, т.е. для оценок Парзена-Розенблатта, соотношения (16) – (17) известны (см. [14, с.315]).

Если соотношение (14) не выполнено, т.е. $a = 0$, то согласно (12) заключаем, что

$$\alpha_n = [g''_{tt}(x, 0)]^2 \left[\int_0^E u^2 K(u) du \right]^2 + \frac{f(x)}{nh_n} \int_0^E K^2(u) du + o\left(h_n^4 + \frac{1}{nh_n}\right). \quad (21)$$

Оптимальное по порядку сходимости h_n определяется из условия

$$h_n^4 = \frac{1}{nh_n}, h_n = n^{-1/5}, \quad (22)$$

и при этом

$$\alpha_n \cong C_1 n^{-4/5} \quad (23)$$

при некоторой константе C_1 .

Соотношения (22) – (23) совпадают с известными результатами для весьма частного случая $Z = R^1$, т.е. для оценок Парзена-Розенблатта (см. [14, с. 316]).

4. Примеры

Из сравнения формул (17) и (23) ясно, что сходимость убыстряется при $a = 0$, где a определено в (14). Поскольку α – произведение двух сомножителей, то $a = 0$ тогда и только тогда, когда $g'_t(x, 0) = 0$ или

$$\int_0^E uK(u)du = 0. \quad (24)$$

Роль сомножителей разная: первый определяется свойствами пространства с мерой, показателя различия (другими словами, меры близости) и плотности распределения случайной величины, а второй – свойствами ядра.

С целью выявления свойств первого сомножителя рассмотрим два примера.

Пример 1. Рассмотрим множество действительных чисел $Z = R^1$. Пусть p – мера Лебега, d – расстояние Евклида (показатель различия),

$F(x)$ – функция распределения случайной величины X , причем ее плотность f (по мере Лебега, т.е. в обычном смысле) дважды непрерывно дифференцируема. Тогда

$$G(x, t) = P \left\{ X \in \left(x - \frac{t}{2}, x + \frac{t}{2} \right) \right\} = F \left(x + \frac{t}{2} \right) - F \left(x - \frac{t}{2} \right), \quad (25)$$

а потому

$$g(x, t) = \frac{1}{2} \left\{ f \left(x + \frac{t}{2} \right) + f \left(x - \frac{t}{2} \right) \right\}. \quad (26)$$

Продифференцируем (26):

$$g'(x, t) = \frac{1}{4} \left\{ f' \left(x + \frac{t}{2} \right) - f' \left(x - \frac{t}{2} \right) \right\}. \quad (27)$$

При $t = 0$ при всех x

$$g'_t(x, 0) = 0. \quad (28)$$

Пример 2. Рассмотрим $Z = [0, +\infty)$. Пусть мера p и расстояние d получены из рассмотренных в примере 1 сужением на $[0, +\infty)$. Пусть функция распределения и плотность обладают теми же свойствами, что и в примере 1. Тогда

$$G(0, t) = F(t), \quad g(0, t) = f(t), \quad g'_t(0, t) = f'(t). \quad (29)$$

Следовательно, первый сомножитель в (14), вообще говоря, отличен от 0.

Можно показать, что для конечномерного пространства $Z = R^k$, меры Лебега p , евклидова расстояния d и дважды дифференцируемой плотности f первый сомножитель в (14) обращается в 0, а для мер p , отличных от Лебеговой, вообще говоря, не обращается (при сохранении прочих перечисленных в примерах 1 и 2 условий).

Из сказанного следует, что для введенных нами оценок (2) – (3) в случае конечномерного пространства $Z = R^k$, меры Лебега p , евклидова расстояния d и дважды дифференцируемой плотности f скорость сходимости задается формулой (23), а для классических оценок Парзена–Розенблатта – формулой (17) (см. также [14, с. 315-316]), т.е. введенные нами оценки сходятся гораздо быстрее, чем оценки Парзена–Розенблатта.

5. Улучшение скорости сходимости

Поскольку статистик может сам выбирать ядро, то для повышения скорости сходимости целесообразно принять условие (24). Однако скорость сходимости можно еще более повысить за счет выбора ядра из более узкого класса.

При более высокой гладкости круговой плотности $g(x, t)$ можно получить более высокую скорость сходимости среднего квадрата ошибки ядерной оценки плотности α_n , соответствующим образом выбирая ядро $K(u)$. Предположим, что круговая плотность $g(x, t)$ допускает разложение

$$g(x, t) = f(x) + tg'_t(x, 0) + \frac{t^2}{2}g''_{tt}(x, 0) + \frac{t^3}{3!}g'''_{ttt}(x, 0) + \dots + \frac{t^k}{k!}g^{(k)}_{t^{(k)}}(x, 0) + o(h_n^k), \quad (30)$$

причем остаточный член равномерно ограничен на отрезке $[0, h_n E]$. Тогда

$$Mf_n(x) = f(x) + h_n g'_t(x, 0) \int_0^E u K(u) du + \frac{h_n^2}{2} g''_{tt}(x, 0) \int_0^E u^2 K(u) du + \frac{h_n^3}{3!} g'''_{ttt}(x, 0) \int_0^E u^3 K(u) du + \dots + \frac{h_n^k}{k!} g^{(k)}_{t^{(k)}}(x, 0) \int_0^E u^k K(u) du + \theta_n h_n^k, \quad (31)$$

где $\theta_n \rightarrow 0$ при $n \rightarrow \infty$.

Пусть теперь

$$\int_0^E u^i K(u) du = 0, \quad i = 1, 2, \dots, k-1. \quad (32)$$

Тогда

$$Mf_n(x) - f(x) = \frac{h_n^k}{k!} g^{(k)}_{t^{(k)}}(x, 0) \int_0^E u^k K(u) du + o(h_n^k) \quad (33)$$

Следовательно,

$$\alpha_n = h_n^{2k} \left(\frac{1}{k!} g^{(k)}_{t^{(k)}}(x, 0) \int_0^E u^k K(u) du \right)^2 + \frac{f(x)}{nh_n} \int_0^E K^2(u) du + o \left(h_n^{2k} + \frac{1}{nh_n} \right) \cong Ah_n^2 k + \frac{B}{nh_n} \quad (34)$$

при соответствующих A и B . Оптимальная по порядку скорость сходимости будет при

$$h_n^2 k = \frac{1}{nh_n}, \quad h_n = n^{-1/(2k+1)} \quad (35)$$

(в предположении $A \neq 0, B \neq 0$). При этом

$$\alpha_n \cong n^{-2k/(2k+1)} = n^{-1+1/(2k+1)}. \quad (36)$$

Этот результат – продвинутое обобщение теоремы 4.1 в книге И.А. Ибрагимова и Р.З. Хасьминского [14, с.316], относящейся к оцениванию плотности одномерной случайной величины. Порядок сходимости в общем случае тот же, что и в указанной теореме, но условия наложены не на плотность $f(x)$, а на круговую плотность $g(x, t)$. Это существенно в случае $Z \neq R^k$, т.к. в R^k имеются традиционно выделенные мера (Лебега) и расстояние (Евклида).

С прикладной точки зрения предположение (30) о гладкости круговой плотности представляется достаточно естественным. Напомним, что вплоть до XIX в. математики практически не делали различия между непрерывными, дифференцируемыми и аналитическими функциями. Инженеры не делают такого различия и сейчас. Отсюда методологическое предложение: в прикладных задачах допустимо использовать математические модели с той степенью гладкости рассматриваемых функций, которая позволяет наиболее легко обосновать алгоритмы расчетов, разумеется, если эта степень гладкости не противоречит фактам соответствующей предметной области. Другой пример подобного методологического подхода: шкала любого прибора конечна, поэтому результаты первичных измерений целесообразно моделировать с помощью финитных случайных величин; такие величины имеют все моменты, а это существенно облегчает получение для них предельных теорем. Методологическим вопросам посвящены наши работы [15, 16].

Из соотношения (36) следует, что для любого $\varepsilon > 0$ существует ядро $K(u)$ такое, что для соответствующей оценки

$$\alpha_n = O(n^{-1+\varepsilon}), \quad (37)$$

а также

$$\lim_{n \rightarrow \infty} n^{1/2-\varepsilon} (f_n(x) - f(x)) = 0 \quad (38)$$

по вероятности. Хотя при любом k множество ядер $K(u)$, удовлетворяющих соотношениям (32), бесконечно, не существует ядра, удовлетворяющего этим соотношениям сразу при всех k .

Перенос полученных результатов на случай конечных пространств объектов нечисловой природы осуществляется тем же путем, что и в статье [8]. Грубо говоря, необходимо, чтобы

$$\alpha_{mn}(g) = o(h_n^k), \quad (39)$$

где m – параметр дискретности (см. [8]). Поскольку принципиальных трудностей, как ясно из рассуждений статьи [8], в указанном переносе нет, мы его здесь не приводим.

Библиографический список

1. Орлов А.И. О новой парадигме прикладной математической статистики // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. / Перм. гос. нац. иссл. ун-т. – Пермь, 2013. – Вып. 25. – С. 162–176.
2. Орлов А.И. Устойчивость в социально-экономических моделях. – М.: Наука, 1979. – 296 с.
3. Орлов А.И. Статистика объектов нечисловой природы и экспертные оценки. – В сб.: Экспертные оценки / Вопросы кибернетики. Вып.58. – М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1979. – С. 17–33.
4. Орлов А.И. Организационно-экономическое моделирование: учебник: в 3-х ч. Часть 1: Нечисловая статистика. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. – 541 с.
5. Орлов А.И. О развитии статистики объектов нечисловой природы // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2013. – №93. – С. 41–50.
6. Орлов А.И. Оценки плотности в пространствах произвольной природы // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. / Перм. гос. нац. иссл. ун-т. – Пермь, 2013. – Вып. 25. – С. 21–33.
7. Орлов А.И. Ядерные оценки плотности в пространствах произвольной природы // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. / Перм. гос. нац. иссл. ун-т. – Пермь, 2015. – Вып. 26. – С. 43–57.
8. Орлов А.И. Ядерные оценки плотности в конечных пространствах // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. / Перм. гос. нац. иссл. ун-т. – Пермь, 2016. – Вып. 27. – С. 24–37.
9. Вероятность и математическая статистика: Энциклопедия / Гл. ред. Ю.В. Прохоров. – М.: Большая Российская Энциклопедия, 1999. – 910 с.
10. Орлов А.И. Статистика объектов нечисловой природы // Теория вероятностей и ее применения. – 1980. – Т. XXV. – №3. – С. 655–656.
11. Орлов А.И. Непараметрические оценки плотности в топологических пространствах // Прикладная статистика. Ученые записки по статистике, т.45. – М.: Наука, 1983. – С. 12–40.
12. Rosenblatt M. Remarks on some nonparametric estimates of a density function // Ann. Math. Statist. – 1956. – V. 27. – №5. – P. 832–837.
13. Parzen E. On estimation of a probability density function and mode // Ann. Math. Statist. – 1962. – V. 33. – №6. – P. 1065–1076.

14. *Ибрагимов И.А., Хасъминский Р.З.* Асимптотическая теория оценивания. – М.: Наука, 1979. – 528 с.
15. *Орлов А.И.* О развитии методологии статистических методов // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. / Перм. гос. нац. иссл. ун-т. – Пермь, 2001. – Вып. 15. – С. 118–131.
16. *Орлов А.И.* О влиянии методологии на последствия принятия решений // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – №125. – С. 319–345.

The rate of convergence of kernel density estimates in spaces of arbitrary nature

A.I. Orlov

Bauman Moscow State Technical University, Institute of high statistical technologies and econometrics; Moscow Physics-Technical Institute; Moscow, Russia;
prof-orlov@mail.ru; <http://orlovs.pp.ru>; +7(916)8305117; 123104, Moscow, the Sytinsky lane, house 7/14, apartment 14.

Abstract. *Nonparametric estimates of the probability distribution density in spaces of arbitrary nature are one of the main tools of non-numerical statistics. Their particular cases are considered - kernel density estimates in spaces of arbitrary nature. The purpose of this article is to study the mean square error of the kernel density estimate and, in order to maximize the order of its decrease, the choice of the kernel function and the sequence of the blur indicators. The basic concepts are the circular distribution function and the circular density. The order of convergence in the general case is the same as in estimating the density of a numerical random variable, but the main conditions are imposed not on the density of a random variable, but on the circular density.*

Keywords: *non-numeric statistics, probability density function, the space of an arbitrary nature, nuclear density estimates, limit theorems, rate of convergence.*

УДК 519.2

О некоторых свойствах порядковых статистик**В. И. Пагурова**

Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; pagurova@yandex.ru; Россия, 119992, г. Москва, ГСП-1, Воробьевы горы, МГУ им. М.В.Ломоносова

Аннотация. *Исследуются асимптотические свойства нормированной разности промежуточных порядковых статистик, когда их ранги отличаются на некоторую величину, а объём выборки неограниченно возрастает. Для центральных порядковых статистик подобная задача рассматривалась в работе [1].*

Ключевые слова: *промежуточная порядковая статистика, гамма-распределение, слабая сходимость.*

1. Введение

Пусть независимые одинаково распределенные случайные величины X_1, \dots, X_n , определенные на некотором вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$, имеют общую функцию распределения (ф.р.) $F(x) = \mathbf{P}\{X_1 \leq x\}$ и плотность распределения $f(x) = F'(x)$, $-\infty < x < \infty$. Обозначим

$$i = [n^\alpha], \quad 0 < \alpha < 1,$$

$[x]$ означает целую часть числа x . Введём промежуточные порядковые статистики $X_i^{(n)}$ и $X_{i+k}^{(n)}$ рангов i и $i+k$ соответственно в вариационном ряду, построенном по случайным величинам X_1, \dots, X_n . Ниже исследуются асимптотические свойства при $n \rightarrow \infty$ и $k = o(n^\alpha)$ нормализованной разности $X_{i+k}^{(n)} - X_i^{(n)}$.

В дальнейшем мы будем рассматривать класс B_1 распределений, для которых выполняются асимптотические представления вида

$$F(x) \sim a|x|^\gamma \exp(-b|x|^\Delta), \quad f(x) \sim ab\Delta|x|^{\gamma+\Delta-1} \exp(-b|x|^\Delta)$$

при $x \rightarrow -\infty$, $a, b, \Delta > 0$. Например, к классу B_1 относятся нормальное распределение с плотностью $f(x) = \exp\{-x^2/(2\sigma^2)\}/(\sqrt{2\pi}\sigma)$ и распределение Лапласа с плотностью $f(x) = b \exp(-b|x|)/2$, $b > 0$, $|x| < \infty$.

Введём величины

$$d_n = - \left(\frac{(1 - \alpha) \ln n}{b} \right)^{1/\Delta} \left(1 + \frac{\gamma \ln \ln n + \ln c}{\Delta^2 (1 - \alpha) \ln n} \right), \quad c = \left(\frac{1 - \alpha}{b} \right)^\gamma a^\Delta,$$

$$c_n = \left(\frac{(1 - \alpha) \ln n}{b} \right)^{(1-\Delta)/\Delta} \frac{1}{\Delta b n^{\alpha/2}}, \quad (1)$$

$$A_n = n f(d_n) = \left(\frac{(1 - \alpha) \ln n}{b} \right)^{(\Delta-1)/\Delta} \Delta b n^\alpha. \quad (2)$$

Известно, что для распределений из рассматриваемого класса B_1 промежуточная порядковая статистика $X_i^{(n)}$ при $n \rightarrow \infty$ асимптотически нормальна [2]:

$$L_n = (X_i^{(n)} - d_n)/c_n \xrightarrow{d} \eta \sim \mathcal{N}(0, 1). \quad (3)$$

2. Основной результат

Теорема 1. Пусть для некоторого целого числа $m \geq 2$ существуют непрерывные производные $F^{(1)}(x), \dots, F^{(2m-1)}(x)$, производная $F^{(2m)}(x)$ ограничена в окрестности точки d_n и $\alpha m > 1$. Тогда при $n \rightarrow \infty$ и $k = o(n^\alpha)$

$$T_n = A_n (X_{i+k}^{(n)} - X_i^{(n)}) \sim \gamma_k, \quad (4)$$

где γ_k означает случайную величину, имеющую гамма-распределение с параметром $k > 0$ и плотностью

$$\frac{e^{-x} x^{k-1}}{\Gamma(k)}, \quad x > 0,$$

$\Gamma(k)$ означает гамма-функцию Эйлера, A_n определяется соотношением (2).

Доказательство. Рассмотрим

$$I_n = \mathbf{P}\{T_n \leq x\} = \iint_B g(u, v) du dv, \quad (5)$$

$B = \{(v - u) \leq x/A_n, u \leq v\}$, $g(u, v)$ – совместная плотность величин $X_i^{(n)}, X_{i+k}^{(n)}$. Покажем, что величину I_n можно представить в виде

$$I_n = \int_{-\infty}^{\infty} g_i(u) \mathbf{P}\{\beta_{k, n-n^\alpha-k+1} \leq h_n(u, x)\} du, \quad (6)$$

где $g_i(u)$ – плотность распределения величины $X_i^{(n)}$, $\beta_{r,s}$ имеет бета-распределение с параметрами (r, s) ,

$$h_n(u, x) = (F(u + x/A_n) - F(u))/(1 - F(u)). \quad (7)$$

Известно [3], что имеют место следующие представления

$$g(u, v) = CF^{i-1}(u)(F(v) - F(u))^{k-1}f(u)(1 - F(v))^{n-i-k}f(v), \quad u \leq v,$$

$$g_i(u) = CF^{i-1}(u)(1 - F(u))^{n-i}f(u),$$

тогда

$$g(u, v) = g_i(u) C \left(\frac{F(v) - F(u)}{1 - F(u)} \right)^{k-1} \left(\frac{1 - F(v)}{1 - F(u)} \right)^{n-i-k} \frac{f(v)}{(1 - F(u))},$$

где C – общее обозначение постоянной. Учитывая полученное выражение для $g(u, v)$, в двойном интеграле (5) делаем замену переменных $\frac{F(v) - F(u)}{1 - F(u)} = t$ и интегрируем по t от 0 до $h_n(u, x)$, тогда выражение (5) принимает вид (6), что и требовалось.

Рассмотрим выражение (6). Предположим, что при $n \rightarrow \infty$

$$h_n(u, x) = \frac{x}{n} + o\left(\frac{1}{n}\right). \quad (8)$$

Применим гамма-аппроксимацию к величине $\beta_{r,s}$, когда $s \rightarrow \infty$, а $r = o(s)$, тогда $s\beta_{r,s} \sim \gamma_r$. В нашем случае $n\beta_{k,n-n^\alpha-k+1} \sim \gamma_k$, и в выражении (6) с учётом (8) имеем $\mathbf{P}\{\beta_{k,n-n^\alpha-k+1} \leq h_n(u, x)\} = \mathbf{P}\{\gamma_k \leq x + o(1)\}$. Таким образом, для завершения доказательства теоремы 1 нужно доказать соотношение (8).

Обозначим $H(u) = F(u + x/A_n) - F(u)$ и разложим $H(u)$ в ряд Тейлора в точке d_n , получим

$$H(u) = \sum_{j=0}^{2m-1} \frac{H^{(j)}(d_n)(u - d_n)^j}{j!} + \frac{H^{(2m)}(\xi)}{(2m)!}(u - d_n)^{2m}, \quad (9)$$

ξ лежит в окрестности d_n . Сначала рассмотрим последнее слагаемое в правой части (9). Из соотношения (3) следует, что при $n \rightarrow \infty$

$$(X_i^{(n)} - d_n)^{2m} \sim c_n^{2m} \eta^{2m},$$

С учетом (1), условия $\alpha t > 1$ и в силу ограниченности $F^{(2m)}(x)$ в окрестности точки d_n получим, что $H^{2m}(\xi)(u - d_n)^{2m} = o\left(\frac{1}{n}\right)$. Из условий теоремы 1 и для рассматриваемого класса B_1 распределений имеем при $n \rightarrow \infty$

$$|f^{(i)}(d_n)| \sim \left(\frac{(1 - \alpha) \ln n}{b} \right)^{(i+1)(\Delta-1)/\Delta} \frac{(b\Delta)^{i+1}}{n^{1-\alpha}}, \quad i = 0, 1, \dots, 2m - 2.$$

Тогда

$$H(d_n) = F(d_n + x/A_n) - F(d_n) = \frac{x}{n} + \sum_{j=1}^{m-2} \frac{f^{(j)}(d_n)}{(j+1)!} \left(\frac{x}{A_n}\right)^{j+1} + \frac{f^{(m-1)}(\xi)}{m!} \left(\frac{x}{A_n}\right)^m = \frac{x}{n} + o\left(\frac{1}{n}\right).$$

Аналогичным образом,

$$H^{(1)}(d_n) = \sum_{j=1}^{m-1} \frac{f^{(j)}(d_n)}{j!} \left(\frac{x}{A_n}\right)^j + \frac{f^{(m)}(\xi)}{m!} \left(\frac{x}{A_n}\right)^m,$$

поэтому при $n \rightarrow \infty$ имеем

$$H^{(1)}(d_n)(u - d_n) = o\left(\frac{1}{n}\right).$$

Точно так же оцениваются другие величины, входящие в соотношение (9):

$$H^{(2)}(d_n)(u - d_n)^2 = o\left(\frac{1}{n}\right), \dots, H^{(2m-1)}(d_n)(u - d_n)^{2m-1} = o\left(\frac{1}{n}\right).$$

Кроме того, при $n \rightarrow \infty$ в окрестности точки d_n имеем $\frac{1}{1-F(u)} = 1 + o(1)$. Таким образом, соотношение (8), где $h_n(u, x)$ определяется из соотношения (7), доказано. Теорема 1 доказана.

3. Следствия из основного результата

Рассмотрим величину

$$Q_n = A_n(X_i^{(n)} - X_{i-l}^{(n)}), \quad (10)$$

где A_n определяется из соотношения (2). Точно так же, как и для распределения величины T_n , доказывається, что при $n \rightarrow \infty$ и $l = o(n^\alpha)$ в условиях теоремы 1 величина Q_n асимптотически имеет гамма-распределение с параметром l .

Теорема 2. В условиях теоремы 1 при $n \rightarrow \infty$ и $k = o(n^\alpha)$, $l = o(n^\alpha)$ величины L_n , T_n и Q_n асимптотически независимы. Здесь L_n , T_n , Q_n определяются соотношениями (3), (4) и (10) соответственно.

Доказательство теоремы 2 аналогично доказательству теоремы 1. Обозначим $g(u, v, w)$, $u \leq v \leq w$, совместную плотность величин

$(X_{i-l}^{(n)}, X_i^{(n)}, X_{i+k}^{(n)})$, тогда [3]

$$g(u, v, w) = g_i(v) C \frac{F^{i-l-1}(u)(F(v) - F(u))^{l-1}}{F^{i-1}(v)} \times \\ \times \left(\frac{F(w) - F(v)}{1 - F(v)} \right)^{k-1} \left(\frac{1 - F(w)}{1 - F(v)} \right)^{n-i-k} \frac{f(u)f(w)}{(1 - F(v))},$$

где $g_i(v)$ – плотность распределения величины $X_i^{(n)}$. Рассмотрим

$$J_n = \mathbf{P}\{(X_i^{(n)} - d_n)/c_n \leq x, A_n(X_i^{(n)} - X_{i-l}^{(n)}) \leq y, A_n(X_{i+k}^{(n)} - X_i^{(n)}) \leq z\} = \\ = \int \int \int_M g(u, v, w) du dv dw, \quad (11)$$

$$M = \{(v - d_n)/c_n \leq x, v - y/A_n \leq u \leq v, v \leq w \leq v + z/A_n\}.$$

В интеграле (11) делаем замену $\frac{F(w)-F(v)}{1-F(v)} = t$ и интегрируем по t от 0 до $h_n(v, z) = (F(v + z/A_n) - F(v))/(1 - F(v))$. Получаем, что

$$C \int_0^{h_n(v, z)} t^{k-1} (1-t)^{n-i-k} dt = \mathbf{P}\{\beta_{k, n-n\alpha-k+1} \leq h_n(v, z)\} = J_{n,1}.$$

Аналогично тому, как это было сделано при доказательстве теоремы 1, получаем, что при $n \rightarrow \infty$ и $k = o(n^\alpha)$ имеем $J_{n,1} = \mathbf{P}\{\gamma_k \leq z + o(1)\}$. Далее в интеграле (11) делаем замену $\frac{F(u)}{F(v)} = t$ и интегрируем по t от $H_n(v, y) = F(v - y/A_n)/F(v)$ до 1. Получим, что

$$C \int_{H_n(v, y)}^1 t^{i-l-1} (1-t)^{l-1} dt = \mathbf{P}\{H_n(v, y) \leq \beta_{n\alpha-l, l} \leq 1\} = \\ = \mathbf{P}\{\beta_{l, n\alpha-l} \leq 1 - H_n(v, y)\} = J_{n,2}.$$

Имеем

$$1 - H_n(v, y) = \frac{F(v) - F(v - y/A_n)}{F(v)}.$$

Аналогично тому, как это было сделано при доказательстве теоремы 1, получим при $n \rightarrow \infty$

$$F(v - y/A_n) - F(v) = -y/n + o(1/n).$$

Кроме того, разлагая $F(v)$ в точке d_n в ряд Тейлора, получим, что

$$F(v) = \frac{1}{n^{1-\alpha}} + o\left(\frac{1}{n^{1-\alpha}}\right).$$

Таким образом, при $n \rightarrow \infty$ имеем

$$1 - H_n(v, y) = \frac{y}{n^\alpha} + o\left(\frac{1}{n^\alpha}\right).$$

Так как при $n \rightarrow \infty$ и $l = o(n^\alpha)$ имеем $(n^\alpha - l)\beta_{l, n^\alpha - l} \sim \gamma_l$, тогда $J_{n,2} = \mathbf{P}\{\gamma_l \leq y + o(1)\}$. Таким образом, при $n \rightarrow \infty$, $l = o(n^\alpha)$, $k = o(n^\alpha)$ имеем

$$J_n = \mathbf{P}\{(X_i^{(n)} - d_n)/c_n \leq x\} J_{n2} J_{n1} \sim \Phi(x) \mathbf{P}\{\gamma_l \leq y\} \mathbf{P}\{\gamma_k \leq z\},$$

$\Phi(x)$ означает функцию стандартного нормального распределения. Теорема 2 доказана.

Следствие. Так как при $n \rightarrow \infty$, $l = o(n^\alpha)$, $k = o(n^\alpha)$ величины

$$T_n = A_n(X_{i+k}^{(n)} - X_i^{(n)}) \sim \gamma_k, \quad Q_n = A_n(X_i^{(n)} - X_{i-l}^{(n)}) \sim \gamma_l,$$

T_n и Q_n асимптотически независимы, тогда асимптотически

$$T_n + Q_n = A_n(X_{i+k}^{(n)} - X_{i-l}^{(n)}) \sim \gamma_{k+l}$$

и не зависит от $X_i^{(n)}$. Если предположить, что $k = o(n^\alpha) \rightarrow \infty$, $l = o(n^\alpha) \rightarrow \infty$ при $n \rightarrow \infty$, тогда величина

$$\frac{A_n(X_{i+k}^{(n)} - X_{i-l}^{(n)}) - (k+l)}{\sqrt{k+l}}$$

асимптотически стандартно нормальна.

Рассмотрим класс распределений B_2 , для которых выполняются асимптотические представления вида

$$F(x) \sim a|x|^{-\Delta}, \quad f(x) \sim a\Delta|x|^{-\Delta-1} \text{ при } x \rightarrow -\infty, \quad a, \Delta > 0.$$

Например, к классу B_2 относится распределение Коши с плотностью

$$f(x) = \frac{b}{\pi(b^2 + x^2)}, \quad b > 0, \quad |x| < \infty.$$

Нормирующие константы $c_n > 0$ и d_n имеют вид

$$d_n = -(an^{1-\alpha})^{1/\Delta}, \quad c_n = a^{1/\Delta}/(\Delta n^{-(1-\alpha)/\Delta + \alpha/2}),$$

величина $(X_i^{(n)} - d_n)/c_n$ при $n \rightarrow \infty$ асимптотически нормальна [2], но $X_i^{(n)} - d_n \xrightarrow{P} 0$ при $n \rightarrow \infty$ только при условии $2/(2+\Delta) < \alpha < 1$, поэтому асимптотический анализ, аналогичный вышеизложенному, для распределений величин T_n и Q_n из класса B_2 применим лишь при указанном ограничении на параметр α .

Заключение

Рассмотрены условия, при которых нормализованная разность двух промежуточных порядковых статистик при увеличении объема выборки асимптотически имеет гамма-распределение. Установлены также некоторые другие асимптотические свойства функций от промежуточных порядковых статистик.

Библиографический список

1. *Siddiqui M. M.* Distribution of quantiles in samples from a bivariate population // J. Res. Nat. Bur. Stand. – 1960. – В. 64. – Р. 145–150.
2. *Пагурова В. И.* О предельном многомерном распределении промежуточных порядковых статистик // Вестник Московского государственного университета. Сер. 15. Вычислительная математика и кибернетика. – 2017. – № 3. – С. 25–28.
3. *Дейвид Г.* Порядковые статистики. – М.: Наука, 1979.

On some asymptotic properties of intermediate order statistics

V. I. Pagurova

Faculty of Computational Mathematics and Cybernetics of Moscow State University; Russia, 119992, Moscow, GSP-1, Vorobyovi gory, Lomonosov MSU; e-mail: pagurova@yandex.ru

Abstract. *We give conditions under which the normalized difference between two intermediate order statistics asymptotically has a gamma distribution when a sample size tends to infinity. We investigate also some another asymptotic properties for functions of intermediate order statistics.*

Key words: *intermediate order statistics, gamma distribution, weak convergence.*

УДК 519.2

Сравнение UMVUE и MLE с помощью степенной функции риска в случае однопараметрического экспоненциального семейства непрерывных распределений

В. В. Чичагов

Пермский государственный национальный исследовательский университет, г. Пермь, Россия
chichagov@psu.ru; +7(342) 2396214; 614990, г. Пермь, ул. Букирева, 15

Аннотация. При некоторых условиях регулярности с использованием степенной функции риска получен асимптотически ожидаемый дефект оценки максимального правдоподобия (MLE) относительно несмещенной оценки с равномерно минимальной дисперсией (UMVUE) заданной однопараметрической функции, которая оценивается по выборке из абсолютно непрерывного распределения экспоненциального типа.

Ключевые слова: асимптотически ожидаемый дефект, степенная функция риска, экспоненциальное семейство, оценка максимального правдоподобия, несмещенная оценка.

1. Введение

Пусть $\vec{X} = (X_1, \dots, X_n)$ – повторная выборка из генеральной совокупности X экспоненциального типа, имеющей плотность распределения вида

$$f[x; a] = \exp \{ \Phi_1[a]T[x] + \Phi_2[a] + d[x] \}, \quad x \in \mathbb{X}_G \subset \mathbf{R}, \quad a \in \mathbb{A}, \quad (1)$$

по отношению к некоторой мере Лебега, где $\Phi_1[\cdot]$, $\Phi_2[\cdot]$, $T[\cdot]$, $d[\cdot]$ – некоторые действительные борелевские функции, \mathbb{A} – параметрическое пространство, $a = \mathbf{E}T[X]$.

Наиболее часто для оценивания заданной параметрической функции $G[a]$, $a \in \mathbb{A}$, используются оценка максимального правдоподобия (MLE) и несмещенная оценка с равномерно минимальной дисперсией (UMVUE), которые, как это показано в [1], при определенных условиях регулярности асимптотически эквивалентны в смысле асимптотической относительной эффективности.

Поэтому в данной работе с целью выяснения, какая из этих оценок лучше, будет использоваться понятие асимптотически ожидаемого дефекта, введенное Ходжесом и Леманом [2]. Эта характеристика определяется следующим образом. Пусть $T_1[\vec{X}]$, $T_2[\vec{X}]$ – две различные оценки функции $G[a]$, а функции риска этих оценок $V_n[T_i[\vec{X}]]$, $i = 1, 2$, являются строго монотонно убывающими последовательностями, представимыми при $n \rightarrow \infty$ следующим образом:

$$V_n[T_i[\vec{X}]] = \alpha n^{-q} + b_i n^{-(q+s)} + o(n^{-(q+s)}), \quad i = 1, 2, \quad q > 0, \quad s > 0. \quad (2)$$

Тогда существует единственное значение k_n , такое, что $k_n/n \rightarrow 1$ и $V_{k_n}[T_2[\vec{X}]] = V_n[T_1[\vec{X}]]$, если $n \rightarrow \infty$. При этом асимптотически ожидаемый дефект (**AED**) (от англ. *Asymptotic Expected Deficiency*) оценки T_2 относительно оценки T_1 при заданной функции риска $V_n[\cdot]$ определяется выражением

$$\mathbf{AED}[T_2, T_1] = \lim_{n \rightarrow \infty} (k_n - n)$$

при условии, что предел существует. При этом

$$\mathbf{AED}[T_2, T_1] = \begin{cases} (b_2 - b_1)/(q\alpha), & s = 1, \\ \infty, & 0 < s < 1, \\ 0, & s > 1. \end{cases} \quad (3)$$

При ряде ограничений значение асимптотического ожидаемого дефекта **MLE** относительно **UMVUE** при квадратичной функции риска впервые найдено в [3]. В [10] и [9] предложен иной подход к решению этой задачи, основывающийся на асимптотических разложениях высокого порядка статистических оценок **MLE** и **UMVUE**. Новый подход позволил в работах [7] и [9] осуществить сравнение этих оценок, основываясь на абсолютной функции риска.

В данной работе сравнение оценок осуществляется с помощью степенной функции риска **MPE** (от англ. *Power Risk Function*), что обобщает значительную часть упомянутых выше результатов, за исключением основных результатов работы [9].

Степенную функцию риска оценки $\tilde{G}[a]$ параметрической функции $G[a]$ определим следующим образом:

$$\mathbf{PRF}[\tilde{G}[a]; r] = \mathbf{E} \left| \tilde{G}[a] - G[a] \right|^r, \quad r \geq 1. \quad (4)$$

Ниже приводятся впервые полученные асимптотические разложения степенной функции риска оценки максимального правдоподобия и несме-

щенной оценки, а также средний ожидаемый дефект оценки максимального правдоподобия относительно несмещенной оценки на основе этой функции риска.

2. Основные результаты

Чтобы сформулировать основные результаты работы, наложим на распределение (1) ряд ограничений.

(C₁). Для каждого $a \in \mathbb{A}$ существует $n_0 \in \mathbf{N}$ такое, что случайная величина

$$Z_n = \frac{S_n - na}{b\sqrt{n}}, \quad \text{где } S_n = \sum_{j=1}^n T[X_j], \quad b^2 = \mathbf{D}T[X],$$

имеет непрерывную ограниченную плотность $f_{Z_n}[\cdot]$ при $n \geq n_0$.

(C₂). \mathbb{X}_G – носитель распределения (1) не зависит от параметра a и содержит некоторый интервал, принадлежащий \mathbf{R} .

(C₃). $a\Phi'_1[a] + \Phi'_2[a] = 0$ и $\Phi'_1[a] > 0$ для всех $a \in \mathbb{A}$.

(C₄). Значение выборочного среднего S_n/n с вероятностью 1 является внутренней точкой области \mathbb{A} .

(C₅). Функция $\Phi_1[a]$ бесконечно дифференцируема по a в области \mathbb{A} .

Отметим, что в условиях (C₁) – (C₄) статистика S_n является достаточной для параметра $a \in \mathbb{A}$. При этом несмещенная оценка для заданной параметрической функции, являющаяся функцией от этой достаточной статистики, имеет равномерно минимальную дисперсию.

Несмещенную оценку и оценку максимального правдоподобия функции $G[a]$ по выборке \bar{X} будем обозначать соответственно

$$\hat{G}[a] = \hat{G}[a; Z_n], \quad \check{G}[a] = \check{G}[a; Z_n].$$

Теорема. Пусть выполнены условия (C₁) – (C₅), ε – сколь угодно малое положительное число, а последовательности степенных ошибок $\mathbf{E} \left| \hat{G}[a] - G[a] \right|^r$ и $\mathbf{E} \left| \check{G}[a] - G[a] \right|^r$, $r \geq 1$, начиная с некоторого $n \geq n_0$, равномерно ограничены некоторыми постоянными. Тогда при $n \rightarrow \infty$ и $G'[a] \neq 0$

1) справедливо следующее разложение степенной функции риска несмещенной оценки $\hat{G}[a]$ функции $G[a]$:

$$\begin{aligned} \mathbf{PRF} \left[\hat{G}[a]; r \right] &= \varphi[0] 2^{(r-1)/2} \Gamma \left[\frac{r+1}{2} \right] \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \times \\ &\times \left\{ 2 + \frac{r}{4n\Phi'[a]} \left(\frac{(r^2 - 2r + 2) G''[a]^2}{G'[a]^2} - \frac{2(r-2)(r+2)G''[a]\Phi''[a]}{3G'[a]\Phi'[a]} + \right. \right. \\ &\left. \left. + \frac{(r-2)(r+5)\Phi''[a]^2}{9\Phi'[a]^2} + \frac{4(r-2)G^{(3)}[a]}{3G'[a]} - \frac{(r-2)\Phi^{(3)}[a]}{3\Phi'[a]} \right) \right\} + \mathbf{O} \left(n^{-(r+1)/2+\varepsilon} \right); \end{aligned} \quad (5)$$

2) справедливо следующее разложение степенной функции риска оценки максимального правдоподобия $\check{G}[a]$ функции $G[a]$:

$$\begin{aligned} \mathbf{PRF}[\check{G}[a]; r] &= \varphi[0] 2^{(r-1)/2} \Gamma \left[\frac{r+1}{2} \right] \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \times \\ &\times \left\{ 2 + \frac{r}{4n\Phi'[a]} \left(\frac{(r^2 - 1)G''[a]^2}{G'[a]^2} - \frac{2r(r+1)G''[a]\Phi''[a]}{3G'[a]\Phi'[a]} + \frac{(2-r)\Phi^{(3)}[a]}{3\Phi'[a]} + \right. \right. \\ &\left. \left. + \frac{(r^2 + 3r - 10)\Phi''[a]^2}{9\Phi'[a]^2} + \frac{4(r+1)G^{(3)}[a]}{3G'[a]} \right) \right\} + \mathbf{O} \left(n^{-(r+1)/2+\varepsilon} \right). \end{aligned} \quad (6)$$

При $n \rightarrow \infty$, $G'[a] = 0$, но $G''[a] \neq 0$, верны разложения

$$\mathbf{PRF} \left[\hat{G}[a]; r \right] = 2 \left(\frac{|G''[a]|}{2n\Phi_1'[a]} \right)^r \int_0^\infty |H_2[z]|^r \varphi[z] dz + \mathbf{O} \left(n^{-r-1/2+\varepsilon} \right), \quad (7)$$

$$\mathbf{PRF}[\check{G}[a]; r] = 2 \left(\frac{|G''[a]|}{2n\Phi_1'[a]} \right)^r \int_0^\infty z^{2r} \varphi[z] dz + \mathbf{O} \left(n^{-r-1/2+\varepsilon} \right). \quad (8)$$

Доказательство теоремы приведено в разделе 3.

Следствие 1. Пусть $G'[a] \neq 0$ и выполнены условия $(\mathbf{C}_1) - (\mathbf{C}_5)$. Тогда \mathbf{AED}_{PRF} – асимптотически ожидаемый дефект оценки максимального правдоподобия относительно несмещенной оценки при функции риска (4) имеет вид:

$$\mathbf{AED}_{PRF} = \frac{r}{\Phi'[a]} \left\{ \frac{(2r-3)G''[a]^2}{4G'[a]^2} - \frac{(r+4)G''[a]\Phi''[a]}{6G'[a]\Phi'[a]} + \frac{G^{(3)}[a]}{G'[a]} \right\}. \quad (9)$$

Доказательство. Так как \mathbf{PRF} обеих оценок представима в виде

$$\mathbf{PRF}[\check{G}[a]; r] = \alpha n^{-r/2} + b_2 n^{-r/2-1} + \mathbf{O} \left(n^{-(r+1)/2+\varepsilon} \right),$$

$$\mathbf{PRF} \left[\hat{G}[a]; r \right] = \alpha n^{-r/2} + b_1 n^{-r/2-1} + \mathbf{O} \left(n^{-(r+1)/2+\varepsilon} \right),$$

то \mathbf{AED}_{PRF} вычисляется согласно (3) по формуле:

$$\mathbf{AED}_{PRF} \equiv \mathbf{AED} \left[\check{G}[a], \hat{G}[a] \right] = 2(b_2 - b_1)/(r\alpha), \quad (10)$$

где

$$\begin{aligned} \alpha &= \varphi[0] 2^{(r+1)/2} \Gamma \left[\frac{r+1}{2} \right] \left(\frac{|G'[a]|}{\sqrt{\Phi'_1[a]}} \right)^r, \\ b_2 - b_1 &= \frac{r\varphi[0]}{\Phi'_1[a]} 2^{(r-1)/2} \Gamma \left[\frac{r+1}{2} \right] \left(\frac{|G'[a]|}{\sqrt{n\Phi'_1[a]}} \right)^r \times \\ &\times \left\{ \frac{(2r-3)G''[a]^2}{4G'[a]^2} - \frac{(r+4)G''[a]\Phi''[a]}{6G'[a]\Phi'[a]} + \frac{G^{(3)}[a]}{G'[a]} \right\}. \end{aligned}$$

Подставляя выражения для α , $b_2 - b_1$ в (10), получим (9).
Следствие доказано.

3. Доказательство теоремы

Предварительно отметим, что все используемые в ходе доказательства теоремы асимптотические разложения для функций с аргументом z будут рассматриваться при $|z| \leq \sqrt{2\alpha \ln n}$, $n \rightarrow \infty$ и $\alpha \geq 3$. Кроме того, при проведении преобразований часто будет использоваться оценка вида

$$|z|^\beta n^{-1/2} = \mathbf{O} \left(n^{-1/2+\delta} \right),$$

которая справедлива при $|z| \leq \sqrt{2\alpha \ln n}$, $n \rightarrow \infty$, любых $\alpha > 0$, $\beta > 0$ и сколь угодно малом $\delta > 0$. В этой связи при записи асимптотических разложений будут использоваться постоянные δ_i , $i = 1, 6$, которые могут быть выбраны сколь угодно малыми, меньшими ε .

Приступим непосредственно к доказательству теоремы.

Поскольку в модели (1) существуют моменты любого порядка q случайной величины $T[X]$, а по теореме Амосовой [5, с. 309] при $q > 2\alpha + 2$ и $\alpha > 0$ верно разложение

$$\mathbf{P} \left(|Z_n| \geq \sqrt{2\alpha \ln n} \right) = \frac{1}{n^\alpha \sqrt{\pi\alpha \ln n}} \left(1 + \mathbf{O} \left(\frac{1}{\ln n} \right) \right),$$

то для любого $\alpha > 0$ справедлива оценка

$$\mathbf{P} \left(|Z_n| > \sqrt{2\alpha \ln n} \right) = \mathbf{o} \left(n^{-\alpha} \right). \quad (11)$$

Отсюда в силу неравенства Коши-Буняковского справедливы неравенства

$$\begin{aligned} \mathbf{E} \left\{ \left| \hat{G}[a; Z_n] - G[a] \right|^r I \left(|Z_n| > \sqrt{2\alpha \ln n} \right) \right\} &\leq \\ &\leq \sqrt{\mathbf{E} \left| \hat{G}[a; Z_n] - G[a] \right|^{2r} \mathbf{P} \left(|Z_n| > \sqrt{2\alpha \ln n} \right)}, \end{aligned} \quad (12)$$

$$\begin{aligned} \mathbf{E} \left\{ \left| \check{G}[a; Z_n] - G[a] \right|^r I \left(|Z_n| > \sqrt{2\alpha \ln n} \right) \right\} &\leq \\ &\leq \sqrt{\mathbf{E} \left| \check{G}[a; Z_n] - G[a] \right|^{2r} \mathbf{P} \left(|Z_n| > \sqrt{2\alpha \ln n} \right)}, \end{aligned} \quad (13)$$

Используя (11)–(13), в условиях теоремы имеем

$$\mathbf{E} \left\{ \left| \hat{G}[a; Z_n] - G[a] \right|^r I \left(|Z_n| > \sqrt{2\alpha \ln n} \right) \right\} = \mathbf{o} \left(n^{-\alpha/2} \right), \quad (14)$$

$$\mathbf{E} \left\{ \left| \check{G}[a; Z_n] - G[a] \right|^r I \left(|Z_n| > \sqrt{2\alpha \ln n} \right) \right\} = \mathbf{o} \left(n^{-\alpha/2} \right). \quad (15)$$

Далее нам потребуются уточнение локальной предельной теоремы для нормированной суммы Z_n независимых одинаково распределенных непрерывных случайных величин

$$f_{Z_n}[z] = \varphi[z] \left(1 + \frac{\rho_3 H_3[z]}{6\sqrt{n}} + \frac{\rho_4 H_4[z]}{24n} + \frac{\rho_3^2 H_6[z]}{72n} \right) + \mathbf{O} \left(n^{-3/2} \right) \quad (16)$$

из [5] и асимптотические разложения функций, определяющих разложения погрешности несмещенной оценки и оценки максимального правдоподобия, из [7], [9], [10]:

$$\hat{G}[a; z] - G[a] = \frac{G'[a]}{\sqrt{n\Phi_1'[a]}} \left\{ z + \frac{AH_2[z]}{\sqrt{n}} + \frac{B_1 H_3[z] + B_2 z}{n} + \mathbf{O} \left(n^{-3/2+\delta_1} \right) \right\}, \quad (17)$$

$$\check{G}[a; z] - G[a] = \frac{G'[a]}{\sqrt{n\Phi_1'[a]}} \left\{ z + \frac{Az^2}{\sqrt{n}} + \frac{B_1 z^3}{n} + \mathbf{O} \left(n^{-3/2+\delta_2} \right) \right\}, \quad (18)$$

где

$$b^2 = \frac{1}{\Phi_1'[a]}, \quad \rho_3 = -\frac{\Phi^{(2)}[a]}{\Phi^{(1)}[a]^{3/2}}, \quad \rho_4 = \frac{3\Phi^{(2)}[a]^2}{\Phi^{(1)}[a]^3} - \frac{\Phi^{(3)}[a]}{\Phi^{(1)}[a]^2}, \quad (19)$$

$$A = \frac{G^{(2)}[a]}{2G^{(1)}[a]\Phi^{(1)}[a]^{1/2}}, \quad B_1 = \frac{G^{(3)}[a]}{6G^{(1)}[a]\Phi^{(1)}[a]}, \quad B_2 = \frac{G^{(2)}[a]\Phi^{(2)}[a]}{2G^{(1)}[a]\Phi^{(1)}[a]^2}. \quad (20)$$

В лемме 4 из [6] было показано, что уравнение

$$z + \frac{AH_2[z]}{\sqrt{n}} + \frac{B_1 H_3[z] + B_2 z}{n} = 0 \quad (21)$$

имеет один корень $z_0 = \frac{A}{\sqrt{n}} + \mathbf{O}(n^{-3/2})$ в области $|z| \leq \sqrt{2\alpha \ln n}$ при достаточно больших значениях n . Поэтому левую часть (21) можно представить следующим образом:

$$\begin{aligned} z + \frac{AH_2[z]}{\sqrt{n}} + \frac{B_1H_3[z] + B_2z}{n} &= \\ &= (z - z_0) \left(1 + \frac{Az}{\sqrt{n}} + \frac{B_2 - 3B_1 + A^2 + B_1z^2}{n} \right) + \mathbf{O}\left(\frac{1}{n^{3/2-\delta_3}}\right). \end{aligned} \quad (22)$$

Используя разложение бинома Ньютона и соотношения

$$z = H_1[z], z^2 = H_2[z] + 1, z^3 = H_3[z] + 3H_1[z], \quad (23)$$

$$z^4 = H_4[z] + 6H_2[z] + 3, z^6 = H_6[z] + 15H_4[z] + 45H_2[z] + 15, \quad (24)$$

получим при $\delta_4 > \delta_3$ разложения следующих выражений:

$$\begin{aligned} \left(1 + \frac{Az}{\sqrt{n}} + \frac{B_2 - 3B_1 + A^2 + B_1z^2}{n} + \mathbf{O}\left(\frac{1}{n^{3/2-\delta}}\right) \right)^r &= \\ = 1 + \frac{Azr}{n^{1/2}} + \frac{r(B_2 - 3B_1 + A^2 + B_1z^2)}{n} + \frac{r(r-1)A^2z^2}{2n} + \mathbf{O}\left(\frac{1}{n^{3/2-\delta_4}}\right), \end{aligned} \quad (25)$$

$$\begin{aligned} \psi[z] &= \left(1 + \frac{Azr}{n^{1/2}} + \frac{r(B_2 - 3B_1 + A^2 + B_1z^2)}{n} + \frac{r(r-1)A^2z^2}{2n} \right) \times \\ &\times \left(1 + \frac{\rho_3H_3[z]}{6\sqrt{n}} + \frac{\rho_4H_4[z]}{24n} + \frac{\rho_3^2H_6[z]}{72n} \right) = \\ &= 1 + \frac{1}{n^{1/2}} \left(Arz - \frac{z\rho_3}{2} + \frac{z^3\rho_3}{6} \right) + \frac{1}{72n} (72A^2r - 216B_1r + 72B_2r - \\ &- 36A^2rz^2 + 72B_1rz^2 + 36A^2r^2z^2 - 36Arz^2\rho_3 + 12Arz^4\rho_3 - 15\rho_3^2 + \\ &+ 45z^2\rho_3^2 - 15z^4\rho_3^2 + z^6\rho_3^2 + 9\rho_4 - 18z^2\rho_4 + 3z^4\rho_4) + \mathbf{O}\left(\frac{1}{n^{3/2-\delta_5}}\right) = \\ &= 1 + \frac{1}{n^{1/2}} \left(ArH_1[z] + \frac{\rho_3H_3[z]}{6} \right) + \frac{A^2r(r+1+H_2[z](r-1))}{2n} + \frac{\rho_4H_4[z]}{24n} + \\ &+ \frac{\rho_3^2H_6[z]}{72n} + \frac{rA\rho_3(3H_2[z] + H_4[z])}{6n} + \frac{r(B_2 + B_1(H_2[z] - 2))}{n} + \mathbf{O}\left(\frac{1}{n^{3/2-\delta_5}}\right). \end{aligned} \quad (26)$$

Используя (14) при $\alpha = 3$, (23)–(26), представим функционал $\mathbf{E} \left| \hat{G}[a; Z_n] - G[a] \right|^r$ при $\delta_6 > \delta_5$ следующим образом:

$$\begin{aligned}
\mathbf{E} \left| \hat{G}[a; Z_n] - G[a] \right|^r &= \mathbf{E} \left\{ \left(\hat{G}[a; Z_n] - G[a] \right)^r I \left(|Z_n| \leq \sqrt{2\alpha \ln n} \right) \right\} + \quad (27) \\
&+ \mathbf{o} \left(n^{-3/2} \right) = \int_{|z| \leq \sqrt{2\alpha \ln n}} \left| \hat{G}[a; z] - G[a] \right|^r f_{Z_n}[z] dz + \mathbf{O} \left(n^{-3/2+\delta_6} \right) = \\
&= \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \int_{|z| \leq \sqrt{2\alpha \ln n}} |z - z_0|^r \psi[z] \varphi[z] dz + \mathbf{O} \left(n^{-3/2+\delta_6} \right) = \\
&= \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \int_{-\infty}^{\infty} |z - z_0|^r \psi[z] \varphi[z] dz + \mathbf{O} \left(n^{-3/2+\delta_6} \right) = \\
&= \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \left(\int_{-\infty}^{z_0} (z_0 - z)^r \psi[z] \varphi[z] dz + \int_{z_0}^{\infty} (z - z_0)^r \psi[z] \varphi[z] dz \right) + \\
&+ \mathbf{o} \left(n^{-3/2+\delta} \right) = \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \times \\
&\times \int_0^{\infty} t^r (\psi[z_0 - t] \varphi[z_0 - t] dt + \psi[z_0 + t] \varphi[z_0 + t]) dt + \mathbf{O} \left(n^{-3/2+\delta_6} \right).
\end{aligned}$$

В ходе преобразований в (27) использована оценка интеграла

$$\int_{|z| \geq \sqrt{2\alpha \ln n}} |z|^\beta \varphi[z] dz = \mathbf{O} \left(n^{-\alpha+\delta_7} \right),$$

которая справедлива для любого $\beta > 0$.

Имея в виду, что функция $\psi[z]$ представляет собой линейную комбинацию полиномов Чебышева–Эрмита $\{H_{2k}[x], k = 0, 1, 2, 3\}$ и $\{H_{2k+1}[x], k = 0, 1\}$, выполним следующие разложения в точке t :

$$\begin{aligned}
H_{2k}[t + z_0] \varphi[t + z_0] + H_{2k}[z_0 - t] \varphi[z_0 - t] &= H_{2k}[t + z_0] \varphi[t + z_0] + \quad (28) \\
+ H_{2k}[t - z_0] \varphi[t - z_0] &= H_{2k}[t] \varphi[t] + H_{2k}[-t] \varphi[-t] - \\
- (H_{2k+1}[t] \varphi[t] + H_{2k+1}[-t] \varphi[-t]) z_0 &+ (H_{2k+2}[t] \varphi[t] + H_{2k+2}[-t] \varphi[-t]) \frac{z_0^2}{2} - \\
- (H_{2k+3}[\zeta_1] \varphi[\zeta_1] + H_{2k+3}[\zeta_2] \varphi[\zeta_2]) \frac{z_0^3}{6} &= \\
= 2H_{2k}[t] \varphi[t] + H_{2k+2}[t] \varphi[t] z_0^2 - \frac{z_0^3}{6} &(H_{2k+3}[\zeta_1] \varphi[\zeta_1] + H_{2k+3}[\zeta_2] \varphi[\zeta_2]),
\end{aligned}$$

$$\begin{aligned}
& H_{2k+1}[t+z_0]\varphi[t+z_0] + H_{2k+1}[z_0-t]\varphi[z_0-t] = H_{2k+1}[t+z_0]\varphi[t+z_0] + \\
& + H_{2k+1}[t-z_0]\varphi[t-z_0] = H_{2k+1}[t]\varphi[t] + H_{2k+1}[-t]\varphi[-t] - \\
& - (H_{2k+2}[t]\varphi[t] + H_{2k+2}[-t]\varphi[-t])z_0 + (H_{2k+3}[t]\varphi[t] + H_{2k+3}[-t]\varphi[-t])\frac{z_0^2}{2} - \\
& - (H_{2k+4}[\zeta_3]\varphi[\zeta_3] + H_{2k+4}[\zeta_4]\varphi[\zeta_4])\frac{z_0^3}{6} = \\
& = -2z_0H_{2k}[t]\varphi[t] - \frac{z_0^3}{6}(H_{2k+4}[\zeta_3]\varphi[\zeta_3] + H_{2k+4}[\zeta_4]\varphi[\zeta_4]),
\end{aligned} \tag{29}$$

где ζ_1, ζ_3 – некоторые значения между t и $t+z_0$, а ζ_2, ζ_4 – некоторые значения между $-t$ и $-t+z_0$. При этом использовано соотношение

$$(H_k[x]\varphi[x])' = -H_{k+1}[x]\varphi[x].$$

Применяя (28)–(29) и (23)–(24), представим (27) следующим образом:

$$\mathbf{E} \left| \hat{G}[a; Z_n] - G[a] \right|^r = \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \int_0^\infty t^r \psi_1[t]\varphi[t]dt + o\left(n^{-3/2+\delta_8}\right),$$

где

$$\begin{aligned}
\psi_1[z] &= 2 + \frac{A^2H_2[z] - 2rA^2H_2[z]}{n} + \frac{2}{n} \left(r(B_2 - 2B_1) + \frac{1}{2}A^2r(r+1) \right) - \\
& - \frac{\rho_3}{3n}AH_4[z] + \frac{\rho_3^2}{36n}H_6[z] + \frac{2H_2[z]}{n} \left(rB_1 + \frac{A^2(r-1)r}{2} + \frac{Ar\rho_3}{2} \right) + \\
& + \frac{2H_4[z]}{n} \left(\frac{Ar\rho_3}{6} + \frac{\rho_4}{24} \right) = 2 - \frac{A^2}{n} + \frac{4rA^2}{n} - \frac{6rB_1}{n} + \frac{2rB_2}{n} - \frac{A\rho_3}{n} - \frac{5\rho_3^2}{12n} + \\
& + \frac{\rho_4}{4n} + z^2 \left(\frac{A^2}{n} - \frac{3rA^2}{n} + \frac{2rB_1}{n} + \frac{A^2r^2}{n} - \frac{(r-2)A\rho_3}{n} + \frac{5\rho_3^2}{4n} - \frac{\rho_4}{2n} \right) + \\
& + z^4 \left(\frac{(r-1)A\rho_3}{3n} - \frac{5\rho_3^2}{12n} + \frac{\rho_4}{12n} \right) + \frac{z^6\rho_3^2}{36n}.
\end{aligned} \tag{30}$$

Наконец, учитывая, что $B_2 = -A\rho_3$, а также соотношение

$$\int_0^\infty t^\beta \varphi[t]dt = 2^{(\beta-1)/2} \varphi[0] \Gamma \left[\frac{\beta+1}{2} \right], \text{ если } \beta > 0, \tag{31}$$

завершим доказательство (5):

$$\begin{aligned}
\mathbf{E} \left| \hat{G}[a; Z_n] - G[a] \right|^r &= \left(\frac{|G'[a]|}{\sqrt{n\Phi'_1[a]}} \right)^r \frac{2^{(r-5)/2}}{9} \varphi[0] \Gamma \left[\frac{r+1}{2} \right] \times \\
&\times \left\{ 72 + \frac{1}{n} (12Ar (r^2 - 4) \rho_3 + r (8 - 6r + r^2) \rho_3^2 + \right. \\
&+ 72B_1(r-2)r + 36A^2r(r^2 - 2r + 2) + 3(r-2)r\rho_4) \left. \right\} = \\
&= \left(\frac{|G'[a]|}{\sqrt{n\Phi'_1[a]}} \right)^r 2^{(r-5)/2} \varphi[0] \Gamma \left[\frac{r+1}{2} \right] \times \\
&\times \left(8 + \frac{1}{n\Phi'[a]} \left(\frac{2rG''[a]^2}{G'[a]^2} - \frac{2r^2G''[a]^2}{G'[a]^2} + \frac{r^3G''[a]^2}{G'[a]^2} + \right. \right. \\
&+ \frac{8rG''[a]\Phi''[a]}{3G'[a]\Phi'[a]} - \frac{2r^3G''[a]\Phi''[a]}{3G'[a]\Phi'[a]} - \frac{10r\Phi''[a]^2}{9\Phi'[a]^2} + \frac{r^2\Phi''[a]^2}{3\Phi'[a]^2} + \frac{r^3\Phi''[a]^2}{9\Phi'[a]^2} - \\
&\left. \left. - \frac{8rG^{(3)}[a]}{3G'[a]} + \frac{4r^2G^{(3)}[a]}{3G'[a]} + \frac{2r\Phi^{(3)}[a]}{3\Phi'[a]} - \frac{r^2\Phi^{(3)}[a]}{3\Phi'[a]} \right) \right) = \\
&= \left(\frac{|G'[a]|}{\sqrt{n\Phi'_1[a]}} \right)^r 2^{(r-5)/2} \varphi[0] \Gamma \left[\frac{r+1}{2} \right] \times \\
&\times \left(8 + \frac{r}{n\Phi'[a]} \left(\frac{(r^2 - 2r + 2) G''[a]^2}{G'[a]^2} - \frac{2(r-2)(r+2)G''[a]\Phi''[a]}{3G'[a]\Phi'[a]} + \right. \right. \\
&\left. \left. + \frac{(r-2)(r+5)\Phi''[a]^2}{9\Phi'[a]^2} + \frac{4(r-2)G^{(3)}[a]}{3G'[a]} - \frac{(r-2)\Phi^{(3)}[a]}{3\Phi'[a]} \right) \right).
\end{aligned}$$

Перейдем к доказательству (6).

В лемме 3 из [6] было показано, что уравнение

$$z + \frac{Az^2}{\sqrt{n}} + \frac{B_1z^3}{n} = 0$$

при достаточно больших значениях n и значениях $|z| \leq \sqrt{2\alpha \ln n}$ имеет только один корень $z = 0$. Поэтому в этих условиях верно соотношение

$$\begin{aligned}
\left| z + \frac{Az^2}{\sqrt{n}} + \frac{B_1z^3}{n} + \mathbf{O} \left(n^{-3/2+\delta_2} \right) \right|^r &= \\
&= |z|^r \left(1 + \frac{Az}{\sqrt{n}} + \frac{B_1z^2}{n} \right)^r + \mathbf{O} \left(n^{-3/2+\delta_9} \right).
\end{aligned} \tag{32}$$

При этом

$$\left(1 + \frac{Az}{\sqrt{n}} + \frac{B_1z^2}{n} \right)^r = 1 + \frac{Azr}{n^{1/2}} + \frac{r}{n} \left(B_1 + \frac{(r-1)A^2}{2} \right) z^2 + \mathbf{O} \left(n^{-3/2+\delta_{10}} \right). \tag{33}$$

С учетом (30), определим функцию

$$\begin{aligned} \psi_2[z] &= \left\{ 1 + \frac{Azr}{n^{1/2}} + \frac{r}{n} \left(B_1 + \frac{(r-1)A^2}{2} \right) z^2 \right\} \times \\ &\times \left(1 + \frac{\rho_3 H_3[z]}{6\sqrt{n}} + \frac{\rho_4 H_4[z]}{24n} + \frac{\rho_3^2 H_6[z]}{72n} \right) = 1 + \frac{1}{n^{1/2}} \left(Arz - \frac{z\rho_3}{2} + \frac{z^3\rho_3}{6} \right) + \\ &+ \frac{1}{72n} \left\{ z^2 (72B_1r - 36A^2r + 36A^2r^2 - 36Ar\rho_3 + 45\rho_3^2 + 9\rho_4 - 18\rho_4) + \right. \\ &\left. + z^6\rho_3^2 + z^4 (12Ar\rho_3 - 15\rho_3^2 + 3\rho_4) - 15\rho_3^2 + 9\rho_4 \right\} + \mathbf{O} \left(n^{-3/2+\delta_{11}} \right). \end{aligned} \quad (34)$$

Используя (15) при $\alpha = 3$ и (32)–(34), получим асимптотическое разложение

$$\begin{aligned} \mathbf{E} |\check{G}[a; Z_n] - G[a]|^r &= \mathbf{E} \left\{ (\check{G}[a; Z_n] - G[a])^r I \left(|Z_n| \leq \sqrt{2\alpha \ln n} \right) \right\} + \\ &+ \mathbf{o} \left(n^{-3/2} \right) = \int_{|z| \leq \sqrt{2\alpha \ln n}} |\check{G}[a; z] - G[a]|^r f_{Z_n}[z] dz + \mathbf{o} \left(n^{-3/2+\delta_{12}} \right) = \\ &= \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \int_{|z| \leq \sqrt{2\alpha \ln n}} |z|^r \psi_2[z] \varphi[z] dz + \mathbf{o} \left(n^{-3/2+\delta_{13}} \right) = \\ &= \left(\frac{|G'[a]|}{\sqrt{n\Phi_1'[a]}} \right)^r \int_{-\infty}^{\infty} |z|^r \psi_2[z] \varphi[z] dz + \mathbf{o} \left(n^{-3/2+\delta_{13}} \right), \end{aligned} \quad (35)$$

где

$$\begin{aligned} \int_{-\infty}^{\infty} |z|^r \psi_2[z] \varphi[z] dz &= 2 \int_0^{\infty} z^r \varphi[z] dz + \frac{9\rho_4 - 15\rho_3^2}{36n} \int_0^{\infty} z^r \varphi[z] dz + \\ &+ \frac{72rB_1 + 36(r-1)rA^2 - 36rA\rho_3 + 45\rho_3^2 - 18\rho_4}{36n} \int_0^{\infty} z^{r+2} \varphi[z] dz + \\ &+ \frac{12Ar\rho_3 - 15\rho_3^2 + 3\rho_4}{36n} \int_0^{\infty} z^{r+4} \varphi[z] dz + \frac{\rho_3^2 z^6}{36n} \int_0^{\infty} z^{r+6} \varphi[z] dz = \\ &= \varphi[0] 2^{(r-1)/2} \Gamma \left[\frac{r+1}{2} \right] \left\{ 2 + \frac{r(1+r)}{n} (2B_1 + (r-1)A^2) + \right. \\ &\left. + \frac{r}{36n} (12Ar(1+r)\rho_3 + (8 + (r-6)r)\rho_3^2 + 3(r-2)\rho_4) \right\}. \end{aligned}$$

Подставляя (19)–(20) в (35) и используя (31), получим (6).

Если $G'[a] = 0$, но $G''[a] \neq 0$, то из (17)–(18) при $|z| \leq \sqrt{2\alpha \ln n}$ и $n \rightarrow \infty$ следует справедливость разложений

$$\hat{G}[a; z] - G[a] = \frac{G''[a] H_2[z]}{2n\Phi'_1[a]} + \mathbf{O}\left(n^{-3/2+\delta_{14}}\right), \quad (36)$$

$$\check{G}[a; z] - G[a] = \frac{G''[a] z^2}{2n\Phi'_1[a]} + \mathbf{O}\left(n^{-3/2+\delta_{15}}\right). \quad (37)$$

Опираясь на разложения (36)–(37), нетрудно убедиться в справедливости заключительных соотношений теоремы (7) и (8), проведя рассуждения по той же схеме, что и при доказательстве (5)–(6).

Следствие 2. В условиях теоремы при $r = 1$ справедливы следующие разложения средних абсолютных погрешностей несмещенной оценки и оценки максимального правдоподобия:

$$\mathbf{E} \left| \hat{G}[a; Z_n] - G[a] \right| = \frac{2\varphi[0] |G'[a]|}{\sqrt{n\Phi'_1[a]}} \left\{ 1 + \frac{1}{n\Phi'[a]} \times \right. \quad (38)$$

$$\left. \times \left(\frac{G''[a]^2}{8G'[a]^2} + \frac{G''[a]\Phi''[a]}{4G'[a]\Phi'[a]} - \frac{\Phi''[a]^2}{12\Phi'[a]^2} - \frac{G^{(3)}[a]}{6G'[a]} + \frac{\Phi^{(3)}[a]}{24\Phi'[a]} \right) + \mathbf{O}\left(n^{-3/2+\varepsilon}\right) \right\},$$

$$\mathbf{E} \left| \check{G}[a; Z_n] - G[a] \right| = \frac{2\varphi[0] |G'[a]|}{\sqrt{n\Phi'_1[a]}} \left\{ 1 + \frac{1}{n\Phi'[a]} \times \quad (39)$$

$$\times \left(\frac{G^{(3)}[a]}{3G'[a]} - \frac{G''[a]\Phi''[a]}{6G'[a]\Phi'[a]} - \frac{\Phi''[a]^2}{12\Phi'[a]^2} + \frac{\Phi^{(3)}[a]}{24\Phi'[a]} \right) + \mathbf{O}\left(n^{-3/2+\varepsilon}\right) \right\}.$$

Впервые разложения (38)–(39) были получены в [6].

Следствие 3. В условиях теоремы при $r = 2$ справедливы следующие разложения средних абсолютных погрешностей несмещенной оценки и оценки максимального правдоподобия:

$$\mathbf{E} \left| \hat{G}[a; Z_n] - G[a] \right|^2 = \frac{G'[a]^2}{n\Phi'_1[a]} \left(1 + \frac{G''[a]^2}{2n\Phi'[a]G'[a]^2} + \mathbf{O}\left(n^{-3/2+\varepsilon}\right) \right), \quad (40)$$

$$\mathbf{E} \left| \check{G}[a; Z_n] - G[a] \right|^2 = \quad (41)$$

$$= \frac{G'[a]^2}{n\Phi'_1[a]} \left\{ 1 + \frac{3G''[a]^2}{4nG'[a]^2\Phi'[a]} - \frac{G''[a]\Phi''[a]}{nG'[a]\Phi'[a]^2} + \frac{G^{(3)}[a]}{nG'[a]\Phi'[a]} + \mathbf{O}\left(n^{-3/2+\varepsilon}\right) \right\}.$$

Впервые разложения (40)–(41) были получены в [8], [10].

Библиографический список

1. Portnoy S. Asymptotic efficiency of minimum variance unbiased estimators// Annals of Statistics. – 1977. – Vol. 5, №3. – P. 522–529.

2. *Hodges J.L., Lehmann E.L.* Deficiency// *Ann. Math. Statist.* – 1970. – Vol. 41. – P. 783–801.
3. *Hwang T.-Y., Hu C.-Y.* More comparisons of MLE with UMVUE for exponential families// *Ann. Inst. Statist. Math.* – 1990. Vol. 42. – P. 65–75.
4. *Кудряцев Л.Д.* Курс математического анализа. Т. II. – М.: Высшая школа, 1981.
5. *Петров В.В.* Суммы независимых случайных величин. – М.: Наука, 1972.
6. *Федосеева Н.П., Чичагов В.В.* Сравнение UMVUE и MLE с помощью абсолютной функции потерь в случае однопараметрического экспоненциального семейства непрерывных распределений// *Статистические методы оценивания и проверки гипотез: Межвуз. сб. науч. тр./ ПГНИУ.* – Пермь, 2012. – С. 96–109.
7. *Чичагов В.В.* Стохастические разложения несмещенных оценок в случае однопараметрического экспоненциального семейства// *Информатика и ее применения.* – 2008. – Т. 2, вып. 2. – С. 62–70.
8. *Чичагов В.В.* Асимптотические разложения высокого порядка для несмещенных оценок и их дисперсий в модели однопараметрического экспоненциального семейства// *Информ. и ее примен.* – 2015. – Т.9, вып. 3. – С. 75–87.
9. *Чичагов В. В.* Асимптотические разложения средней абсолютной ошибки несмещенной оценки с равномерно минимальной дисперсией и оценки максимального правдоподобия в модели однопараметрического экспоненциального семейства решетчатых распределений// *Информ. и ее примен.* – 2016. – Т.10, вып. 3. – С. 66–76.
10. *Chichagov V.* On comparison of UMVUE and MLE risks on high order asymptotic expansions for one-parameter exponential family. XXIX International Seminar on Stability Problems for Stochastic Models (ISSPSM'2011) and V International Workshop "Applied Problems in Theory of Probabilities and Mathematical Statistics related to modeling of information systems"(АТП + MS'2011).: Book of abstracts. – М.: IPI RAS, 2011. – P. 12–14.

Comparison UMVUE and MLE with power risk function in the case of a one-parameter exponential family continuous distributions

V.V. Chichagov

Perm State University, Perm, Russia

chichagov@psu.ru; +7(342) 2396214; 15, Bukireva str., Perm, Russia, 614990

Abstract. The asymptotic expected deficiency of the maximum likelihood estimation relative to the uniformly minimum variance unbiased estimator is obtained under certain regularity conditions with an power risk function for a given one-parameter function, as estimated by a sample from an absolutely continuous distribution of exponential type.

Key words: asymptotic expected deficiency, exponential family, maximum likelihood estimation, uniformly minimum variance unbiased estimator.

УДК 519.2

Неравенства типа Каца–Петрова для некоторых случайных сумм¹

А.В. Дорофеева

Факультет вычислительной математики и кибернетики МГУ имени М.В. Ломоносова; 119991, Москва, ГСП-1, Ленинские горы, МГУ, 2-й уч. корп., ф-т ВМК; alex.dorofeyeva@gmail.com

Аннотация. Доказаны аналоги неравенств Каца–Петрова, устанавливающие неравномерные оценки скорости сходимости в центральной предельной теореме для сумм случайного числа независимых одинаково распределенных случайных величин для случаев, когда индекс суммирования (число слагаемых в сумме) имеет биномиальное или пуассоновское распределение и стохастически независим от слагаемых. Рассмотрена ситуация, когда на слагаемые накладываются ослабленные моментные условия, в частности, слагаемые могут не иметь моменты степенных порядков, больших второго. Указаны конкретные числовые значения абсолютных констант, входящих в оценки. На биномиальные и пуассоновские суммы перенесен результат В. В. Петрова, уточняющий упомянутые неравенства Каца–Петрова, из которого, в частности, вытекает, что если слагаемые имеют момент порядка $2+\delta$ с $\delta \in (0, 1)$, то остаточный член в соответствующей версии центральной предельной теоремы имеет более высокий порядок малости (по числу слагаемых и аргументу), нежели степенные порядки, устанавливаемые неравенствами типа Нагаева–Бикялиса.

Ключевые слова: центральная предельная теорема; нормальная аппроксимация; случайная сумма; биномиальное распределение; распределение Пуассона; скорость сходимости.

Введение

Оценки точности нормальной аппроксимации для распределений сумм случайных величин традиционно являются объектом пристального внимания специалистов в области теории вероятностей, поскольку они играют важную роль во многих прикладных задачах. Такие оценки помогают осознанно принимать решения об адекватности или неадекватности нормальной модели для наблюдаемых статистических закономерностей. При этом особый интерес представляет ситуация, в которой распределения слагаемых могут иметь довольно тяжелые хвосты. Эта ситуация в статье

¹Работа выполнена при поддержке РФФИ (проект 18-07-01405)

формализована в виде ослабленных моментных условий, налагаемых на слагаемые. В частности, слагаемые могут не иметь моменты степенных порядков, больших второго. Пусть X_1, X_2, \dots – независимые случайные величины с $\mathbf{E}X_i = 0$ и $0 < \mathbf{E}X_i^2 \equiv \sigma_i^2 < \infty$, $i = 1, 2, \dots$. Для $n \in \mathbb{N}$ обозначим $S_n = X_1 + \dots + X_n$, $B_n^2 = \sigma_1^2 + \dots + \sigma_n^2$. Стандартную нормальную функцию распределения обозначим $\Phi(x)$:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz, \quad x \in \mathbb{R}.$$

Обозначим

$$\Delta_n(x) = |\mathbf{P}(S_n < xB_n) - \Phi(x)|, \quad \Delta_n = \sup_x \Delta_n(x).$$

Всюду далее символ $\mathbb{I}(A)$ будет обозначать индикаторную функцию события A .

Пусть \mathcal{G} – класс вещественных функций $g(x)$ аргумента $x \in \mathbb{R}$ таких, что

- функция $g(x)$ четна;
- функция $g(x)$ неотрицательна при всех x и $g(x) > 0$ при $x > 0$;
- функция $g(x)$ не убывает при $x > 0$;
- функция $x/g(x)$ не убывает при $x > 0$.

Классу \mathcal{G} принадлежат, например, функции $g(x) \equiv \text{const} > 0$, $g(x) = |x|^\delta$ при $0 \leq \delta \leq 1$, $g(x) = \min\{|x|, \text{const}\}$.

В 1963 г. М. Кац [11] доказал, что, какой бы ни была функция $g \in \mathcal{G}$, если случайные величины X_1, X_2, \dots одинаково распределены, причем $\mathbf{E}X_1^2 g(X_1) < \infty$, то существует конечная положительная абсолютная постоянная C_1 такая, что

$$\Delta_n \leq C_1 \cdot \frac{\mathbf{E}X_1^2 g(X_1)}{\sigma_1^2 g(\sigma_1 \sqrt{n})}. \quad (1)$$

В 1965 г. этот результат был обобщен В. В. Петровым [3] на случай неодинаково распределенных случайных величин (также см. [4]): какой бы ни была функция $g \in \mathcal{G}$, если $\mathbf{E}X_k^2 g(X_k) < \infty$, $k = 1, \dots, n$, то существует конечная положительная абсолютная постоянная C_2 такая, что

$$\Delta_n \leq \frac{C_2}{B_n^2 g(B_n)} \sum_{k=1}^n \mathbf{E}X_k^2 g(X_k). \quad (2)$$

В 1979 г. В. В. Петров [5] доказал неравномерный аналог неравенства (2) (также см. [6]): какой бы ни была функция $g \in \mathcal{G}$, если $\mathbf{E}X_k^2 g(X_k) < \infty$, $k = 1, \dots, n$, то существует конечная положительная абсолютная постоянная C_3 такая, что для любого $x \in \mathbb{R}$

$$\Delta_n(x) \leq \frac{C_3}{B_n^2(1 + |x|)^2 g(B_n(1 + |x|))} \sum_{k=1}^n \mathbf{E}X_k^2 g(X_k). \quad (3)$$

В частности, если случайные величины одинаково распределены и $\mathbf{E}X_1^2 g(X_1) < \infty$ для некоторой функции $g \in \mathcal{G}$, то существует абсолютная постоянная C такая, что для любого $x \in \mathbb{R}$

$$\Delta_n(x) \leq \frac{C \cdot \mathbf{E}X_1^2 g(X_1)}{\sigma_1^2(1 + |x|)^2 g(\sigma_1 \sqrt{n}(1 + |x|))}. \quad (4)$$

Неравенство (3) получено в [5] как следствие оценки

$$\Delta_n(x) \leq C \sum_{k=1}^n \left[\frac{\mathbf{E}X_k^2 \mathbb{I}(|X_k| \geq (1 + |x|)B_n)}{B_n^2(1 + |x|)^2} + \frac{\mathbf{E}|X_k|^3 \mathbb{I}(|X_k| < (1 + |x|)B_n)}{B_n^3(1 + |x|)^3} \right], \quad (5)$$

которая для случая одинаково распределенных слагаемых принимает вид

$$\Delta_n(x) \leq \frac{C}{\sigma^2(1 + |x|)^2} \cdot \mathbf{E}X_1^2 \min \left\{ 1, \frac{|X_1|}{\sigma \sqrt{n}(1 + |x|)} \right\}. \quad (6)$$

(Интересно отметить, что в свою очередь, неравенства (5) и (6) являются частными случаями неравенств (3) и (4) соответственно со специальной функцией $g(y) = \min\{|y|, 1\}$).

В 2001 году неравенство (5) было передоказано другим методом в статье [10]. В нескольких работах предпринимались попытки оценить значение константы C в неравенстве (5). В частности, в работах [14, 15] была получена оценка $C \leq 76,17$. Эта оценка была существенно уточнена в работе [8], где было показано, что в случае одинаково распределенных слагаемых константа не превосходит 39,25. Наконец, недавно было показано, что в случае одинаково распределенных слагаемых, рассматриваемом в настоящей статье, константа C не превосходит 36,62 (см. [1]).

Цель настоящей работы – распространить неравенство (4) на случайные суммы, в которых число слагаемых имеет биномиальное или пуассоновское распределение. При этом будет существенно использоваться подход, развитый в работах [1, 13]. Также на биномиальные и пуассоновские суммы будет перенесен результат В. В. Петрова, уточняющий упомянутые неравенства Каца–Петрова, из которого, в частности, вытекает, что если слагаемые имеют момент порядка $2 + \delta$ с $\delta \in (0, 1)$, то остаточный член

в соответствующей версии центральной предельной теоремы имеет более высокий порядок малости (по числу слагаемых и аргументу), нежели степенные порядки, устанавливаемые неравенствами типа Нагаева–Бикялиса.

Неравномерные оценки для биномиальных случайных сумм

Всюду далее рассматриваются независимые одинаково распределенные случайные величины X_1, X_2, \dots с $\mathbf{E}X_i = 0$ и $0 < \mathbf{E}X_i^2 \equiv \sigma^2 < \infty$. Пусть $p \in (0, 1]$ – произвольно. Пусть ξ_1, \dots, ξ_n – независимые случайные величины, такие что

$$\xi_j = \begin{cases} 1 & \text{с вероятностью } p, \\ 0 & \text{с вероятностью } 1 - p, \end{cases} \quad j = 1, \dots, n.$$

Случайная величина $N_{n,p} = \xi_1 + \dots + \xi_n$ может интерпретироваться как число успехов в схеме испытаний Бернулли с вероятностью успеха p . Случайная величина $N_{n,p}$ имеет биномиальное распределение с параметрами n и p :

$$\mathbf{P}(N_{n,p} = k) = C_n^k p^k (1-p)^{n-k}, \quad k = 0, \dots, n.$$

Предположим, что при каждом $n \in \mathbb{N}$ случайные величины $N_{n,p}, X_1, X_2, \dots$ взаимно независимы. В данном разделе основным объектом изучения будут *биномиальные случайные суммы* вида

$$S_{N_{n,p}} = X_1 + \dots + X_{N_{n,p}}.$$

При этом если $N_{n,p} = 0$, то $S_{N_{n,p}} = 0$.

Для $j \in \mathbb{N}$ введем случайные величины \tilde{X}_j , полагая

$$\tilde{X}_j = \begin{cases} X_j & \text{с вероятностью } p, \\ 0 & \text{с вероятностью } 1 - p. \end{cases}$$

Несложно видеть, что $\tilde{X}_j \stackrel{d}{=} \xi_j X_j$, где сомножители в правой части независимы (здесь и далее символ $\stackrel{d}{=}$ обозначает совпадение распределений).

Пусть $F(x)$ – общая функция распределения случайных величин X_j , $E_0(x)$ – функция распределения с единственным единичным скачком в нуле. Тогда, очевидно,

$$\mathbf{P}(\tilde{X}_j < x) = pF(x) + (1-p)E_0(x), \quad x \in \mathbb{R}, \quad j \in \mathbb{N}.$$

При этом $\mathbf{E}\tilde{X}_j = 0$,

$$\mathbf{D}\tilde{X}_j = \mathbf{E}\tilde{X}_j^2 = p\sigma^2. \quad (4)$$

Лемма 1. Для любых $n \in \mathbb{N}$ и $p_j \in (0, 1]$

$$S_{N_{n,p}} \stackrel{d}{=} \tilde{X}_1 + \dots + \tilde{X}_n, \quad (5)$$

где случайные величины в правой части (5) независимы.

Доказательство леммы 1 представляет собой простое упражнение на свойства характеристических функций.

С учетом (4) и (5) легко заметить, что $DS_{N_{n,p}} = n\sigma^2$. Обозначим

$$\Delta_{n,p}(x) = |\mathbf{P}(S_{N_{n,p}} < x\sigma\sqrt{np}) - \Phi(x)|.$$

В работе [1] с помощью леммы 1 и соотношения (6) доказано следующее утверждение.

Лемма 2 [1]. Для любых $n \in \mathbb{N}$ и $p \in (0, 1]$ справедливо неравенство

$$\begin{aligned} & \Delta_{n,p}(x) \leq \\ & \leq 36,62 \left[\frac{\mathbf{E}X_1^2 \mathbb{I}(|X_1| \geq (1 + |x|)\sigma\sqrt{np})}{\sigma^2(1 + |x|)^2} + \frac{\mathbf{E}|X_1|^3 \mathbb{I}(|X_1| < (1 + |x|)\sigma\sqrt{np})}{\sigma^3\sqrt{np}(1 + |x|)^3} \right] = \\ & = \frac{36,62}{\sigma^2(1 + |x|)^2} \cdot \mathbf{E}X_1^2 \min \left\{ 1, \frac{|X_1|}{\sigma\sqrt{np}(1 + |x|)} \right\}. \end{aligned}$$

Теорема 1. Предположим, что $\mathbf{E}X_1^2 g(X_1) < \infty$ для некоторой функции $g \in \mathcal{G}$. Тогда для любых $n \in \mathbb{N}$, $p \in (0, 1]$ и $x \in \mathbb{R}$ справедливо неравенство

$$\Delta_{n,p}(x) \leq \frac{36,62 \mathbf{E}X_1^2 g(X_1)}{\sigma^2(1 + |x|)^2 g((1 + |x|)\sigma\sqrt{np})}.$$

Доказательство. Имеем

$$\begin{aligned} \frac{\mathbf{E}X_1^2 \mathbb{I}(|X_1| \geq (1 + |x|)\sigma\sqrt{np})}{\sigma^2(1 + |x|)^2} &= \mathbf{E} \frac{X_1^2 g(X_1) \mathbb{I}(|X_1| \geq (1 + |x|)\sigma\sqrt{np})}{\sigma^2(1 + |x|)^2 g(X_1)} \leq \\ &\leq \frac{\mathbf{E}X_1^2 g(X_1) \mathbb{I}(|X_1| \geq (1 + |x|)\sigma\sqrt{np})}{\sigma^2(1 + |x|)^2 g((1 + |x|)\sigma\sqrt{np})}, \end{aligned} \quad (6)$$

$$\begin{aligned} & \frac{\mathbf{E} \{ |X_1|^3 \mathbb{I}(|X_1| < (1 + |x|)\sigma\sqrt{np}) \}}{\sigma^3\sqrt{np}(1 + |x|)^3} = \\ &= \mathbf{E} \left[\frac{|X_1|^2 g(X_1) \mathbb{I}(|X_1| < (1 + |x|)\sigma\sqrt{np})}{\sigma^3\sqrt{np}(1 + |x|)^3} \cdot \frac{|X_1|}{g(X_1)} \right] \leq \\ &\leq \frac{\mathbf{E} \{ |X_1|^2 g(X_1) \mathbb{I}(|X_1| < (1 + |x|)\sigma\sqrt{np}) \}}{\sigma^3\sqrt{np}(1 + |x|)^3} \cdot \frac{(1 + |x|)\sigma\sqrt{np}}{g((1 + |x|)\sigma\sqrt{np})} = \\ &= \frac{\mathbf{E} \{ |X_1|^2 g(X_1) \mathbb{I}(|X_1| < (1 + |x|)\sigma\sqrt{np}) \}}{\sigma^2(1 + |x|)^2 g((1 + |x|)\sigma\sqrt{np})}. \end{aligned} \quad (7)$$

Объединяя (6) и (7) с учетом леммы 2, завершаем доказательство теоремы.

Неравномерные оценки для пуассоновских случайных сумм

Пусть N_λ – случайная величина, имеющая распределение Пуассона с параметром $\lambda > 0$:

$$\mathbf{P}(N_\lambda = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \mathbb{N} \cup \{0\}.$$

Предположим, что при каждом λ случайные величины $N_\lambda, X_1, X_2, \dots$ независимы. Рассмотрим пуассоновскую случайную сумму

$$S_{N_\lambda} = X_1 + \dots + X_{N_\lambda}.$$

Если $N_\lambda = 0$, то полагаем $S_{N_\lambda} = 0$. Несложно убедиться, что $\mathbf{E}S_{N_\lambda} = 0$ и $\mathbf{D}S_{N_\lambda} = \lambda\sigma^2$. Точность нормальной аппроксимации для распределений пуассоновских случайных сумм изучалась многими авторами (см. исторические обзоры в работах [9, 12]). Равномерные оценки точности нормальной аппроксимации для распределений пуассоновских случайных сумм при ослабленных моментных условиях получены в статье [13]. Насколько известно автору, неравномерные оценки величины

$$\Delta_\lambda(x) = |\mathbf{P}(S_{N_\lambda} < x\sigma\sqrt{\lambda}) - \Phi(x)|$$

для такой ситуации впервые были получены в работе [1], где доказан следующий результат.

Лемма 3 [1]. Для любых $\lambda > 0$ и $x \in \mathbb{R}$ справедлива оценка

$$\begin{aligned} \Delta_{n,p}(x) &\leq \\ &\leq 36,62 \left[\frac{\mathbf{E}X_1^2 \mathbb{I}(|X_1| \geq (1 + |x|)\sigma\sqrt{\lambda})}{\sigma^2(1 + |x|)^2} + \frac{\mathbf{E}|X_1|^3 \mathbb{I}(|X_1| < (1 + |x|)\sigma\sqrt{\lambda})}{\sigma^3\sqrt{\lambda}(1 + |x|)^3} \right] = \\ &= \frac{36,62}{\sigma^2(1 + |x|)^2} \cdot \mathbf{E}X_1^2 \min \left\{ 1, \frac{|X_1|}{\sigma\sqrt{\lambda}(1 + |x|)} \right\}. \end{aligned}$$

Точно так же, как с помощью леммы 2 доказана теорема 1, с помощью леммы 3 доказывается следующий результат.

Теорема 2. Предположим, что $\mathbf{E}X_1^2 g(X_1) < \infty$ для некоторой функции $g \in \mathcal{G}$. Тогда для любых $\lambda > 0$ и $x \in \mathbb{R}$ справедливо неравенство

$$\Delta_\lambda(x) \leq \frac{36,62 \mathbf{E}X_1^2 g(X_1)}{\sigma^2(1 + |x|)^2 g((1 + |x|)\sigma\sqrt{\lambda})}.$$

Уточнение неравномерных оценок для специального подкласса класса \mathcal{G} . В статье [2] было доказано, что если $\mathbf{E}|X_1|^{2+\delta} < \infty$ при некотором $\delta \in (0, 1)$, то

$$\Delta_n(x) \leq \frac{Q(\sqrt{n}(1 + |x|))}{n^{\delta/2}(1 + |x|)^{2+\delta}},$$

где $Q(y)$ – ограниченная функция, такая что $Q(y) \rightarrow 0$ при $y \rightarrow \infty$. Другими словами, если $\mathbf{E}|X_1|^{2+\delta} < \infty$ при $0 < \delta < 1$, то $\Delta_n(x) = o(n^{-\delta/2}(1+|x|)^{-(2+\delta)})$. Условие $\delta < 1$ существенно, при $\delta = 1$ этот результат не имеет места.

В работе [7] упомянутая оценка была обобщена. А именно, пусть \mathcal{G}_0 – класс, содержащий функции $g(x)$ такие, что

- $g(x) \in \mathcal{G}$;
- $\lim_{x \rightarrow \infty} \frac{g(x^2)}{xg(x)} = 0$.

Классу \mathcal{G}_0 принадлежит, например, функция $g_\delta(x) \equiv |x|^\delta$ при $0 < \delta < 1$. В то же время функция $g_1(x) \equiv |x|$, принадлежащая классу \mathcal{G} , не входит в класс \mathcal{G}_0 . Еще одним примером функции из \mathcal{G}_0 является $g(x) = \exp\{(\log(e + |x|))^p\}$ при $0 < p < 1$ (см. [7]).

В работе [7] было доказано, что если $\mathbf{E}X_1^2g(X_1) < \infty$ для некоторой функции $g(x) \in \mathcal{G}_0$, то

$$\Delta_n(x) \leq \frac{Q(\sigma\sqrt{n}(1+|x|))}{(1+|x|)^2g(\sigma\sqrt{n}(1+|x|))}$$

для всех $n \in \mathbb{N}$ и $x \in \mathbb{R}$, где $Q(y)$ – ограниченная функция, такая, что $Q(y) \rightarrow 0$ при $y \rightarrow \infty$. Цель данного раздела – следуя логике доказательства этого результата в [7], перенести этот результат на биномиальные и пуассоновские случайные суммы и установить вид функции $Q(y)$.

Сначала рассмотрим биномиальные случайные суммы. Снова воспользуемся леммой 2.

Из неравенства (6) непосредственно вытекает, что

$$\frac{\mathbf{E}X_1^2\mathbb{I}(|X_1| \geq (1+|x|)\sigma\sqrt{np})}{\sigma^2(1+|x|)^2} \leq \frac{Q_1((1+|x|)\sigma\sqrt{np})}{\sigma^2(1+|x|)^2g((1+|x|)\sigma\sqrt{np})}, \quad (8)$$

где для $y > 0$

$$Q_1(y) = \mathbf{E}X_1^2g(X_1)\mathbb{I}(|X_1| \geq y). \quad (9)$$

При этом, если $\mathbf{E}X_1^2g(X_1) < \infty$, то $\lim_{y \rightarrow \infty} Q_1(y) = 0$.

Рассмотрим величину

$$\frac{\mathbf{E}|X_1|^3\mathbb{I}(|X_1| < (1+|x|)\sigma\sqrt{np})}{\sigma^3\sqrt{np}(1+|x|)^3} \equiv A_1 + A_2,$$

где

$$A_1 = \frac{\mathbf{E}|X_1|^3\mathbb{I}[|X_1| < \sqrt{\sigma\sqrt{np}(1+|x|)}]}{\sigma^3\sqrt{np}(1+|x|)^3},$$

$$A_2 = \frac{\mathbf{E}|X_1|^3 \mathbb{I}[\sqrt{\sigma\sqrt{np}(1+|x|)} \leq |X_1| < \sigma\sqrt{np}(1+|x|)]}{\sigma^3\sqrt{np}(1+|x|)^3}.$$

Используя тот же прием, что при доказательстве теоремы 1, получаем оценку

$$\begin{aligned} A_2 &\leq \frac{\sigma\sqrt{np}(1+|x|)\mathbf{E}X_1^2g(X_1)\mathbb{I}[|X_1| > \sqrt{\sigma\sqrt{np}(1+|x|)}]}{\sigma^3\sqrt{np}(1+|x|)^3g((1+|x|)\sigma\sqrt{np})} = \\ &= \frac{Q_2((1+|x|)\sigma\sqrt{np})}{\sigma^2(1+|x|)^2g((1+|x|)\sigma\sqrt{np})}, \end{aligned} \quad (10)$$

где

$$Q_2(y) = \mathbf{E}X_1^2g(X_1)\mathbb{I}(|X_1| > \sqrt{y}). \quad (11)$$

При этом, если $\mathbf{E}X_1^2g(X_1) < \infty$, то $\lim_{y \rightarrow \infty} Q_2(y) = 0$.

Наконец, рассмотрим A_1 . Имеем

$$\begin{aligned} A_1 &= \frac{\mathbf{E}|X_1|^3 \mathbb{I}[|X_1| < \sqrt{\sigma\sqrt{np}(1+|x|)}]}{\sigma^3\sqrt{np}(1+|x|)^3} = \\ &= \frac{1}{\sigma^3\sqrt{np}(1+|x|)^3} \mathbf{E} \left[X_1^2g(X_1) \cdot \frac{|X_1|}{g(X_1)} \cdot \mathbb{I}(|X_1| < \sqrt{\sigma\sqrt{np}(1+|x|)}) \right] \leq \\ &\leq \frac{\sqrt{\sigma\sqrt{np}(1+|x|)} \mathbf{E} \left[X_1^2g(X_1)\mathbb{I}(|X_1| < \sqrt{\sigma\sqrt{np}(1+|x|)}) \right]}{\sigma^3\sqrt{np}(1+|x|)^3g(\sqrt{\sigma\sqrt{np}(1+|x|)})} \leq \\ &\leq \frac{\mathbf{E}X_1^2g(X_1)}{\sigma^{5/2}(np)^{1/4}(1+|x|)^{5/2}g(\sqrt{\sigma\sqrt{np}(1+|x|)})}. \end{aligned} \quad (12)$$

Обозначим $y = \sqrt{\sigma\sqrt{np}(1+|x|)}$. Тогда, продолжая цепочку (12), получим

$$\begin{aligned} A_1 &\leq \frac{np\mathbf{E}X_1^2g(X_1)}{y^5g(y)} = \frac{1}{\sigma^2(1+|x|)^2g((1+|x|)\sigma\sqrt{np})} \cdot \frac{g(y^2)\mathbf{E}X_1^2g(X_1)}{yg(y)} \equiv \\ &\equiv \frac{Q_3((1+|x|)\sigma\sqrt{np})}{\sigma^2(1+|x|)^2g((1+|x|)\sigma\sqrt{np})}, \end{aligned} \quad (13)$$

где

$$Q_3(y) = \frac{g(y^2)\mathbf{E}X_1^2g(X_1)}{yg(y)}. \quad (14)$$

При этом, если $\mathbf{E}X_1^2g(X_1) < \infty$, то из свойств функции $g \in \mathcal{G}_0$ вытекает, что $\lim_{y \rightarrow \infty} Q_3(y) = 0$.

Из леммы 2 и неравенств (8), (10) и (13) вытекает следующее утверждение.

Теорема 3. *Предположим, что $EX_1^2 g(X_1) < \infty$ для некоторой функции $g \in \mathcal{G}_0$. Тогда для любых $n \in \mathbb{N}$, $p \in (0, 1]$ и $x \in \mathbb{R}$ справедливо неравенство*

$$\Delta_{n,p}(x) \leq \frac{Q((1+|x|)\sigma\sqrt{np})}{\sigma^2(1+|x|)^2 g((1+|x|)\sigma\sqrt{np})},$$

где $Q(y) = 36,62[Q_1(y) + Q_2(y) + Q_3(y)]$, а функции $Q_1(y)$, $Q_2(y)$ и $Q_3(y)$ определены в (9), (11) и (14) соответственно. При этом $\lim_{y \rightarrow \infty} Q(y) = 0$.

Рассуждениями, практически дословно повторяющими доказательство теоремы 3 (с заменой леммы 2 на лемму 3, а np на λ), аналогичный результат можно получить и для пуассоновских случайных сумм. А именно, справедлива следующая теорема.

Теорема 4. *Предположим, что $EX_1^2 g(X_1) < \infty$ для некоторой функции $g \in \mathcal{G}_0$. Тогда для любых $\lambda > 0$ и $x \in \mathbb{R}$ справедливо неравенство*

$$\Delta_\lambda(x) \leq \frac{Q((1+|x|)\sigma\sqrt{\lambda})}{\sigma^2(1+|x|)^2 g((1+|x|)\sigma\sqrt{\lambda})},$$

где $Q(y) = 36,62[Q_1(y) + Q_2(y) + Q_3(y)]$, а функции $Q_1(y)$, $Q_2(y)$ и $Q_3(y)$ определены в (9), (11) и (14) соответственно с формальной заменой np на λ . При этом $\lim_{y \rightarrow \infty} Q(y) = 0$.

Библиографический список

1. Королев В. Ю., Дорофеева А.В. О неравномерных оценках точности нормальной аппроксимации для распределений некоторых случайных сумм при ослабленных моментных условиях // Информатика и ее применения. – 2018. – Т. 12, Вып. 4. – С. 91–96.
2. Осипов Л.В., Петров В.В. Об оценке остаточного члена в центральной предельной теореме // Теория вероятностей и ее применения. – 1967. – Т. 12, Вып. 2. – С. 322–329.
3. Петров В.В. Одна оценка отклонения распределения суммы независимых случайных величин от нормального закона // Доклады АН СССР. – 1965. – Т. 160, Вып. 5. – С. 1013–1015.
4. Петров В. В. Суммы независимых случайных величин. – М.: Наука, 1972. – 416 с.
5. Петров В.В. Одна предельная теорема для сумм независимых неодинаково распределенных случайных величин // Записки научных семинаров ЛОМИ. – 1979. – Т. 85. – С. 188–192.
6. Петров В.В. Предельные теоремы для сумм независимых случайных величин. – М.: Наука, 1987. – 320 с.

7. *Петров В.В.* Об оценке остаточного члена в центральной предельной теореме // Записки научных семинаров ЛОМИ. – 2007. – Т. 341. – С. 142–146.
8. *Попов С.В.* Оценки скорости сходимости в центральной предельной теореме при ослабленных моментных условиях: Дис... канд. физ.-мат. наук. – М: МГУ, 2011.
9. *Шевцова И.Г.* О точности нормальной аппроксимации для обобщенных пуассоновских распределений // Теория вероятностей и ее применения. – 2014. – Т. 58, Вып. 1. – С. 138–158.
10. *Chen L.H.Y., Shao Q.M.* A non-uniform Berry–Esseen bound via Stein’s method // Probability Theory and Related Fields. – 2001. – Vol. 120. – P. 236–254.
11. *Katz M.* Note on the Berry–Esseen theorem // Annals of Math. Statist. – 1963. – Vol. 39, № 4. – P. 1348–1349.
12. *Korolev V.Yu., Shevtsova I.G.* An improvement of the Berry–Esseen inequality with applications to Poisson and mixed Poisson random sums // Scandinavian Actuarial Journal. – 2012. – № 2. – P. 81–105.
13. *Korolev V., Dorofeeva A.* Bounds of the accuracy of the normal approximation to the distributions of random sums under relaxed moment conditions // Lithuanian Mathematical Journal. – 2017. – Vol. 57, № 1. – P. 38–58.
14. *Neammanee K., Thongtha P.* Improvement of the non-uniform version of the Berry–Esseen inequality via Paditz–Shiganov theorems // Journal of Inequalities in Pure and Applied Mathematics. – 2007. – Vol. 8, № 4, Art. 92.
15. *Thongtha P., Neammanee K.* Refinement of the constants in the non-uniform version of the Berry–Esseen theorem // Thai Journal of Mathematics. – 2007. – Vol. 5. – P. 1–13.

Katz-Petrov-type inequalities for some random sums¹

A.V. Dorofeeva

Faculty of Computational Mathematics and Cybernetics of Moscow State University; 119991, Moscow, GSP-1, Leninskie Gory, Moscow State University, Faculty of Computational Mathematics and Cybernetics; alex.dorofeyeva@gmail.com

Abstract. *Analogues of Katz-Petrov inequalities, setting non-uniform convergence rate estimates in central limit theorem for sums of a random number of independent identically distributed random variables for cases where summation index (number of terms in sum) has a binomial or Poisson distribution and stochastically independent of the terms, are proven. The case is considered when weak moment conditions are imposed on the terms, in particular, the terms may not have moments higher than the second. Constants estimates are written explicitly. V.V. Petrov’s result, shurping Katz-Petrov inequalities, are extended on binomial and Poisson random sums. From this result it follows that in case when terms have order of moments equals $2 + \delta$ with $\delta \in (0, 1)$, remainder member in the corresponding version of the central limit theorem has more high order of smallness (in terms of the number of terms and argument) than power orders established by Nagaev-Bikelis-type inequalities.*

Key words: *central limit theorem; normal approximation; random sum; binomial distribution; Poisson distribution; convergence rate.*

УДК 519.24

Применение EM-алгоритма для обработки циклических сигналов¹

Т.В. Захарова

Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, lsa@cs.msu.ru

Аннотация. Часто в прикладных задачах анализа данных возникает необходимость обработки циклических сигналов. Такие задачи актуальны, например, в медицине, где при обследованиях организма регистрируется большое число разнообразных биосигналов. Существующие методы позволяют эффективно обрабатывать слабо зашумленные нестационарные сигналы. Биосигналы же сложны тем, что имеют низкое отношение сигнал/шум. В данной работе предложен новый метод, основанный на комбинации CUSUM-статистик и стохастического EM-алгоритма разделения смесей вероятностных распределений. Он показывает высокую точность выявления момента разладки в сигнале. Новый метод применялся для изучения активности головного мозга, а именно определения расположения и взаимосвязи его функциональных зон.

Ключевые слова: миограмма; смеси вероятностных распределений; стохастический EM-алгоритм; CUSUM-статистики.

1. Введение

Точная локализация невосполнимых областей в мозге является основой планирования хирургического вмешательства. Она необходима для минимизации риска возникновения постоперационных осложнений. Кроме того, полученная информация может использоваться для того, чтобы понять, какие из зон головного мозга затронуты болезнью или повреждены. В силу индивидуальной пластичности мозга или вследствие полученных травм и различных поражений центральной нервной системы предварительная локализация по анатомическим маркерам не может быть выполнена с необходимой точностью. С этой целью используется магнитоэнцефалография в комбинации с магнитно-резонансной томографией.

¹Работа выполнена при поддержке РФФИ (проект 18-07-00252).

©Захарова Т.В., 2018

Регистрируемое магнитное поле, возникающее в результате электрической активности мозга, очень сильно зашумлено. В роли шума в данном случае выступает суперпозиция аппаратного (от датчиков, усилителей, аналого-цифровых преобразователей, сетевые помехи и т.д.) и физиологического (фоновая мозговая активность) шумов. Для увеличения отношения полезного сигнала к шуму в медицинских исследованиях принято использовать метод возбужденных потенциалов, состоящий из многократного повторения эксперимента и дальнейшего усреднения результатов [1, 5]. Сигналы, полученные таким способом, имеют циклическую нестационарную природу. Перейдем к более точной постановке задачи.

2. Постановка задачи

Кратко опишем схему эксперимента. Пациент кладет руку на стол и в течение нескольких минут нажимает пальцем на кнопку, тем самым возбуждая зоны интереса в головном мозге. В то же время множество сенсоров на поверхности головы записывают магнитоэнцефалограмму (рис.1). К каждой записи эксперимента прилагается дополнительная информация, это может быть миограмма (рис.2), регистрирующая электрические сигналы мышечных сокращений, или внешние данные о положении пальца (с акселерометра, кнопки, фотоэлемента и пр.). Все сигналы берутся синхронно со строго определенной частотой дискретизации.

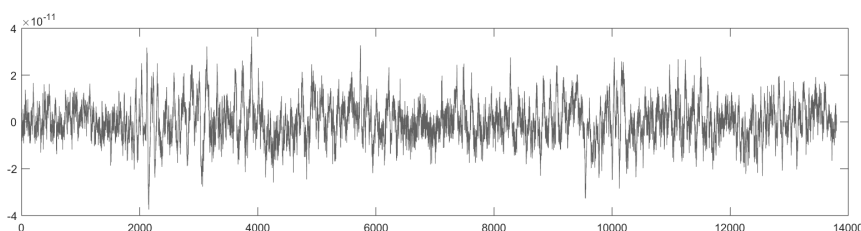


Рис. 1: Сигнал МЭГ

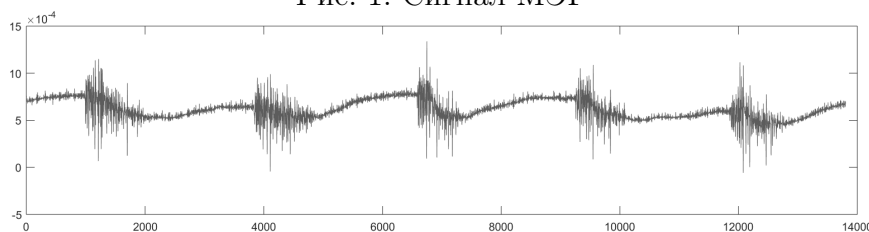


Рис. 2: Сигнал миограммы

На МЭГ-сигнале различить периоды покоя и движения не представляется возможным, а на соответствующем ему сигнале миограммы явно различимы все совершенные движения, хотя сигнал и зашумлен.

Главной целью анализа экспериментальных данных является обнаружение участков первичной моторной коры головного мозга, ответственной за подачу команды на движение. Она относится к зонам мозга, не способным к саморегенерации, и любое ее повреждение может лишить человека возможности управлять какой-либо группой мышц.

Усреднив сигнал МЭГ по точкам начала движения, можно определить магнитометры, на которых отклик на движение является наибольшим, и далее для локализации активных зон решить так называемую обратную задачу, то есть найти источник сигнала по характеристикам поля, порожденного данным источником. Методы решения обратных задач подробно рассмотрены в работах [2, 6].

В силу неустойчивости обратных задач по начальным данным неточное определение точек привязки может привести к значительным ошибкам локализации. Именно поэтому проблеме поиска опорных точек уделяется так много внимания.

3. Задача о разладке

Будем интерпретировать сигнал миограммы как наблюдения за случайным процессом, тогда опорную точку на отдельно взятой эпохе можно найти, решив так называемую задачу о разладке [3]. Приведем ее классическую постановку.

Пусть последовательно во времени проводятся измерения информации о случайном процессе, и в какой-то заранее неизвестный момент времени происходит изменение его вероятностных характеристик. Задача обнаружения произошедшего изменения быстрее всего носит название задачи скорейшего обнаружения момента появления разладки.

Один из методов ее решения приведен в [7]. Предположим, что

$$x_1, x_2, \dots, x_{\theta-1}, x_{\theta}, x_{\theta+1}, \dots$$

– это наблюдения над независимыми случайными величинами

$$\xi_1, \xi_2, \dots, \xi_{\theta-1}, \xi_{\theta}, \xi_{\theta+1}, \dots$$

такими, что в моменты времени $k = 1, 2, \dots, \theta - 1$ они имеют распределение с плотностью $f_{\infty}(x)$, а в моменты $k = \theta, \theta + 1, \dots$ – с плотностью $f_0(x)$. Параметр θ считается заранее не известным и принимающим значения $1, \dots, \infty$.

Для $n \geq 1, \theta > 1$ положим

$$p_{\theta}(x_1, \dots, x_n) = f_{\infty}(x_1) \cdot \dots \cdot f_{\infty}(x_{\theta-1}) f_0(x_{\theta}) \cdot \dots \cdot f_0(x_n).$$

При $\theta = 1$

$$p_1(x_1, \dots, x_n) = f_0(x_1) \cdots f_0(x_n)$$

и при $\theta = \infty$

$$p_\infty(x_1, \dots, x_n) = f_\infty(x_1) \cdots f_\infty(x_n).$$

Введем следующую статистику:

$$\begin{aligned} \gamma_n &= \sup_{\theta \geq 1} \frac{p_\theta(x_1, \dots, x_n)}{p_\infty(x_1, \dots, x_n)} = \\ &= \max \left\{ \frac{p_1(x_1, \dots, x_n)}{p_\infty(x_1, \dots, x_n)}, \frac{p_1(x_2, \dots, x_n)}{p_\infty(x_2, \dots, x_n)}, \dots, \frac{p_1(x_n)}{p_\infty(x_n)}, 1 \right\} = \\ &= \max \left\{ 1, \max_{1 \leq \theta \leq n} \prod_{k=\theta}^n \frac{f_0(x_k)}{f_\infty(x_k)} \right\} \end{aligned}$$

и ПОЛОЖИМ

$$T_n = \log \gamma_n = \max \left\{ 0, \max_{1 \leq \theta \leq n} \sum_{k=\theta}^n \log \frac{f_0(x_k)}{f_\infty(x_k)} \right\} = \max \left\{ 0, \max_{1 \leq \theta \leq n} \sum_{k=\theta}^n \zeta_k \right\}, \quad (1)$$

где

$$\zeta_k = \log \frac{f_0(x_k)}{f_\infty(x_k)}.$$

Введенные статистики $T_n, n \geq 1$, обладают важным рекуррентным свойством

$$T_n = \max[0, T_{n-1} + \zeta_n], \quad (2)$$

которое удобно использовать при реализации алгоритма обнаружения разладки.

Статистики $\gamma = (\gamma_n)_{n \geq 1}$ называют обобщенными отношениями правдоподобия, а $T = (T_n)_{n \geq 1}$ – статистиками CUSUM (англ. *cumulative sums*).

На рис.3 приведен типичный вид графика T_n . При $n < \theta$ она принимает неотрицательные значения, близкие к нулю, а при $n > \theta$ начинает в среднем возрастать.

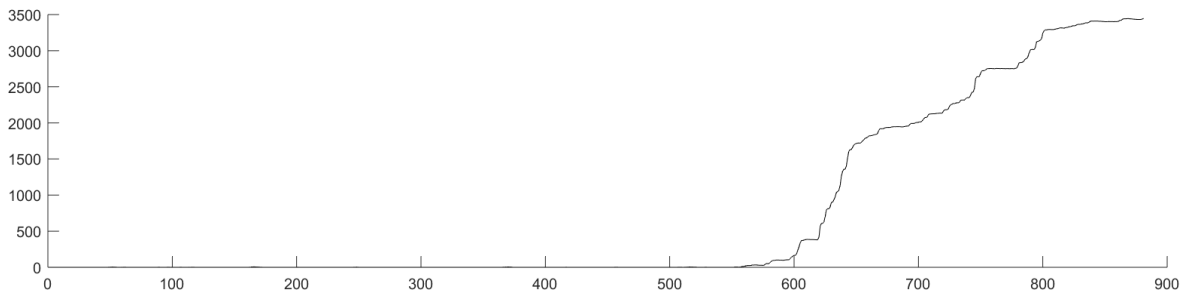


Рис. 3: График статистики CUSUM

Для применения этих статистик к реальным данным надо уметь хорошо оценивать плотности $f_0(x)$ и $f_\infty(x)$, которые точно не известны. На практике плотности распределения имеют сложную форму и не принадлежат одному типу распределения. В таких случаях плотности можно аппроксимировать смесью распределений.

4. Базовые определения и утверждения

Для лучшего понимания излагаемого материала приведем основные определения и утверждения, изложенные в книге [4].

4.1 Основные определения

Рассмотрим функцию $F(x, y)$, определенную на множестве $\mathbb{R} \times \mathbb{Y}$. Полагаем для простоты, что \mathbb{Y} – это некоторое подмножество m -мерного евклидова пространства, $m \geq 1$, снабженное борелевской σ -алгеброй Σ . Дополнительно предположим, что при каждом фиксированном y функция $F(x, y)$ является функцией распределения по x , а при каждом фиксированном x функция $F(x, y)$ измерима по y , то есть для любых $x \in \mathbb{R}$ и $c \in \mathbb{R}$ выполнено условие $\{y : F(x, y) < c\} \in \Sigma$. Пусть Q – вероятностная мера, определенная на измеримом пространстве (\mathbb{Y}, Σ) . Функция распределения

$$H(x) = \int_{\mathbb{Y}} F(x, y) Q(dy), \quad x \in \mathbb{R},$$

называется *смесью* функции распределения $F(x, y)$ по y относительно Q . Распределение $F(x, y)$ называется *смешиваемым*, мера Q задает *смешивающее* распределение. В случае, если \mathbf{Y} – m -мерная тождественная случайная величина ($\mathbf{Y}(y) \equiv y, y \in \mathbb{Y}$), определенная на вероятностном пространстве (\mathbb{Y}, Σ, Q) , то функция распределения $H(x)$ может быть записана в виде

$$H(x) = \mathbb{E}F(x, \mathbf{Y}), \quad x \in \mathbb{R}.$$

Если случайный вектор \mathbf{Y} имеет дискретное распределение, то есть принимает значения y_1, y_2, \dots с вероятностями p_1, p_2, \dots , то получаем *дискретную* смесь:

$$H(x) = \mathbb{E}F(x, \mathbf{Y}) = \sum_{j \geq 1} p_j F(x, y_j), \quad x \in \mathbb{R}.$$

Функции распределения $F(x, y_j)$ называют *компонентами* смеси $H(x)$, а числа p_j – весами соответствующих компонент. В случае, если в дискретной смеси число ненулевых весов конечно, дискретная смесь называется *конечной*.

Вообще говоря, разным значениям параметра y могут соответствовать функции $F(x, y)$, задающие разные типы распределения. Если смесь дискретна и $F(x, y_j) = F_j(x)$, $j = 1, 2, \dots$, а компоненты F_j абсолютно непрерывны и имеют плотности f_j , то смесь $H(x)$ также абсолютно непрерывна с плотностью

$$h(x) = \mathbb{E}f(x, \mathbf{Y}) = \sum_{j \geq 1} p_j f_j(x), \quad x \in \mathbb{R}.$$

Рассмотрим один важный тип смесей – сдвиг/масштабные смеси. Положим $m = 2$ и пусть вектор y имеет вид $y = (u, v)$, где $u > 0$ – параметр масштаба и $v \in \mathbb{R}$ – параметр сдвига, так что функция распределения $F(x, y)$ представима в виде

$$F(x, y) = F\left(\frac{x - v}{u}\right), \quad x \in \mathbb{R}.$$

Тогда функция распределения

$$H(x) = \int_{\mathbb{Y}} F\left(\frac{x - v}{u}\right) Q(du, dv), \quad x \in \mathbb{R} \quad (3)$$

называется *сдвиг/масштабной смесью* функции распределения F относительно Q .

Для определения дискретной сдвиг/масштабной смеси функции распределения $F(x)$ положим $y_j = (\sigma_j, a_j)$, где $a_j \in \mathbb{R}$, $\sigma_j > 0$, $j = 1, \dots, k$, и получим

$$H(x) = \sum_{j=1}^k p_j F\left(\frac{x - a_j}{\sigma_j}\right), \quad x \in \mathbb{R} \quad (4)$$

и соответствующую ей смесь плотностей

$$h(x) = \sum_{j=1}^k \frac{p_j}{\sigma_j} f\left(\frac{x - a_j}{\sigma_j}\right), \quad x \in \mathbb{R}.$$

4.2 Идентифицируемость смесей

Пусть функция $F(x, y)$ определена на множестве $\mathbb{R} \times \mathbb{Y}$, множество \mathbb{Y} и функция F удовлетворяют условиям из определения смеси функций распределения. Пусть \mathcal{Q} – семейство случайных величин, принимающих значения во множестве \mathbb{Y} . Обозначим

$$\mathcal{H} = \{H_Q(x) = \mathbb{E}F(x, Q), x \in \mathbb{R} : Q \in \mathcal{Q}\}.$$

Семейство \mathcal{H} , определяемое ядром F и множеством \mathcal{Q} , называется *идентифицируемым*, если из равенства

$$\mathbb{E}F(x, Q_1) = \mathbb{E}F(x, Q_2), \quad x \in \mathbb{R}, \quad Q_1, Q_2 \in \mathcal{Q}$$

следует, что $Q_1 \stackrel{d}{=} Q_2$.

Теперь сузим определение идентифицируемости на класс конечных сдвиг/масштабных смесей. Семейство смесей

$$\mathcal{H} = \left\{ \sum_{j=1}^k p_j F\left(\frac{x - a_j}{\sigma_j}\right) : \sum_{j=1}^k p_j = 1; p_j \geq 0, a_j \in \mathbb{R}, \sigma_j > 0, j = 1, \dots, k \right\},$$

порожденное ядром F , идентифицируемо, если из равенства

$$\sum_{j=1}^k p_j F\left(\frac{x - a_j}{\sigma_j}\right) = \sum_{i=1}^m q_i F\left(\frac{x - b_i}{\delta_i}\right)$$

вытекает, что

- 1) $k = m$;
- 2) для каждого индекса $j \in \{1, \dots, k\}$ существует индекс $i \in \{1, \dots, k\}$ такой, что

$$p_j = q_i, \quad a_j = b_i, \quad \sigma_j = \delta_i.$$

Справедливо следующее утверждение, доказанное Г. Тейчером: *семейство конечных сдвиг/масштабных смесей нормальных законов идентифицируемо*.

Таким образом, мы можем сделать вывод, что задача разделения конечных сдвиг/масштабных смесей нормальных законов является корректной, так как может иметь лишь единственное решение.

4.3 EM-алгоритм и его модификации

Попытка разделить смесь, используя непосредственно принцип максимума правдоподобия, приводит к слишком громоздкой оптимизационной задаче. Обойти эту трудность позволяет EM-алгоритм.

Пусть \mathbf{X} и \mathbf{Y} – случайные величины, заданные на одном измеримом пространстве (Ω, \mathcal{A}) и принимающие значения в \mathbb{R}^n и \mathbb{R}^m соответственно, $n, m \geq 1$. Иначе говоря, (\mathbf{X}, \mathbf{Y}) – $(n + m)$ -мерный случайный вектор.

Предположим, что на σ -алгебре \mathcal{A} задано семейство вероятностных мер $\{P_\theta : \theta \in \Theta\}$, где Θ – множество произвольной природы.

Пусть \mathcal{B}_n и \mathcal{B}_m – борелевские σ -алгебры на \mathbb{R}^n и \mathbb{R}^m соответственно, $\mu_{\mathbf{X}}$ и $\mu_{\mathbf{Y}}$ – σ -конечные меры, определенные на \mathcal{B}_n и \mathcal{B}_m . Тогда определим меру μ , заданную на борелевской σ -алгебре $\sigma(\mathcal{B})$, порожденной множеством

измеримых прямоугольников $\mathcal{B} = \{A \times B : A \in \mathcal{B}_n, B \in \mathcal{B}_m\}$, как произведение мер $\mu_{\mathbf{X}}$ и $\mu_{\mathbf{Y}}$:

$$\mu(A \times B) = \mu_{\mathbf{X}}(A) \cdot \mu_{\mathbf{Y}}(B), \quad A \in \mathcal{B}_n, B \in \mathcal{B}_m.$$

Предположим, что при каждом $\theta \in \Theta$ совместное распределение $(n + m)$ -мерного случайного вектора (\mathbf{X}, \mathbf{Y}) абсолютно непрерывно относительно меры μ , соответствующую плотность распределения обозначим $f_{\theta}(x, y)$. Иными словами, справедливо представление

$$P_{\theta}(\mathbf{X} \in A, \mathbf{Y} \in B) = \int_{A \times B} f_{\theta}(x, y) \mu(dx \times dy), \quad \forall \theta \in \Theta, \forall A \in \mathcal{B}_n, \forall B \in \mathcal{B}_m.$$

Маргинальной плотностью случайной величины \mathbf{X} относительно меры $\mu_{\mathbf{X}}$ называется функция

$$f_{\theta}^{\mathbf{X}}(x) = \int_{\mathbb{R}^m} f_{\theta}(x, y) \mu_{\mathbf{Y}}(dy).$$

Считаем, что случайная величины \mathbf{X} имеет смысл наблюдаемых данных, а \mathbf{Y} – ненаблюдаемая случайная величина, играющая вспомогательную роль. Зная совместную плотность $f_{\theta}(x, y)$ и значение наблюдаемой величины $\mathbf{X} = x$, можно ввести *полную функцию правдоподобия*

$$L(\theta; x, y) = f_{\theta}(x, y), \quad \theta \in \Theta. \quad (5)$$

Классическую функцию правдоподобия

$$L(\theta; x) = f_{\theta}^{\mathbf{X}}(x) \quad (6)$$

в таком случае можно трактовать как *неполную* функцию правдоподобия.

EM-алгоритм представляет собой итерационную процедуру, в результате выполнения которой вычисляется последовательность значений $\{\theta^{(m)}\}_{m \geq 1}$ параметра θ , максимизирующая функции (5) и (6). Если известно значение $\theta^{(m)}$, то вычисление следующего значения $\theta^{(m+1)}$ происходит в два этапа.

1°. Этап вычисления математического ожидания (E-этап, англ. *expectation*).

Функцию $Q(\theta; \theta^{(m)})$ определим как условное математическое ожидание логарифма полной функции правдоподобия при известном значении наблюдаемой компоненты \mathbf{X} :

$$Q(\theta; \theta^{(m)}) = \mathbb{E}_{\theta^{(m)}}[\log f_{\theta}(\mathbf{X}, \mathbf{Y}) | \mathbf{X}]. \quad (7)$$

Здесь θ является аргументом функции $Q(\theta; \theta^{(m)})$, \mathbf{X} и $\theta^{(m)}$ являются параметрами, в последнем равенстве усреднение идет по \mathbf{Y} относительно меры $P_{\theta^{(m)}}$.

При известном значении $\mathbf{X} = x$ функция $Q(\theta; \theta^{(m)})$ вычисляется по формуле

$$Q(\theta; \theta^{(m)}) = \int_{\mathbb{R}^m} [\log f_{\theta}(x, y)] f_{\theta^{(m)}}(y|x) \mu_{\mathbf{Y}}(dy). \quad (8)$$

2°. Этап максимизации (М-этап, англ. *maximization*).

На этом этапе происходит вычисление

$$\theta^{(m+1)} = \arg \max_{\theta} Q(\theta; \theta^{(m)}).$$

Итерационный процесс останавливается при выполнении заранее выбранного условия, например, близости найденных значений параметров $\theta^{(m)}$ и $\theta^{(m-1)}$.

EM-алгоритм обладает свойством монотонности, то есть каждая его итерация гарантированно увеличивает функцию правдоподобия:

$$L(\theta^{(m+1)}; x) \geq L(\theta^{(m)}; x), \quad m \geq 1.$$

Также при определенных условиях регулярности построенная последовательность оценок параметров сходится к локальному максимуму функции правдоподобия $L(\theta; x)$. Вместе с тем свойство монотонности свидетельствует о сильной зависимости EM-алгоритма от выбора начального приближения.

Область его применения чрезвычайно широка – дискриминантный анализ, восстановление пропусков в данных, кластеризация, обработка сигналов и изображений. В данной работе мы рассматриваем его как инструмент разделения смеси распределений.

Пусть плотность наблюдаемой случайной величины \mathbf{X} имеет вид (для конечных смесей нормальных законов это условие выполнено)

$$f_{\theta}^{\mathbf{X}}(x) = \sum_{i=1}^k p_i \psi_i(x; t_i), \quad (9)$$

где $k \geq 1$ – заданное натуральное число, ψ_1, \dots, ψ_k – известные плотности распределения (*компоненты* смеси), $\theta = (p_1, \dots, p_k, t_1, \dots, t_k)$ – неизвестный параметр, $p_i \geq 0$, $i = 1, \dots, k$ – *веса* компонент, $p_1 + \dots + p_k = 1$, t_i , $i = 1, \dots, k$ – многомерные параметры.

Задачей разделения смеси (9) называют задачу статистического оценивания параметров $\theta = (p_1, \dots, p_k, t_1, \dots, t_k)$ по известным реализациям случайной величины \mathbf{X} .

Отметим особенности EM-алгоритма, существенно ограничивающие возможность его применения в чистом виде на практике. Максимизируемый функционал $Q(\theta)$ может иметь большое количество локальных экс-

тремумов, поэтому для EM-алгоритма характерны такие черты, как застревание в локальных максимумах, зависимость решения от начального приближения и медленная сходимость при его неудачном выборе. В связи с этим существует большое количество модификаций EM-алгоритма.

4.4 SEM-алгоритм

Как было сказано ранее, классический EM-алгоритм находит не глобальный максимум функции правдоподобия, а локальный, являющийся ближайшим к начальному приближению, т. е. относится к категории “жадных” алгоритмов. Одним из эффективных способов противодействия этому свойству является случайная, но целенаправленная “перетасовка” выборки на каждой итерации. Этот механизм лежит в основе модификации, носящей название SEM-алгоритма (англ. *stochastic EM-algorithm*).

Предположим, что помимо наблюдаемой случайной величины \mathbf{X} задана ненаблюдаемая случайная величина \mathbf{Y} , содержащая информацию о номерах компонент, в соответствии с которыми “генерируются” значения случайной величины \mathbf{X} . То есть при каждом наблюдении $j = 1, \dots, n$ сначала реализуется значение $y_j \in \{1, \dots, k\}$ случайной величины \mathbf{Y} – номер компоненты смеси, являющейся распределением случайной величины \mathbf{X} при j -ом наблюдении, а затем только значение x_j .

Считаем, что пары значений (x_j, y_j) являются стохастически независимыми реализациями пары случайных величин (\mathbf{X}, \mathbf{Y}) .

Представим ненаблюдаемую информацию в другой форме, по сути, эквивалентной старой. Будем считать, что каждому наблюдению x_j соответствует вектор $\vec{y}_j = (y_{1j}, y_{2j}, \dots, y_{kj})$, $j = 1, \dots, n$, где k – число компонент смеси, n – объем выборки, и

$$y_{ij} = \begin{cases} 1, & \text{если наблюдение } x_j \text{ порождено } i\text{-й компонентой смеси,} \\ 0, & \text{иначе.} \end{cases}$$

В терминах введенных величин $y = \vec{y}_j = (y_{1j}, y_{2j}, \dots, y_{kj})$, $j = 1, \dots, n$ логарифм полной функции правдоподобия для модели (9) принимает вид

$$\begin{aligned} \log L(\theta; x, y) &= \sum_{j=1}^n \sum_{i=1}^k y_{ij} \log[p_i \psi_i(x_j; t_i)] = \\ &= \sum_{i=1}^k \log p_i \sum_{j=1}^n y_{ij} + \sum_{i=1}^k \sum_{j=1}^n y_{ij} \log \psi_i(x_j; t_i). \end{aligned} \quad (10)$$

Векторы \vec{y}_j , $j = 1, \dots, n$ разбивают исходные наблюдения x на k кластеров $\mathcal{K}_1, \dots, \mathcal{K}_k$:

$$x = \mathcal{K}_1 \cup \dots \cup \mathcal{K}_k.$$

Для каждого $i = 1, \dots, k$ во множество \mathcal{K}_i входят те наблюдения x_j , которым соответствует $y_{ij} = 1$, причем $\mathcal{K}_i \cap \mathcal{K}_j = \emptyset$ при $i \neq j$. Обозначим число наблюдений, попавших в кластер \mathcal{K}_i , $i = 1, \dots, k$ через ν_i :

$$\nu_i = \sum_{j=1}^n y_{ij}.$$

Ясно, что $\nu_1 + \dots + \nu_k = n$. Продолжая (10), имеем

$$\log L(\theta; x, y) = \sum_{i=1}^k \nu_i \log p_i + \sum_{i=1}^k \sum_{j: x_j \in \mathcal{K}_i} \log \psi_i(x_j; t_i). \quad (11)$$

Предположим, что величины y_{ij} известны. Слагаемые в правой части (11) зависят от непересекающихся групп параметров, и, следовательно, максимизировать функцию правдоподобия по θ можно по каждому из слагаемых в отдельности. Максимум первого по набору $p_1 + \dots + p_k = 1$ согласно методу неопределенных множителей Лагранжа достигается при

$$p_i^* = \frac{\nu_i}{n}. \quad (12)$$

Далее заметим, что

$$\sum_{j: x_j \in \mathcal{K}_i} \log \psi_i(x_j; t_i) = \log \prod_{j: x_j \in \mathcal{K}_i} \psi_i(x_j; t_i) \equiv \log L_i(t_i; \mathcal{K}_i),$$

где $L_i(t_i; \mathcal{K}_i)$ – это функция правдоподобия параметра t_i , построенная по кластеру \mathcal{K}_i в предположении, что каждый элемент подвыборки имеет плотность распределения $\psi_i(x; t_i)$. Положим

$$t_i^* = \arg \max_t L_i(t; \mathcal{K}_i), \quad i = 1, \dots, k. \quad (13)$$

Таким образом, если мы умеем моделировать величины y_{ij} , то оценки наибольшего правдоподобия параметров принятой модели (9) определяются соотношениями (12) и (13).

5. Выбор начального приближения EM-алгоритма

Скорость сходимости и точность полученного решения для EM-алгоритма и его модификаций сильно зависит от того, из какой точки стартует итерационный процесс. На практике время, потраченное на поиск хорошего начального приближения, может оказаться незначительно малым по сравнению со временем, сэкономленным за счет уменьшения числа итераций.

Предположим, что параметры смеси (сдвига, масштаба или веса компонент) задают последовательность ξ_1, \dots, ξ_n одинаково распределенных случайных величин, имеющих нормальное распределение с известными параметрами: $\xi_i \sim N(a, \sigma^2)$. Будем считать, что случайные величины в последовательности зависимы и известен коэффициент корреляции между соседними с.в. $\alpha = \text{corr}(\xi_i, \xi_{i-1})$. Рассмотрим две соседние эпохи и обозначим соответствующие параметры смеси через ξ_1 и ξ_2 . Нас интересует вероятность того, что разброс значений параметров на соседних эпохах окажется больше заданного x :

$$\mathbb{P}(|\xi_2 - \xi_1| > x).$$

Обозначим $\eta = \xi_2 - \xi_1$. Случайная величина η представляет собой нормально распределенную с.в. с параметрами \tilde{a} и $\tilde{\sigma}^2$, где

$$\tilde{a} = \mathbb{E}\eta = \mathbb{E}(\xi_2 - \xi_1) = \mathbb{E}\xi_2 - \mathbb{E}\xi_1 = a - a = 0,$$

$$\tilde{\sigma}^2 = \mathbb{D}\eta = \mathbb{D}(\xi_2 - \xi_1) = \mathbb{D}\xi_2 + \mathbb{D}\xi_1 - 2\text{cov}(\xi_1, \xi_2) = \sigma^2 + \sigma^2 - 2\alpha\sigma\sigma = 2(1 - \alpha)\sigma^2.$$

Тогда искомую вероятность отклонения можно выразить через ф.р. стандартного нормального закона:

$$\begin{aligned} \mathbb{P}(|\xi_2 - \xi_1| > x) &= \mathbb{P}(|\eta| > x) = \mathbb{P}(\eta > x) + \mathbb{P}(\eta < -x) = \\ &= 1 - \Phi\left(\frac{x - \tilde{a}}{\tilde{\sigma}}\right) + \Phi\left(\frac{-x - \tilde{a}}{\tilde{\sigma}}\right) = 2 - 2\Phi\left(\frac{x}{\sqrt{2(1 - \alpha)}\sigma}\right) \end{aligned}$$

Теперь предположим, что в качестве начального приближения мы берем не значение параметра с предыдущей эпохи, а фиксированное значение, например, математическое ожидание $\mathbb{E}\xi_1 = a$. Найдём разброс в этом случае:

$$\mathbb{P}(|\xi_2 - a| > x) = \{\xi_2 - a \sim N(0, \sigma^2)\} = 2 - 2\Phi\left(\frac{x}{\sigma}\right).$$

Иными словами, разность параметров в первом случае представляет собой нормальную с.в. с нулевым математическим ожиданием и дисперсией $2(1 - \alpha)\sigma^2$, а во втором с дисперсией σ^2 .

Утверждение. Пусть параметры смеси задают последовательность ξ_1, \dots, ξ_n одинаково распределенных случайных величин, $\xi_i \sim N(a, \sigma^2)$. Рассмотрим две соседние эпохи и обозначим соответствующие параметры смеси через ξ_1 и ξ_2 . Если коэффициент корреляции между с.в. $\alpha = \text{corr}(\xi_1, \xi_2) \geq 0.5$, то $\mathbb{D}(|\xi_2 - \xi_1|) \leq \mathbb{D}(|\xi_2 - a|)$.

Замечание. При положительном коэффициенте корреляции $\alpha > 0.5$ предпочтительнее использование значения с предыдущей эпохи в качестве начального приближения.

6. SEM-алгоритм с начальным приближением с предыдущей эпохи

Вернемся к задаче определения точек привязки. Напомним, что нам удалось разбить весь массив данных на эпохи, содержащие по одному движению.

Благодаря внешнему сигналу акселерометра нам известна окрестность, в которой нужно искать опорную точку. “Отступив” от нее на некоторый интервал влево, мы гарантированно окажемся на участке покоя. Хотя плотность можно было бы оценивать по всей эпохе с самого ее начала, мы ограничим “обучающую выборку”, отобрав только наблюдения, ближайшие к моменту начала движения.

Аналогичную процедуру сделаем для оценки плотности сигнала на участке движения с тем лишь отличием, что отступ берется вправо.

Далее распределение приращений миограммы на участках покоя и движения были аппроксимированы сдвиг-масштабными смесями нормальных законов.

В данной версии SEM-алгоритма мы будем использовать другую “точку входа” в итерационный процесс, а именно E-этап. В качестве начальных приближений параметров p_i , a_i и σ_i будем брать оценки, полученные на выходе SEM-алгоритма на предыдущей эпохе (для всех эпох, кроме первой; для нее по-прежнему начинаем с S-этапа).

6.1 Поиск точек привязки

Итак, мы получили оценки плотности на участках покоя и движения (в обозначениях (1) это $f_\infty(x)$ и $f_0(x)$ соответственно), и все готово для решения задачи разладки. Используя соотношение (2), вычислим значения T_n . За искомым момент разладки примем момент первого пересечения статистики T_n с нулем, если идти справа (рис. 4).

На рисунке 5 показан исходный сигнал с расставленными опорными точками.

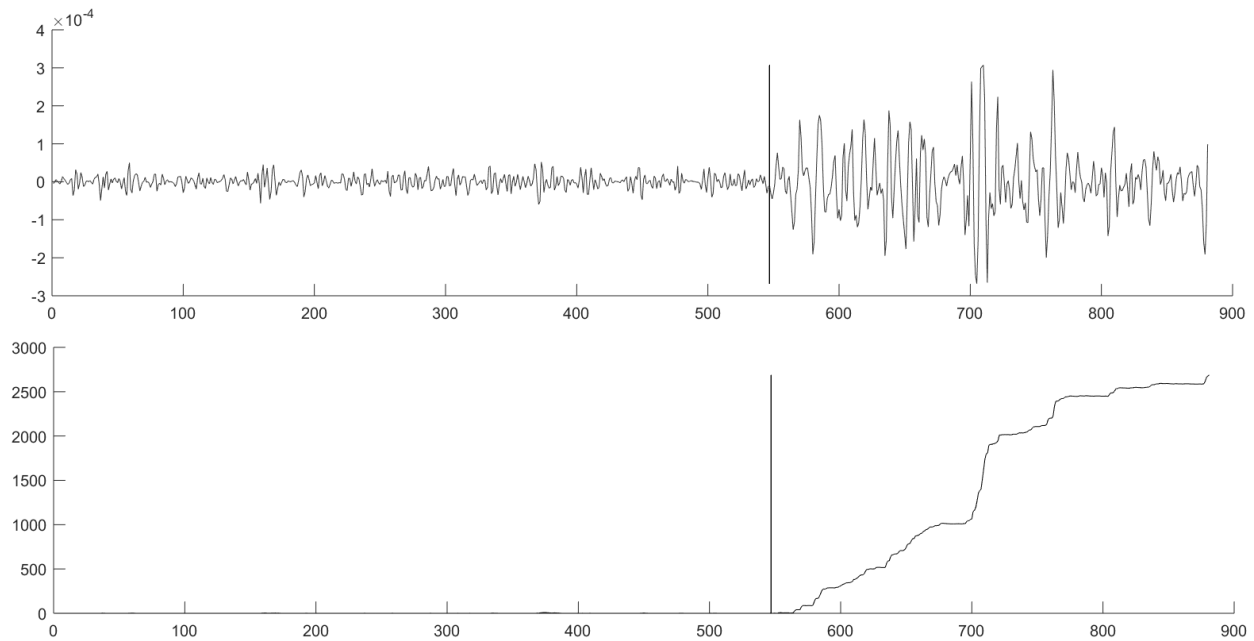


Рис. 4: Поиск момента разладки: приращения миограммы и график статистики T_n

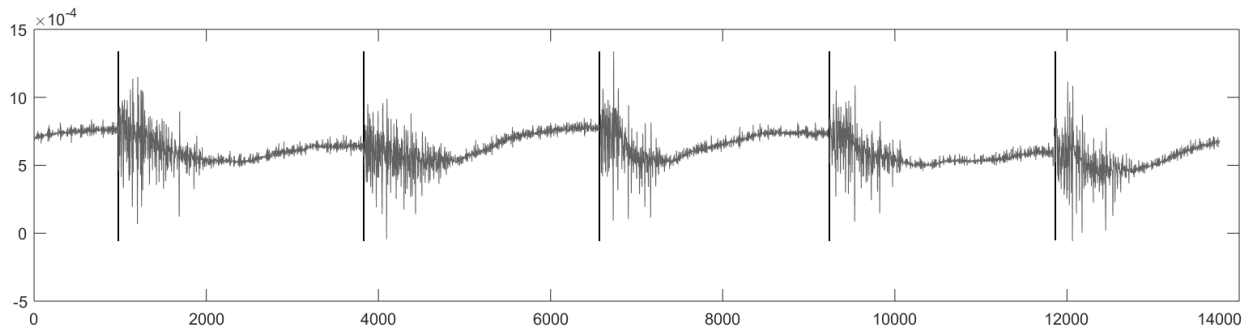


Рис. 5: Миограмма с найденными точками привязки

В данной статье предложена модификация EM-алгоритма, учитывающая особенности циклических нестационарных сигналов. Благодаря выбору начального приближения удалось значительно увеличить скорость сходимости итерационного процесса. Данный алгоритм применялся к задаче расстановки опорных точек по сигналу миограммы, играющей важную роль в локализации первичной моторной коры головного мозга.

В качестве математической модели сигнала миограммы были использованы конечные сдвиг-масштабные смеси нормальных распределений.

Математический аппарат смесей вероятностных законов может оказаться удобным инструментом для дальнейшего решения актуальных задач в важных областях современной медицины.

Библиографический список

1. Захарова Т. В., Никифоров С. Ю., Гончаренко М. Б., Драницына М. А., Климов Г. А., Хазиахметов М. Ш., Чаянов Н. В. Методы обработки сигналов для локализации невосполнимых областей головного мозга // Системы и средства информатики. – 2012. – Т. 22, №2. – С. 157–176.
2. Захарова Т. В., Гончаренко М. Б., Никифоров С. Ю. Метод решения обратной задачи магнитоэнцефалографии, основанный на кластеризации поверхности мозга // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. / Перм. гос. нац. иссл. ун-т. – Пермь, 2013. – Вып. 25. – С. 120–125.
3. Захарова Т. В., Подлесный М. М. Смеси нормальных законов в задаче поиска опорных точек по сигналу миограммы // Системы и средства информатики. – 2016. – Т. 26, № 3. – С. 106–121.
4. Королев В. Ю. EM-алгоритм, его модификации и их применение к задаче разделения смесей вероятностных распределений. Теоретический обзор. – М.: ИПИ РАН, 2007. – 94 с.
5. Fabiani M., Gratton G., Federmeier K. D. Event-Related Brain Potentials: Methods, Theory, and Applications // Handbook of Psychophysiology. Ed. Cacioppo J. T., Tassinari L. G., Berntson G. G. – Cambridge: Cambridge University Press. – 2007. – P. 85–119.
6. Бенинг В. Е., Драницына М. А., Захарова Т. В., Карпов П. И. Решение обратной задачи в многодипольной модели источников магнитоэнцефалограмм методом независимых компонент // Информатика и ее применения. – 2014. – Т. 8., Вып. 2. – С. 79–87.
7. Ширяев А. Н. Вероятностно-статистические методы в теории принятия решений. – М.: ФМОП, МЦНМО, 2011. – 144 с.

Processing cyclic signals by EM-algorithm

T.V. Zakharova

Faculty of Computational Mathematics and Cybernetics of Moscow State University; 119991, Moscow, GSP-1, Leninskie Gory, Moscow State University, Faculty of Computational Mathematics and Cybernetics; Institute for Informatics Problems of FRC IC RAS; e-mail: lsa@cs.msu.ru

Abstract. *Often, in data analysis applications, it is necessary to process cyclic signals. For example, when examining an organism, a large number of various biosignals are recorded. Existing methods allow to efficiently process low-noise non-stationary signals. Biosignals already are complex because they have a low signal-to-noise ratio. This paper proposes a new method based on a combination of CUSUM-statistics and a stochastic EM-algorithm for separating mixtures of probability distributions. It shows the high accuracy of detecting the moment of disorder in the signal. A new method is used to study the activity of the brain.*

Key words: myogram; mixtures of probability distributions; stochastic EM-algorithm; CUSUM-statistics.

УДК 519.2

Об асимптотически оптимальных размещениях станций обслуживания на прямой

Т.В. Захарова¹, И.М. Меньших²

¹Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, lsa@cs.msu.ru

²Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики

Аннотация. В статье рассматривается неклассическая задача теории массового обслуживания. На прямой возникают вызовы и для их обслуживания имеется n станций. Вызов поступает на ближайшую к нему станцию. Время обслуживания вызова есть некоторая возрастающая функция, зависящая от расстояния между вызовом и станцией. Требуется так расположить станции, чтобы минимизировать время обслуживания появляющихся вызовов. Найти оптимальные размещения станций для задач подобного рода удастся лишь в исключительных ситуациях. Но можно получить удовлетворительное для практики асимптотическое решение задачи. В статье приведен алгоритм построения асимптотически оптимальных второго порядка размещений для семейства критериев оптимальности с вещественным положительным параметром.

Ключевые слова: оптимальное размещение; станция обслуживания; зона влияния; критерий оптимальности.

1. Введение

Интерес к задачам размещения станций обслуживания вызван в большей степени практической потребностью оптимального распределения ограниченных ресурсов. Исследования этой задачи для нахождения наилучших размещений на различных областях проведены, например, в [1], [2], [3], [4].

Наиболее весомый прогресс на сегодняшний день достигнут в задачах размещения станций обслуживания на вещественной прямой и ее отрезках. Примером практического применения размещений на одномерном множестве является, например, расстановка автосервисов на магистралях,

а также выбор расположений ремонтных станций вдоль нефте- и газопроводов.

В данной статье приводятся теоретические выкладки, а также построенный на их основе алгоритм размещения станций обслуживания, обеспечивающий эффективное расположение станций (т.н. асимптотически оптимальное второго порядка размещение) для семейства критериев оптимальности. В качестве фундаментального факта использовано необходимое условие оптимального размещения на отрезке, полученное в работе [1] и развитое в [5]. В статье обобщен результат, полученный в [6], на случай вещественного положительного критериального параметра.

2. Постановка задачи

В пространстве возникают вызовы и для их обслуживания имеется n станций. Вызов поступает на ближайшую к нему станцию. Время обслуживания вызова есть некоторая возрастающая функция, зависящая от расстояния между вызовом и станцией. Требуется так расположить станции, чтобы минимизировать время обслуживания появляющихся вызовов.

- Задача поиска оптимального размещения состоит в нахождении такого расположения *заданного* количества станций, при котором достигается минимум среднего времени обслуживания.
- Задача поиска асимптотически оптимального размещения состоит в указании такого алгоритма расположения *произвольного* числа станций, чтобы при его увеличении качество размещения стремилось к идеальному.

Определение 1. Размещением n станций в пространстве назовем множество $\{x_1, \dots, x_n\}$ точек пространства, в которых эти станции расположены.

Обозначать размещение станций будем символом x , то есть $x = \{x_1, \dots, x_n\}$. Станцию обслуживания и точку пространства, где она расположена, будем обозначать одним и тем же символом.

Определение 2. Зоной влияния станции x_i назовем множество C_i тех точек пространства \mathbb{R} , для которых эта станция является ближайшей, т.е.

$$C_i = \{v \in \mathbb{R} : |v - x_i| \leq |v - x_j|, j = \overline{1, n}\}.$$

Точка ξ пространства \mathbb{R} , из которой поступает вызов, является случайной величиной с плотностью распределения ρ . Станции обслуживают вызовы только из своих зон влияния. Для оценивания оптимальности расположения этих станций задан критерий φ :

$$\varphi(x) = \mathbb{E} \min_{1 \leq i \leq n} |\xi - x_i|^s, \quad s > 0.$$

Определение 3. Размещение $x^* = \{x_1^*, \dots, x_n^*\}$ называется оптимальным по критерию φ среди размещений $x = \{x_1, \dots, x_n\}$, если

$$\varphi(x^*) = \min_x \varphi(x).$$

Определение 4. Пусть φ – некоторый критерий для размещений. Размещение $x = \{x_1, \dots, x_n\}$ называется асимптотически оптимальным порядка k , если

$$\lim_{n \rightarrow \infty} (\varphi(x^*) - \varphi(x))n^{s+k-1} = 0,$$

где $x^* = \{x_1^*, \dots, x_n^*\}$ – оптимальное размещение по критерию φ .

3. Асимптотически оптимальные второго порядка размещения на прямой

Теорема 1. Пусть $[d_1, d_{n+1}]$ – носитель плотности распределения вызовов ρ , положительной и непрерывно дифференцируемой $\hat{s} = [s]$ число раз, $[s]$ – целая часть числа $s > 0$. Тогда для всякого оптимального размещения $x^* = \{x_1^*, \dots, x_n^*\}$ и критерияльного параметра s выполняются соотношения:

$$x_i^* = d_i + \frac{1}{2}\Delta_i + o(\Delta_i), \quad i = \overline{1, n},$$

где $[d_i, d_{i+1}]$ – зона влияния станции x_i^* , $\Delta_i = d_{i+1} - d_i$.

Доказательство. Рассмотрим произвольную станцию x_i^* . В соответствии с необходимым условием оптимального размещения для произвольного параметра $s > 0$ имеет место равенство:

$$\int_{d_i}^{x_i^*} \rho(u)(x_i^* - u)^{s-1} du = \int_{x_i^*}^{d_{i+1}} \rho(u)(u - x_i^*)^{s-1} du. \quad (1)$$

Обозначим $f(u) = \rho(u)(x_i^* - u)^{s-1}$, $g(u) = \rho(u)(u - x_i^*)^{s-1}$. Выполним процедуру отделения от точки x_i^* для левой части (1):

$$\int_{d_i}^{x_i^*} f(u) du = \int_{d_i}^{x_i^* - \varepsilon} f(u) du + \int_{x_i^* - \varepsilon}^{x_i^*} f(u) du. \quad (2)$$

При этом $\varepsilon = \varepsilon(n) > 0$ выберем таким, что $x_i^* - d_i - \varepsilon > 0$, $d_{i+1} - x_i^* - \varepsilon > 0$ и

$$\lim_{n \rightarrow \infty} \frac{\varepsilon}{\min\{(x_i^* - d_i), (d_{i+1} - x_i^*)\}} = 0.$$

Тогда второй интеграл из правой части равенства (2) по теореме о среднем можно представить в следующем виде:

$$\int_{x_i^* - \varepsilon}^{x_i^*} f(u) du = \rho(\xi) \int_{x_i^* - \varepsilon}^{x_i^*} (x_i^* - u)^{s-1} du = o(x_i^* - d_i)^s, \quad \xi \in [x_i^* - \varepsilon, x_i^*].$$

Проделав аналогичные действия с правой частью (1) и получив соотношение

$$\int_{x_i^*}^{x_i^* + \varepsilon} g(u) du = o(d_{i+1} - x_i^*)^s,$$

представим необходимое условие оптимальности (1) в виде

$$\int_{d_i}^{x_i^* - \varepsilon} \rho(u) (x_i^* - u)^{s-1} du = \int_{x_i^* + \varepsilon}^{d_{i+1}} \rho(u) (u - x_i^*)^{s-1} du + o(\Delta_i^s), \quad (3)$$

учитывая неравенства $x_i^* - d_i < \Delta_i$, $d_{i+1} - x_i^* < \Delta_i$.

Так как ρ и Δ_i положительны, то из (1) по теореме о среднем следует, что $x_i^* \neq d_i$ и $x_i^* \neq d_{i+1}$. Таким образом, в этих точках существуют первые \hat{s} производных функций f и g . Далее, выполнив разложение функции f в левой части (2) по формуле Тейлора в точке d_i и учитывая, что

$$\int_{d_i}^{x_i^* - \varepsilon} (u - d_i)^k du = \frac{(x_i^* - d_i - \varepsilon)^{k+1}}{k+1},$$

имеем

$$\begin{aligned} \int_{d_i}^{x_i^* - \varepsilon} f(u) du &= f(d_i)(x_i^* - d_i - \varepsilon) + f'(d_i) \frac{(x_i^* - d_i - \varepsilon)^2}{2!} + \dots \\ &\dots + f^{(\hat{s}-1)}(d_i) \frac{(x_i^* - d_i - \varepsilon)^{\hat{s}}}{\hat{s}!} + o(x_i^* - d_i)^{\hat{s}}. \end{aligned}$$

Отсюда, поскольку при сделанных предположениях верно соотношение

$$f^{(k)}(d_i)(x_i^* - d_i - \varepsilon)^{k+1} = \rho(d_i)(x_i^* - d_i)^s (-1)^k (s-1) \dots (s-k) + o(\Delta_i^s)$$

для $k = 1, \dots, \hat{s} - 1$, получим равенство

$$\begin{aligned} \int_{d_i}^{x_i^* - \varepsilon} f(u) du &= \rho(d_i)(x_i^* - d_i)^s \left(1 - \frac{s-1}{2!} + \dots \right. \\ &\left. \dots + (-1)^{\hat{s}-1} \frac{(s-1) \dots (s - (\hat{s}-1))}{\hat{s}!} \right) + o(\Delta_i^s). \end{aligned}$$

Аналогично раскладывая функцию g в ряд Тейлора в точке d_{i+1} и пользуясь тем, что

$$\int_{x_i^* + \varepsilon}^{d_{i+1}} (u - d_{i+1})^k du = \frac{(-1)^k (d_{i+1} - x_i^* - \varepsilon)^{k+1}}{k+1},$$

получим соотношение

$$\begin{aligned} \int_{x_i^* + \varepsilon}^{d_{i+1}} g(u) du &= \rho(d_{i+1})(d_{i+1} - x_i^*)^s \left(1 - \frac{s-1}{2!} + \dots \right. \\ &\quad \left. \dots + (-1)^{\hat{s}-1} \frac{(s-1) \cdots (s - (\hat{s}-1))}{\hat{s}!} \right) + o(\Delta_i^s). \end{aligned}$$

Для наглядности дальнейших выкладок обозначим $a = x_i^* - d_i$, $b = d_{i+1} - x_i^*$. В этих обозначениях $\Delta_i = d_{i+1} - d_i = a + b$. Тогда из (2) с учетом полученных разложений следует, что

$$\rho(d_i)a^s = \rho(d_{i+1})b^s + o(\Delta_i^s).$$

По формуле конечных приращений $\rho(d_{i+1}) = \rho(d_i) + \rho'(\xi)\Delta_i$, $\xi \in (d_i, d_{i+1})$, поэтому

$$a^s - b^s = o(\Delta_i^s). \quad (4)$$

Преобразуем левую часть равенства (4):

$$a^s - b^s = (a+b)^s \left[\left(1 - \frac{b}{a+b} \right)^s - \left(1 - \frac{a}{a+b} \right)^s \right],$$

Воспользуемся разложением в ряд Маклорена бинома

$$(1+x)^r = \sum_{k=0}^{\infty} \binom{r}{k} x^k = 1 + rx + \dots + \frac{r(r-1)\cdots(r-k+1)}{k!} x^k + \dots,$$

которое справедливо при $|x| < 1$, $r > 0$. Полагая $r = s$, а x — равным последовательно $-\frac{b}{a+b}$ и $-\frac{a}{a+b}$, получим разложение

$$a^s - b^s = (a-b) \sum_{k=1}^{\infty} A_k,$$

где

$$A_k = (-1)^{k+1} \frac{s(s-1)\cdots(s-k+1)}{k!} \frac{a^{k-1} + \dots + b^{k-1}}{(a+b)^k}.$$

Очевидно, $\lim_{k \rightarrow \infty} A_k = 0$. При этом отношение

$$\left| \frac{A_{k+1}}{A_k} \right| = \frac{|s - k|(a^k + \dots + b^k)}{(k + 1)(a + b)(a^{k-1} + \dots + b^{k-1})}$$

меньше 1 при $k \geq \max\{\hat{s}, 2\}$. Это значит, что ряд $\sum_{k=1}^{\infty} A_k$ представляет собой сумму конечного числа слагаемых порядка $O(\Delta_i^{-1})$ и ряда Лейбница. Так как сумму последнего можно оценить первым членом, имеющим порядок $O(\Delta_i^{-1})$, получаем

$$a^s - b^s = (a - b)(a + b)^s O(\Delta_i^{-1}) = (a - b) O(\Delta_i^{s-1}).$$

Используя результат (4), можем заключить, что $a - b = o(\Delta_i)$. Таким образом

$$(x_i^* - d_i) - (d_{i+1} - x_i^*) = o(\Delta_i)$$

и

$$x_i^* = d_i + \frac{1}{2}\Delta_i + o(\Delta_i).$$

Последнее равенство верно для всех $i = \overline{1, n}$. Теорема доказана.

Теорема 2. Если плотность ρ имеет непрерывную вторую производную, то для всякой последовательности оптимальных размещений $\{x^*\}$ выполнено неравенство

$$\liminf_{n \rightarrow \infty} n^{s+1} \left(\varphi(x^*) - \frac{1}{(s+1)(2n)^s} \|\rho\| \right) \geq 0, \quad \|\rho\| = \left(\int p^{1/(s+1)}(u) du \right)^{s+1}.$$

Доказательство. Верно соотношение

$$\begin{aligned} \varphi(x^*) &= \sum_{i=1}^n \int_{\Delta_i} |u - x_i^*|^s \rho(u) du = \\ &= \sum_{i=1}^n \int_{\Delta_i} |u - x_i^*|^s (\rho(x_i^*) + \rho'(x_i^*)(u - x_i^*)) du + o(\Delta_i^{s+2}) = \\ &= \sum_{i=1}^n \frac{1}{s+1} \rho(x_i^*) ((x_i^* - d_i)^{s+1} + (d_{i+1} - x_i^*)^{s+1}) + \\ &+ \sum_{i=1}^n \frac{1}{s+2} \rho'(x_i^*) (-(x_i^* - d_i)^{s+2} + (d_{i+1} - x_i^*)^{s+2}) + \sum_{i=1}^n o(\Delta_i^{s+2}). \end{aligned}$$

Упростим это равенство, воспользовавшись теоремой 1:

$$\varphi(x^*) = \frac{1}{(s+1)2^s} \sum_{i=1}^n (\rho(x_i^*) \Delta_i^{s+1} + o(\Delta_i^{s+2})).$$

Используя неравенство $\sum_{i=1}^n x_i^{s+1} \geq n^{-s} \left(\sum_{i=1}^n x_i \right)^{s+1}$, получим

$$\varphi(x^*) \geq \frac{1}{(s+1)2^s} \left(\sum_{i=1}^n \rho^{1/(s+1)}(x_i^*) \Delta_i \right)^{s+1} n^{-s} + o\left(\frac{1}{n^{s+1}}\right).$$

Отсюда следует

$$\liminf_{n \rightarrow \infty} n^{s+1} \left(\varphi(x^*) - \frac{1}{(s+1)(2n)^s} \|\rho\| \right) \geq 0,$$

что завершает доказательство теоремы 2.

Алгоритм построения асимптотически оптимального размещения станций обслуживания.

Для удобства будем обозначать одним символом и множество, и его длину. Итак, введем обозначения: G – исходный отрезок, $M = \max_G \rho(u)$, $m = \min_G \rho(u)$. Выберем также числа a и b таким образом, чтобы выполнялись условия $0.5 < a \leq b < 1$, $3a - b > 1$. Например, $a = 0.6$, $b = 0.7$.

Далее разобьем отрезок G на непересекающиеся множества G_i , $i = \overline{1, \kappa}$, $\kappa = [n^b]$ так, чтобы выполнялись неравенства

$$G_i \leq \frac{G}{n^a}, \quad G_i \geq \frac{M^{1/(s+1)}G}{m^{1/(s+1)}n}, \quad i = \overline{1, \kappa}. \quad (5)$$

Отметим также, что числа n_i , определяемые соотношениями

$$n_i = \frac{\int_{G_i} \rho^{1/(s+1)}(u) du}{\int_G \rho^{1/(s+1)}(u) du} n, \quad i = \overline{1, \kappa}, \quad (6)$$

должны быть натуральными.

Последнего условия можно добиться, выбрав число n достаточно большим, в силу непрерывности интеграла и условия $\sum_{i=1}^{\kappa} n_i = n$. Заметим, что при этом $n_i \geq 1$ для $i = \overline{1, \kappa}$. Далее, каждое G_i делим на n_i равных отрезков Δ_i и помещаем n_i станций в их центры. Таким образом, для возрастающей последовательности значений n , алгоритм позволяет получить последовательность асимптотически оптимальных размещений второго порядка.

Теорема 3. *Построенная последовательность размещений является асимптотически оптимальной второго порядка.*

Доказательство. Оценим сверху значение критерия φ на данном размещении $x = \{x_1, \dots, x_n\}$:

$$\varphi(x) = \int \min_{1 \leq i \leq n} |u - x_i|^s \rho(u) du \leq \sum_{i=1}^{\kappa} \int_{G_i} \min_{1 \leq j \leq n_i} |u - x_j|^s \rho(u) du. \quad (7)$$

Разложим ρ на отрезке G_i по формуле Тейлора относительно середины v_i :

$$\rho(u) = \rho(v_i) + \rho'(v_i)(u - v_i) + O(G_i^2) = \bar{\rho}(u) + O(G_i^2).$$

Тогда слагаемые в (6) будут иметь следующий вид:

$$\begin{aligned} \int_{G_i} \min_{1 \leq j \leq n_i} |u - x_j|^s \rho(u) du &= \sum_{j=1}^{n_i} \int_{\Delta_j} |u - x_j|^s (\bar{\rho}(u) + O(G_i^2)) du = \\ &= \sum_{j=1}^{n_i} \int_{\Delta_j} |u - x_j|^s \bar{\rho}(u) du + O(G_i^2) \sum_{j=1}^{n_i} \frac{\Delta_j^{s+1}}{(s+1)2^s} = \\ &= \sum_{j=1}^{n_i} \int_{\Delta_j} |u - x_j|^s \bar{\rho}(u) du + \frac{n_i \Delta_j^{s+1} O(G_i^2)}{(s+1)2^s}. \end{aligned}$$

Последнее равенство верно в силу того, что Δ_j для $j = \overline{1, n_i}$ равны между собой. Подставляя $\Delta_j = G_i/n_i$ и внося константы под знак O , получаем:

$$\frac{n_i \Delta_j^{s+1} O(G_i^2)}{(s+1)2^s} = O\left(\frac{G_i^{s+3}}{n_i^s}\right).$$

Ввиду линейности $\bar{\rho}(u)$ на множестве Δ_j она представима как $\bar{\rho}(u) = \bar{\rho}(x_j) + \rho'(v_i)(u - x_j)$, поэтому

$$\begin{aligned} \sum_{j=1}^{n_i} \int_{\Delta_j} |u - x_j|^s \bar{\rho}(u) du &= \sum_{j=1}^{n_i} \int_{\Delta_j} |u - x_j|^s (\bar{\rho}(x_j) + \rho'(v_i)(u - x_j)) du = \\ &= \sum_{j=1}^{n_i} \frac{\Delta_j^{s+1}}{(s+1)2^s} \bar{\rho}(x_j) = \frac{1}{(s+1)2^s} G_i^{s+1} n_i^{-(s+1)} \sum_{j=1}^{n_i} \bar{\rho}(x_j) = \\ &= \frac{1}{(s+1)2^s} G_i^{s+1} n_i^{-s} \rho(v_i). \end{aligned}$$

Таким образом, подставляя полученные результаты в (6), получаем:

$$\varphi(x) \leq \frac{1}{(s+1)2^s} \sum_{i=1}^{\kappa} G_i^{s+1} n_i^{-s} \rho(v_i) + \sum_{i=1}^{\kappa} O\left(\frac{G_i^{s+3}}{n_i^s}\right). \quad (8)$$

Оценим второе слагаемое этого неравенства, используя (5):

$$\begin{aligned} &\sum_{i=1}^{\kappa} O\left(\frac{G_i^{s+3}}{n_i^s}\right) \leq \\ &\leq \left(\max_{1 \leq i \leq \kappa} G_i\right)^2 n^{-s} \sum_{i=1}^{\kappa} G_i^{s+1} \left(\int_{G_i} \rho^{1/(s+1)}(u) du\right)^{-s} \left(\int_G \rho^{1/(s+1)}(u) du\right)^s. \end{aligned}$$

Из того, что $\rho(u) \geq m$ на всем отрезке G , следует:

$$\left(\int_{G_i} \rho^{1/(s+1)}(u) du \right)^{-s} \leq \frac{1}{G_i^s m^{s/(s+1)}}.$$

В силу полученного и (4), верхняя оценка для правой части (8) равна

$$\frac{G^2}{n^{2a} n^s} \left(\int_G \rho^{1/(s+1)}(u) du \right)^s \sum_{i=1}^{\kappa} \frac{G_i}{m^{s/(s+1)}} = \frac{G^2}{n^{2a} n^s} \left(\int_G \rho^{1/2}(u) du \right)^s \frac{G}{m^{s/(s+1)}}.$$

Поскольку в последнем выражении G , m и интеграл являются константами, а параметр a удовлетворяет условию $a > 0.5$, получаем:

$$\sum_{i=1}^{\kappa} O\left(\frac{G_i^{s+3}}{n_i^s}\right) = o\left(\frac{1}{n^{s+1}}\right).$$

Далее оценим первое слагаемое в неравенстве (7), снова используя (5):

$$\begin{aligned} & \frac{1}{(s+1)2^s} \sum_{i=1}^{\kappa} G_i^{s+1} n_i^{-s} \rho(v_i) = \\ & = \frac{1}{(s+1)2^s} \left(\int_G \rho^{1/(s+1)}(u) du \right)^s \sum_{i=1}^{\kappa} \frac{G_i^{s+1} \rho(v_i)}{n^s \left(\int_{G_i} \rho^{1/(s+1)}(u) du \right)^s} = \\ & = \frac{1}{(s+1)(2n)^s} \left(\int_G \rho^{1/(s+1)}(u) du \right)^s \sum_{i=1}^{\kappa} \rho^{1/(s+1)}(v_i) G_i \left(\frac{\rho^{1/(s+1)}(v_i) G_i}{\int_{G_i} \rho^{1/(s+1)}(u) du} \right)^s. \end{aligned}$$

Для знаменателя последней скобки по формуле численного интегрирования методом прямоугольников имеем:

$$\int_{G_i} \rho^{1/(s+1)}(v) dv = G_i \rho^{1/(s+1)}(v_i) + O(G_i^3).$$

Подставляя полученную сумму в знаменатель и выполняя сокращение на числитель, получаем дробь $\frac{1}{1 + O(G_i^2)}$, которая при $G_i \rightarrow 0$ ведет себя эквивалентно $1 - O(G_i^2)$. Таким образом,

$$\frac{1}{(s+1)2^s} \sum_{i=1}^{\kappa} G_i^{s+1} n_i^{-s} \rho(v_i) =$$

$$= \frac{1}{(s+1)(2n)^s} \left(\int_G \rho^{1/(s+1)}(u) du \right)^s \sum_{i=1}^{\kappa} \rho^{1/(s+1)}(v_i) G_i (1 + O(G_i^2))^s.$$

Из (4) известно, что $G_i^2 \leq \frac{G}{n^{2a}} = o\left(\frac{1}{n}\right)$, поскольку $a > 0.5$, поэтому верно:

$$(1 + O(G_i^2))^s = 1 + s O(G_i^2) + \dots = 1 + o\left(\frac{1}{n}\right).$$

Далее, используя формулу подсчета интегралов методом прямоугольников, получим

$$\sum_{i=1}^{\kappa} \rho^{1/(s+1)}(v_i) G_i = \int_G \rho^{1/(s+1)}(v) dv + O\left(\sum_{i=1}^{\kappa} G_i^3\right) = \int_G \rho^{1/(s+1)}(v) dv + o\left(\frac{1}{n}\right),$$

поскольку, учитывая условие $3a - b > 1$,

$$\sum_{i=1}^{\kappa} G_i^3 \leq \sum_{i=1}^{\kappa} \frac{G^3}{n^{3a}} = \frac{G^3}{n^{3a-b}} = o\left(\frac{1}{n}\right).$$

Следовательно, верхняя оценка критерия $\varphi(x)$ имеет следующий вид:

$$\begin{aligned} \varphi(x) &\leq \frac{1}{(s+1)(2n)^s} \left(\int_G \rho^{1/(s+1)}(u) du \right)^{s+1} + o\left(\frac{1}{n^{s+1}}\right) = \\ &= \frac{1}{(s+1)(2n)^s} \|\rho\| + o\left(\frac{1}{n^{s+1}}\right). \end{aligned}$$

Сопоставляя полученный результат с результатом теоремы (2), получим утверждение теоремы.

Заключение

В данной работе представлен алгоритм размещения станций обслуживания на отрезке, а также доказана асимптотическая второго порядка оптимальность этого алгоритма. Полученные результаты являются обобщением известных ранее аналогичных результатов для критерия с параметром $s \in \mathbb{N}$. Это обобщение нацелено на повышение гибкости подхода к решению рассмотренной в работе задачи. В частности, появилась возможность строить описанные размещения при $s < 1$, то есть эффективно обрабатывать области с высокой концентрацией вызовов.

Библиографический список

1. Назаров Л. В., Смирнов С. Н. Обслуживание вызовов, распределенных в пространстве // Изв. АН СССР. Техн. кибернет. – 1982. – № 1. – С. 95–99.
2. Захарова Т. В. Оптимизация расположения станций обслуживания в пространстве // Информатика и ее применения. – 2008. – Т. 2. – № 2. – С. 41–46.
3. Захарова Т. В. Оптимизация расположения в пространстве систем массового обслуживания: дис. ... канд. физ.-мат. наук. – М., МГУ, 2008.
4. Матвеева С. С., Захарова Т. В. Сети массового обслуживания с наименьшей длиной очереди // Информатика и ее применения. – 2010. – Т. 4. – № 3. – С. 22–28.
5. Захарова Т. В., Фисак А. А. Необходимое условие оптимального размещения станций обслуживания на отрезке для параметрического семейства критериев оптимальности // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. / Перм. гос. нац. иссл. ун-т. – Пермь, 2015. – Вып. 26. – С. 86–94.
6. Захарова Т. В., Фисак А. А. Об оптимальных размещениях станций обслуживания // Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика. – 2018. – № 2. – С. 40–47.

Asymptotically optimal arrangement of service stations on straight line

T.V. Zakharova¹, I.M. Menshikh²

¹Faculty of Computational Mathematics and Cybernetics of Moscow State University; 119991, Moscow, GSP-1, Leninskie Gory, Moscow State University, Faculty of Computational Mathematics and Cybernetics; Institute for Informatics Problems of FRC IC RAS; e-mail: lsa@cs.msu.ru

²Faculty of Computational Mathematics and Cybernetics of Moscow State University

Abstract. *The article deals with the neoclassical problem of queueing theory. Calls emerge on straight line and there are n stations to service them. The call goes to the nearest station. Call service time is some increasing function depending on the distance between the call and the station. It is necessary to arrange the stations in such a way to minimize the service time of incoming calls. It is possible to find optimal station arrangement for tasks of this kind only in exceptional situations. However, it is possible to obtain satisfactory for practice solution of the problem. The article presents an algorithm for constructing asymptotically optimal second-order arrangements for a family of optimality criteria with a real positive parameter.*

Key words: *optimal arrangement; service station; zone of influence; optimality criterion.*

УДК 519.2

К методике кластеризации крупных городских школ России. Факторы, отвечающие за результаты образовательной деятельности

А. А. Макаров¹, Г. И. Симонова², Д. И. Зверева³

¹Национальный исследовательский университет Высшая школа экономики, г. Москва, Россия; amakarov@hse.ru;

²Национальный исследовательский университет Высшая школа экономики, г. Москва, Россия; gisimon@yandex.ru;

³Центр педагогического мастерства, г. Москва, Россия; daryazvereva@gmail.com

Аннотация. В статье изучены возможные подходы к кластеризации крупных городских школ РФ по результатам их образовательной деятельности. Исследовано, насколько попадание школы в тот или иной кластер обусловлено рядом объективных условий, в которых функционируют школы. Выявлены значимые показатели условий, влияющие на результаты ЕГЭ в школах. Показано, что исследуемых показателей условий работы школ недостаточно для правильной кластеризации школ по агрегированным показателям ЕГЭ.

Ключевые слова: классификация, кластерный анализ, многомерный дисперсионный анализ, оценки параметров, дискриминантный анализ.

1. Введение

К оценке работы государственных общеобразовательных школ России можно подходить с разных позиций и использовать при этом разные критерии и показатели. Это актуальная задача не только для органов образования, но и для всех, кто так или иначе вовлечён в образование, включая учащихся, их родителей, работодателей, университеты, органы государственной власти. Однако, при решении задачи оценки образовательной деятельности и выработке рекомендаций и государственной политики в этой области необходимо учитывать значительную неоднородность государственных школ страны. В России по данным на начало 2017 года насчитывалось чуть более 40 тысяч государственных общеобразовательных школ. Обследование этих школ в начале 2017 г. в рамках программы Национального исследования качества образования (НИКО) [1]

показало, что эти школы распадаются по объективным характеристикам на несколько укрупнённых групп. В табл. 1 приведено распределение городских и сельских школ в стране и обучающихся в них.

Таблица 1. Государственные городские и сельские школы РФ и обучающиеся в них

Тип школы	Число школ	% школ	Обучающихся	% обучающихся
Городские	14957	37%	10564297	72%
Сельские	25689	63%	4103541	28%
Всего	40646	100%	14667838	100%

Из таблицы 1 видно, что хотя процент городских школ в стране менее 40%, но обучается в них 72% всех учащихся. Однако, и городские школы далеко неоднородны. Среди них есть заметная доля школ, которые не имеют старшей школы, а также есть школы, в которых число обучающихся в 10-11 классах малочисленно. По этим школам трудно или невозможно судить об итоговом уровне подготовки учащихся, так как результаты экзаменов в 9-м классе (ОГЭ) не сопоставимы в разных регионах страны, а результаты учащихся по ЕГЭ могут быть неустойчивыми и не вполне объективно отражают положение дел. В настоящей работе будут рассмотрены 9544 городские школы страны, в которых, как правило, есть не менее двух выпускных девятых классов и достаточное число учащихся в 10-11-м классах. В подобных школах обучается 55% всех учащихся страны. В работе будут рассмотрены агрегированные показатели ЕГЭ по русскому языку и профильной математике за 2016 год в указанных школах, исследованы различные возможные подходы к кластеризации школ по их агрегированным показателям и изучено возможное влияние ряда объективных условий на уровень подготовки учащихся по русскому языку и профильной математике.

2. Данные исследования

С учётом заметной разнородности школ, в настоящем исследовании мы рассмотрим только большие городские государственные школы страны, имеющие выпуск в 11-м классе. Такой выбор обусловлен рядом соображений. Во-первых, мы изначально хотим заметно уменьшить объективную неоднородность условий, в которых находятся школы. Во-вторых, относительно надёжными объективными для всех школ показателями образовательной деятельности являются лишь результаты ЕГЭ. (Заметим, что ОГЭ-экзамены, которые сдают учащиеся в 9-м классе, находятся в ведении

субъектов федерации, и каждый регион России имеет право формировать собственные критерии оценивания учащихся. То есть результаты ОГЭ по школам разных регионов просто несопоставимы между собой.) В-третьих, судить об уровне образования в школе по малому количеству выпускников старших классов нам не представляется возможным, так как такие показатели подвержены высокой доле случайной изменчивости.

Исходя из этих принципов, для исследования были отобраны городские не малокомплектные школы, в которых есть хотя бы один 11-й класс и число выпускников 9-го класса не меньше 33, и по которым есть данные по ЕГЭ по русскому языку и математике. В подавляющем большинстве этих школ не менее двух девярых классов. Согласно многолетним статистическим наблюдениям, в среднем примерно половина выпускников девятого класса продолжают обучение в старшей школе. При этом в специализированных школах эта доля заметно выше.

В ходе обследования всех государственных школ страны по проекту НИКО в 2017 году была собрана информация, охватывающая многие стороны функционирования школы. В частности, фиксировались:

- численность населения города, в котором находится школа;
- процент обучающихся с неродным русским языком;
- количество обучающихся в каждой параллели классов;
- число учителей в школе;
- число учителей-предметников;
- есть ли в названии школ слова «лицей», «гимназия», «школа с углубленным изучением каких-либо предметов»;
- другие показатели условий работы школы, включая транспортную доступность населённого пункта и многое другое.

Однако считать собранную информацию исчерпывающей нельзя. Скажем, мы не видим по собранным данным, находится ли школа в неблагополучном районе города или нет. Последнее обстоятельство может очень существенно влиять на показатели её образовательной деятельности даже в крупных городах страны. По части указанных значимых средних показателей уровня подготовки учащихся в дальнейшем проводилась агрегация. О ней более подробно сказано в разделе 4.

3. Подходы к кластеризации школ по итогам образовательной деятельности

Практически единственной сопоставимой информацией, если говорить о результатах образовательной деятельности по всем рассматриваемым школам, являются результаты обязательных экзаменов ЕГЭ по русскому языку и математике. В полном виде эта информация описывается распределением баллов по каждому из этих предметов. В работе [2] описана задача кластеризации регионов РФ по распределениям баллов ЕГЭ. Там, в частности, отмечалось, что обращение ко всему не агрегированному распределению является явно избыточным, так как содержательно нет принципиального различия между близкими баллами ЕГЭ, и предлагалось рассматривать агрегированные распределения. С другой стороны, в министерстве образования РФ и Рособрнадзоре часто для сравнительного анализа регионов предлагают опираться только на средние баллы региона по предмету.

В настоящей работе будет дан сравнительный анализ результатов двух подходов к кластеризации. В первом подходе каждая школа будет описываться только своими средними баллами ЕГЭ по русскому языку и математике, а во втором будут использованы такие обобщённые, высоко устойчивые к случайным колебаниям характеристики распределений баллов по каждому из предметов, как нижняя квартиль, медиана и верхняя квартиль. В обоих случаях для кластеризации школ будет использован метод k -средних. Кластеризация проводилась в пакете SPSS [3].

3.1. Кластеризация школ по средним баллам ЕГЭ по русскому языку и профильной математике

При кластеризации школ по показателям средних баллов ЕГЭ за 2016 год по русскому языку и профильной математике было выделено шесть кластеров. Центры этих кластеров представлены в табл. 2.

Таблица 2. Центры кластеров

Средний балл ЕГЭ	Кластер					
	1	2	3	4	5	6
Русский язык	79,53	74,53	72,91	65,57	64,81	54,57
Математика (профильная)	63,33	53,38	43,55	47,63	36,64	28,43

Номера кластеров проставлены в порядке убывания суммы баллов по русскому языку и профильной математике для центра кластера.

В табл. 3 представлено распределение школ и обучающихся в них по указанным кластерам.

Таблица 3. Число школ и число обучающихся в них школьников в каждом кластере и их проценты

	Кластер						Всего
	1	2	3	4	5	6	
Число школ	1043	2510	2007	1623	1837	524	9544
% школ	10,93%	26,30%	21,03%	17,01%	19,25%	5,49%	100%
Число учащихся	948158	2317794	1727017	1326384	1376733	372648	8068734
% школьников	11,75%	28,73%	21,40%	16,44%	17,06%	4,62%	100%

Графически разброс средних показателей школ в каждом из кластеров можно представить переходя к сумме средних баллов ЕГЭ по русскому языку и профильной математике (см. рис. 1). Именно суммы баллов выпускников по предметам учитываются при приеме в вузы. При этом по большинству естественно-научных, инженерно-технических и социально-управленческих специальностей суммы баллов ЕГЭ по русскому языку и профильной математике являются составной частью конкурсного балла.

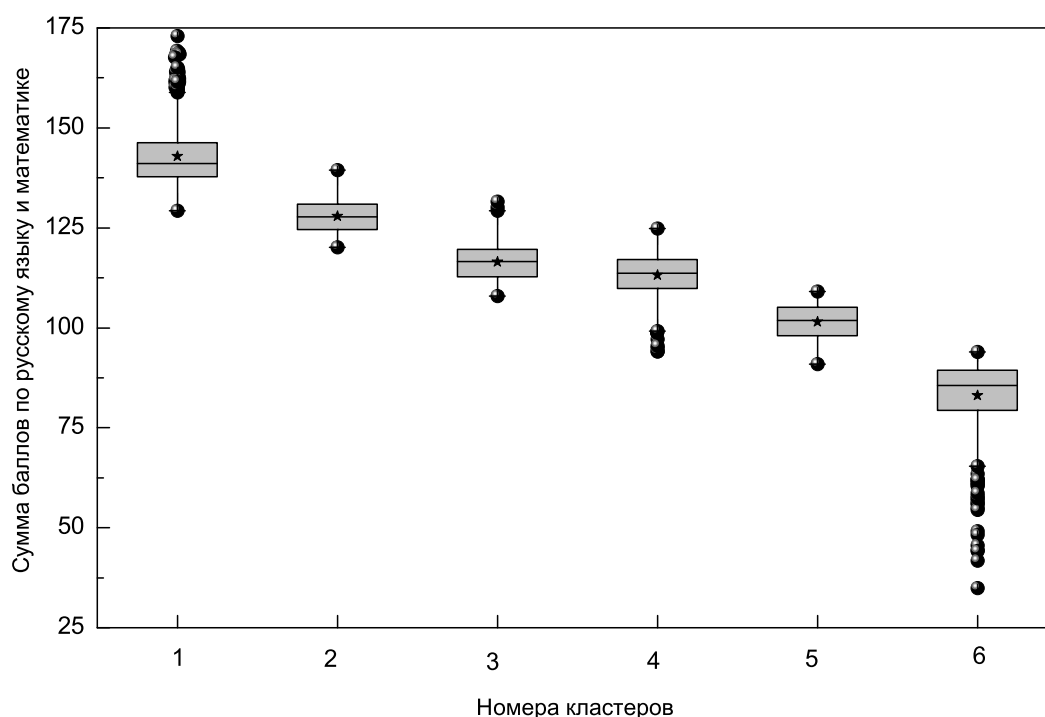


Рис. 1. Ящики с усами для суммарного балла ЕГЭ по русскому языку и профильной математике при классификации по средним баллам

Из рис. 1 видно, что разброс сумм баллов внутри кластеров 2–5 невелик и только в первом и шестом кластере есть школы заметно отличающиеся от остальных в своём кластере. На наш взгляд, учитывая, что межквартильный размах средних в этих кластерах не велик (то есть основная

масса школ в этих кластерах достаточно однородна), нет необходимости вводить дополнительные кластеры для этих выделяющихся школ, так как число объектов в этих дополнительных кластерах будет мало. Также обратим внимание на то, что хотя с точки зрения суммы средних баллов ЕГЭ кластеры 3 и 4 кажутся практически одинаковыми, однако, из таблицы 2 хорошо видно, что центры этих кластеров заметно отличаются.

3.2. Кластеризация школ по базовым характеристикам распределений баллов ЕГЭ по русскому языку и профильной математике

В качестве обобщённых базовых характеристик распределений баллов ЕГЭ по каждому из предметов мы рассматриваем три показателя: нижнюю и верхнюю квартили распределений и их медиану. Таким образом, каждая школа получает шесть параметров, которые будут учитываться в этой кластеризации. Для удобства сравнения двух подходов к выбору характеристик объектов приведём ниже результаты, полученные при выделении шести кластеров. Центры этих кластеров представлены в табл. 4.

Таблица 4. Центры кластеров при кластеризации по базовым характеристикам распределения баллов ЕГЭ по русскому языку и профильной математике

Параметры центра кластера	Кластер					
	1	2	3	4	5	6
Русский язык (нижняя квартиль)	72	66	65	57	58	47
Русский язык (медиана)	80	74	73	65	66	56
Русский язык (верхняя квартиль)	88	84	83	73	74	65
Математика (нижняя квартиль)	53	41	31	35	26	20
Математика (медиана)	67	55	43	47	36	28
Математика (верхняя квартиль)	75	68	58	61	47	38

Распределение школ по кластерам при подобной методике классификации указано в табл. 5, где кластеры так же как и при кластеризации по средним значениям упорядочены по уровню подготовки учащихся по указанным предметам.

Таблица 5. Число наблюдений в каждом кластере

	Кластер						Итого
	1	2	3	4	5	6	
Число наблюдений	1045	2395	1853	1671	1881	699	9544

Сравнение данных табл. 5 и 3 показывает, что по отдельным кластерам (1, 4, 5) числа объектов в них практически совпадают. Это, конечно, еще не означает, что в эти кластеры попали одни и те же школы. В кластерах 2, 3, 6 отличие составляет от 100 до 200 школ.

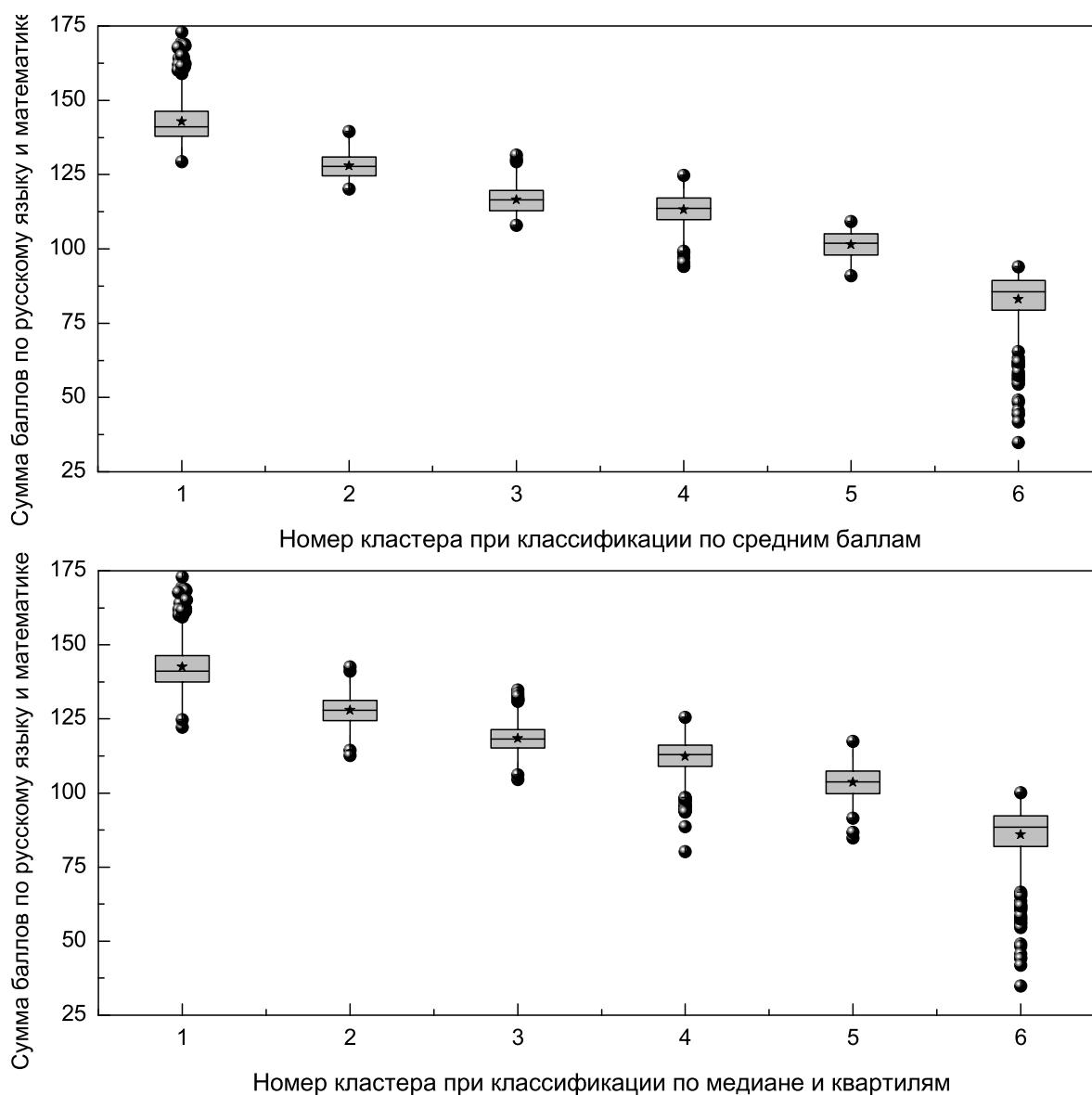


Рис. 2. Ящики с усами для суммарного балла по русскому языку и математике для кластеров, полученных при классификации по средним значениям баллов по русскому языку и математике, и при классификации по медиане и квартилям баллов по русскому языку и математике.

На рис. 2 изображена диаграмма «ящик с усами» для суммарного балла по русскому языку и математике для кластеров, полученных вторым методом (классификация по медиане и квартилям баллов по русскому языку и математике). Для сравнения здесь же изображена такая же диаграмма, полученная первым методом (классификация по средним баллам по русскому языку и математике). Эти диаграммы принципиально не отличаются друг от друга.

Сравнение схожих по смыслу характеристик центров кластеров при кластеризации по средним значениям (таблица 2) и по показателям распределений показывает, что медианный балл по русскому языку в кластерах практически не отличается от среднего балла по русскому языку. Медианы же по профильной математике в первом и втором кластере значительно отличаются от средних в аналогичных кластерах. В остальных кластерах эти различия не существенны.

Очевидно, что более подробная информация о распределении дает дополнительную информацию о дифференциации показателей учеников школы и более детально и содержательно характеризует каждый из кластеров. Однако, если отвлечься от детализации, желательно выяснить, удалось ли получить принципиально иную, более содержательную классификацию школ, переходя к более подробной информации о показателях её выпускников? Для ответа на этот вопрос проведён сравнительный анализ этих двух методик кластеризации.

3.3. Сравнительный анализ двух подходов к кластеризации школ

Выше были сделаны некоторые общие замечания о схожести результатов двух разных кластеризаций, включая схожесть кластерных центров и близость численностей объектов в соответствующих кластерах. Более детальную информацию о соотношении этих кластеризаций дает анализ таблицы сопряжённости двух классификаций, представленной в табл. 6. Составим таблицу сопряжённости и с помощью критерия хи-квадрат проверим гипотезу о независимости двух номинальных переменных – номеров кластеров при кластеризация по средним значениям и номеров кластеров при кластеризация по медиане и квартилям.

Таблица 6. Таблица сопряжённости для двух классификаций

Кластер по средним баллам	Кластер по медиане и квартилям						Всего	
	1	2	3	4	5	6		
Частота	1	958	85	0	0	0	0	1043
% по строке		91,9	8,1	0,0	0,0	0,0	0,0	100
% по столбцу		91,7	3,5	0,0	0,0	0,0	0,0	10,9
	2	85	2133	249	43	0	0	2510
		3,4	85,0	9,9	1,7	0,0	0,0	100
		8,1	89,1	13,4	2,6	0,0	0,0	26,3
	3	0	59	1529	116	303	0	2007
		0,0	2,9	76,2	5,8	15,1	0,0	100
		0,0	2,5	82,5	6,9	16,1	0,0	21,0
	4	2	118	55	1377	70	1	1623
		0,1	7,3	3,4	84,8	4,3	0,1	100
		0,2	4,9	3,0	82,4	3,7	0,1	17,0
	5	0	0	20	131	1502	184	1837
		0,0	0,0	1,1	7,1	81,8	10,0	100
		0,0	0,0	1,1	7,8	79,9	26,3	19,2
	6	0	0	0	4	6	514	524
		0,0	0,0	0,0	0,8	1,1	98,1	100
		0,0	0,0	0,0	0,2	0,3	73,5	5,5
Итого		1045	2395	1853	1671	1881	699	9544
		10,9	25,1	19,4	17,5	19,7	7,3	100
		100	100	100	100	100	100	100

На диагонали таблицы жирным шрифтом указаны проценты совпадений при двух классификациях (т.е. отнесение к кластерам с одинаковыми номерами) по строкам и столбцам. По строкам – это процент наблюдений среди относящихся к соответствующему кластеру первой классификации, которые попадают в соответствующий кластер по второй классификации. По столбцам – это процент наблюдений среди относящихся к соответствующему кластеру второй классификации, которые попадают в соответствующий кластер первой классификации. Можно считать, что обе классификации достаточно хорошо совпадают. При этом общий процент отнесения наблюдений к одному и тому же кластеру при двух классификациях составляет 84%. Значимый коэффициент сопряжённости Крамера-Чупрова, используемый в качестве меры связи между двумя признаками, равен $V = \sqrt{\frac{\chi^2}{n\sqrt{(r-1)(s-1)}}} = 0,823$. Здесь $r = 6, s = 6$ – числа строк и

столбцов в таблице, χ^2 – значение статистики, вычисленной по таблице сопряжённости, которая при независимости признаков имеет асимптотическое распределение хи-квадрат с $f = (r-1)(s-1) = 25$, степенями свободы, $n = 9544$ – общее число наблюдений. Критерий хи-квадрат и коэффициент сопряжённости показывают сильную зависимость двух классификаций.

Схожесть результатов этих кластеризаций можно проиллюстрировать несколькими примерами. Если рассмотреть самый сильный первый кластер, то почти 92% школ из этого кластера по одной классификации попадают в соответствующий первый кластер по другой кластеризации. Лишь около 8% школ из этого кластера переходит в чуть более слабый второй кластер и обратно. Школы, образующие 4-й кластер, совпадают на 82-84%. А самые заметные изменения происходят в шестом кластере. При классификации по медиане и квартилям этот кластер фактически формируется из всех школ шестого кластера по средним и к ним добавляется еще около 170 школ из пятого кластера. То есть классификация по медианам и квартилям дает заметные отличия прежде всего на небольшом количестве школ с самыми слабыми результатами. Общий процент таких школ составляет от 5 до 7% при разных классификациях.

Также можно сформировать таблицу для средних значений суммарного балла по двум предметам в каждом кластере каждой из двух классификаций. Эти данные представлены в табл. 7.

Таблица 7. Средние значения суммарного балла по двум предметам в каждом кластере каждой из двух классификаций

Номер кластера	Число наблюдений для первой классификации	Среднее значение суммы баллов для первой классификации	Среднее значение суммы баллов для второй классификации	Число наблюдений для второй классификации
1	1043	142,86	142,58	1045
2	2510	127,91	127,86	2395
3	2007	116,45	118,42	1853
4	1623	113,19	112,33	1671
5	1837	101,45	103,62	1881
6	524	83,00	85,94	699

Табл. 7 показывает, что с точки зрения сумм баллов по двум предметам, выделенные кластеры заметно отличаются друг от друга, но эти суммы остаются практически неизменными при разных методиках классификации. Можно рассмотреть более детальные и точные классификации,

при которых количества выбранных кластеров при разных методиках будут не совпадать, например, 6 и 8 кластеров соответственно. В этом случае исследование зависимости между двумя номинальными переменными, характеризующими два метода классификации, также показывает значимую высокую зависимость между этими методами. Поэтому при дальнейшем исследовании возможных условий, которые определяют попадание школ в тот или иной кластер, мы остановимся на более простой классификации с 6-ю кластерами, основанной на двух переменных – средний балл по русскому языку и средний балл по математике.

4. Факторы, влияющие на результаты образовательной деятельности

Как видно из предыдущего раздела, крупные городские школы страны, в которых обучается более половины всех школьников страны, весьма неоднородны по агрегированным результатам образовательной деятельности по русскому языку и математике. В настоящем разделе исследовано возможное влияние тех или иных условий, в которых работает школа, на результаты образования. Среди доступных для исследования факторов объективных условий выделены те, которые оказывают значимое влияние на результаты образовательной деятельности. При этом для значимых факторов выявлена их разумная агрегация. Все городские населённые пункты (переменная V_1), как показало исследование, целесообразно разбить на три группы: города с населением не более 250 тыс. человек ($V_1 = 1$), города с населением от 250 тыс. до 500 тыс. человек ($V_1 = 2$) и города с населением более 500 тыс. человек ($V_1 = 3$). Показатель «обучающиеся с неродным русским языком» (переменная V_2) тоже был разбит на три категории: в первую попали школы, где процент обучающихся с неродным русским языком не превышает 20% ($V_2 = 1$), во вторую – процент неродного русского языка составляет от 20% до 80% ($V_2 = 2$) и в третью – процент неродного русского языка среди обучающихся превышает 80% ($V_2 = 3$). По численности десятых классов (переменная V_3) выделены два типа школ: школы с одним 10-ым классом ($V_3 = 0$) и школы с более чем одним десятым классом ($V_3 = 1$). У последних школ есть больше объективных возможностей проводить более гибкую кадровую политику при наборе учителей-предметников в старших классах. Кроме того все школы были разбиты на две группы (переменная V_4) – общеобразовательные школы (ООШ) ($V_4 = 0$) и гимназии и лицеи ($V_4 = 1$). Последнее разбиение отчасти условно, так как далеко не всегда условия и уровень обучения в гимназиях и лицеях соответствует требованиям, которые раньше предъ-

являлись к этому типу образовательных учреждений. В настоящее время официально упразднено разбиение школ на ООШ, гимназии и лицеи, однако за бывшими названиями гимназий и лицеев на практике обычно стоят более сильные и квалифицированные педагогические коллективы, что не может не сказываться на результатах обучения. Для каждого из перечисленных выше показателей были исследованы таблицы сопряжённости между предложенными в разделе 3 шестью кластерами и соответствующими уровнями факторов. Все таблицы сопряжённости для этих переменных сведены в одну табл. 8.

Таблица 8. Таблицы сопряжённости для переменных $V_1 - V_4$, для которых связь с переменной классификации `ncluster6` подтверждается

Переменные	Кластер						Всего	
	1	2	3	4	5	6		
V_1	1	428	1295	1133	968	1056	319	5199
	2	198	423	347	216	307	89	1580
	3	417	792	527	439	474	116	2765
Всего		1043	2510	2007	1623	1837	524	9544
V_2	1	953	2307	1897	1407	1644	364	8572
	2	75	156	91	149	140	51	662
	3	15	47	19	67	53	109	310
Всего		1043	2510	2007	1623	1837	524	9544
V_3	0	273	958	1037	881	1220	346	4715
	1	770	1552	970	742	617	178	4829
Всего		1043	2510	2007	1623	1837	524	9544
V_4	0	282	1439	1478	1313	1632	493	6637
	1	761	1071	529	310	205	31	2907
Всего		1043	2510	2007	1623	1837	524	9544

В табл. 9 приведены значения коэффициентов сопряжённости Крамера-Чупрова между кластерной переменной `ncluster6` и переменными $V_1 - V_4$.

Таблица 9. Коэффициенты сопряжённости Крамера-Чупрова между переменной `ncluster6` и переменными $V_1 - V_4$

	V_1	V_2	V_3	V_4
Коэффициент Крамера-Чупрова	0,086	0,182	0,259	0,416
Значимость	0,000	0,000	0,000	0,000

Из таблицы 9 видно, что наиболее сильная значимая связь наблюдается для признака тип школы (ООШ, гимназия или лицей). Другими словами вклад сильного педагогического коллектива в результаты образования

существеннее, чем все остальные факторы. Отметим, что фактор размера школы (переменная V_3) также оказывает положительное влияние на результаты образовательной деятельности. Самая слабая (хотя и значимая) связь уровня подготовки учащихся наблюдается с размером населённого пункта. Выявленные значимые факторы можно объединить в обобщённую линейную модель, чтобы понять насколько в совокупности эти факторы определяют средние баллы выпускников школы по русскому языку и профильной математике.

5. Обобщённая линейная модель взаимосвязи средних показателей школ в ЕГЭ по русскому языку и профильной математике с условиями, в которых работают школы

Рассмотрим вклад выделенных качественных переменных $V_1 - V_4$ на зависимые переменные Y_p – средний балл по русскому языку и Y_m – средний балл по профильной математике. Применим многомерный дисперсионный анализ с использованием критериев Роя, Хотеллинга, Уилкса [4], [5]. Была исследована полная модель с включением качественных переменных и всевозможных взаимодействий между ними. Значимыми являются все эти переменные и такие взаимодействия между ними: $V_1 \times V_4$, $V_3 \times V_4$, $V_1 \times V_2$, $V_1 \times V_2 \times V_3$, $V_2 \times V_4$. Все эти переменные хотя и дают значимый вклад в зависимые переменные, однако коэффициенты детерминации при этом невелики и равны 24% и 19,8% для переменных Y_p и Y_m соответственно. И значит, только по этим переменным прогноз ЕГЭ-баллов затруднителен. Соответственно и классификация школ по результатам ЕГЭ с использованием только этих переменных будет в большинстве случаев ошибочна. Это подтверждается методом дискриминантного анализа, применение которого описано ниже.

Исследуем, можно ли расклассифицировать школы по выделенным в разделе 3 кластерам, опираясь на выделенные выше значимые показатели влияния и их взаимодействия.

Результаты дискриминантного анализа с указанными выше предикторами представлены в табл. 10.

Таблица 10. Результаты классификации

	Исходный кластер	Предсказанный кластер						Итого
		1	2	3	4	5	6	
Исходные: частота %	1	422 40,5	326 31,3	109 10,5	14 1,3	149 14,3	23 2,2	1043 100
	2	464 18,5	591 23,5	679 27,1	26 1,0	659 26,3	91 3,6	2510 100
	3	168 8,4	357 17,8	598 29,8	12 0,6	811 40,4	61 3,0	2007 100
	4	68 4,2	223 13,7	476 29,3	34 2,1	692 42,6	130 8,0	1623 100
	5	35 1,9	165 9,0	442 24,1	19 1,0	1025 55,8	151 8,2	1837 100
	6	5 1,0	19 3,6	91 17,4	16 3,1	250 47,7	143 27,3	524 100

Лишь 29,5% исходных сгруппированных наблюдений классифицировано правильно, что является неудовлетворительным.

6. Заключение

В работе из всего многообразия российских государственных школ выделена достаточно однородная по формальным условиям группа крупных городских школ, в которых обучается более половины всех учащихся страны. Для этих школ исследованы показатели их образовательной деятельности по подготовке школьников по русскому языку и профильной математике на базе ЕГЭ 2016 года. В разделе 3 показано, что подобные школы довольно неоднородны по своим агрегированным результатам. Сравнительный анализ различных подходов к кластеризации школ показал, что для обобщённой классификации этих школ достаточно средних показателей ЕГЭ по русскому языку и профильной математике и нет необходимости прибегать к дополнительным характеристикам распределений баллов ЕГЭ учащихся по школам. С другой стороны, такие характеристики как верхняя и нижняя квартили распределений баллов ЕГЭ по школам помогают лучше понять реальный уровень подготовки учащихся. Так, в самом слабом шестом кластере, в котором около 700 школ (см. таблицу 4), верхняя квартиль распределения ЕГЭ по профильной математике находится

на уровне 38 баллов, то есть 75% выпускников этих школ не могут продолжать обучение в вузах по естественно-научным, инженерным, социально-экономическим и управленческим специальностям. Одновременно с этим в самом сильном кластере школ, охватывающим около 1000 крупных городских школ, показатели половины выпускников по русскому языку и математике вполне соответствуют требованиям подготовки к учебе в вузах. Настоящая статья не дает более или менее полного ответа на вопрос, почему наблюдается такой большой разброс в результатах образовательной деятельности крупных городских школ. При этом нами исследовано возможное влияние на эти результаты ряда объективных факторов, в которых работают школы. В статье рассмотрены следующие факторы: размер населённого пункта, количество выпускных классов в школе, тип школы (ООШ, гимназия или лицей), доля учащихся с неродным русским языком. Выяснено, что каждый из перечисленных факторов и их взаимодействия значимы для достижения образовательных результатов, однако предопределяют эти результаты только в среднем на 20%. Это свидетельствует о том, что в любых, даже самых неблагоприятных условиях у изученных школ есть все возможности показать высокие образовательные результаты. Как показывают построенные нами в разделе 5 обобщённые линейные модели взаимосвязи результатов и условий, высоким образовательным результатам благоприятствуют (в порядке убывания силы взаимосвязи):

- сильные образовательные коллективы, сформировавшиеся в бывших гимназиях и лицеях. Эти школы (см. табл. 5) составляют около 75% от всех школ самого сильного первого кластера и около 40% второго и третьего кластеров. При этом их процент в самых слабых пятом и шестом кластерах колеблется от 6% до 11%.
- Фактор числа выпускных классов (один или более). В первом и втором кластере процент этих школ колеблется от 60% до 70% и снижается до 36% в шестом кластере.
- Высокий процент учащихся в школе с родным русским языком. Так, около 94% школ первого кластера – это школы, где процент учащихся с родным русским языком составляет 80% и более. Однако, как показывают результаты анализа (см. табл. 5), в самом слабом шестом кластере процент школ с высоким процентом родного русского языка составляет более 70%. То есть слабые показатели крупных городских школ более характерны не для национальных республик, а для вполне русскоговорящих регионов страны.
- Размер населённого пункта оказывает самое слабое и, можно сказать,

не вполне однозначное влияние на образовательные результаты. Так, больше всего школ из самого сильного кластера находится именно в относительно мелких городах (до 250 тыс. человек), а вовсе не в самых крупных (более 500 тыс. человек).

Табл. 5 показывает, что городская школа с любым набором более благоприятных или менее благоприятных внешних условий может попасть в любой из выявленных кластеров. Однако особое внимание должны привлекать к себе школы, попавшие в самые слабые кластеры. В этих школах обучается более 1 миллиона 750 тысяч учащихся, что составляет более 10% всех учащихся школ страны. Ясно, что сбор информации о школах в рамках проекта НИКО не позволяет выявить ряд существенных условий функционирования государственных школ. К таким условиям, в частности, можно отнести неблагоприятные городские районы даже в крупных городах страны, слабые педагогические коллективы и многое другое. Поэтому в моделях, связывающих условия работы школы с её агрегированными образовательными результатами, не стоит ждать высоких коэффициентов детерминации. Однако, на наш взгляд, любые исследования в этой области, с одной стороны, помогают понять, какие компоненты образовательной среды способствуют более высоким образовательным результатам, а с другой – выявить те образовательные организации, в которых при вполне благоприятных условиях образовательные результаты неудовлетворительны.

Библиографический список

1. <https://niko.statgrad.org>
2. Макаров А.А., Симонова Г.И. Сравнительный анализ результатов кластеризации распределений баллов ЕГЭ по математике в регионах РФ в 2009-2010 гг. // Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. / Перм.гос.ун-т. – Пермь. – 2011. – Вып. 23. – С. 89-103.
3. Бьюль А., Цёфель П. SPSS: искусство обработки информации. – М., СПб, Киев: DiaSoft, 2005. – 608 с.
4. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физматлит, 1963. – 500 с.
5. Тюрин Ю.Н. Многомерная статистика: гауссовские линейные модели. – М.: Издательство Московского университета, 2011. – 136 с.

**To the methodology of clustering of large city schools in Russia.
Factors responsible for the results of educational activities.**

A.A. Makarov¹, G.I. Simonova², D.I. Zvereva³

¹ National Research University Higher School Economic; Moscow, Russia; amakarov@hse.ru;

² National Research University Higher School Economic; Moscow, Russia; gisimon@yandex.ru;

³ Center of Pedagogical Skills, Moscow, Russia; daryazvereva@gmail.com

Abstract. *In article possible approaches to clustering of large city schools of the Russian Federation by results of their educational activity are studied. The extent to which a school has entered a particular cluster is determined by a number of objective conditions in which schools operate. Significant indicators of conditions affecting the EGE-results in schools were identified. It is shown that the studied indexes of the working conditions of schools are not sufficient for the correct clustering of schools according to the aggregated EGE-indicators.*

Key words: *Classification, cluster analysis, multivariate ANOVA, estimation of parameters, discriminant analysis*

УДК 519.2

Уравнения для старших моментных характеристик решения линейного стохастического дифференциального уравнения в частных производных с аддитивными и мультипликативными возмущениями

И.Е. Полосков

Пермский государственный национальный исследовательский университет, г. Пермь, Россия;

polosk@psu.ru; +7342 2396560, 614990, г. Пермь, ул. Букирева, 15, ПГНИУ, кафедра высшей математики

Аннотация. Рассматривается задача построения дифференциальных уравнений для старших моментных функций решения линейного стохастического дифференциального уравнения в частных производных (СДУ в ЧП) параболического типа с аддитивными и мультипликативными возмущениями. Проблема решается на основе применения разработанной ранее схемы дискретизации общей системы СДУ в ЧП с последующим возвратом к непрерывной среде, которая позволяет решать ряд задач, связанных как с динамическим анализом поведения стохастических распределенных систем, так и получением стационарных характеристик, в частности, функционалов плотности вероятности.

Ключевые слова: стохастическое дифференциальное уравнение в частных производных, система с распределенными параметрами, случайное поле, моментные функции, параметрическое возмущение, уравнение Фоккера–Планка–Колмогорова.

1. Введение

Быстрый рост вычислительной мощности в течение последних десятилетий дает возможность моделировать и прогнозировать сложные физические процессы, происходящие в технических устройствах, природе, обществе и др. Между тем в математических моделях большинства физических процессов исследователи сталкиваются с неопределенностями из-за неполноты моделей, ошибок измерений и ограниченности знаний, связанных с входными данными. Случайные и стохастические дифференциальные уравнения в частных производных (СлДУ в ЧП, СДУ в ЧП) [1,2] дают

возможность учесть неточные данные и внешние и/или внутренние возмущения в математических моделях для последующего компьютерного анализа.

Примеры такого подхода широко известны. Например, большое число работ посвящено изучению напряженно-деформируемого состояния и вибраций пластин и оболочек [3], случайных колебаний упругих стержней [4], машиностроительных конструкций [5], элементов оборудования [6], транспортных систем [7]; изучению случайного деформирования и колебаний в упруго-пластических материалах [8, 9], стохастических режимов в лазерах [10] и химической технологии [11], движения волн (в том числе сейсмических) в случайно неоднородных средах [12, 13], изменения статистических характеристик полей в гидромеханике [14, 15]; исследованию процессов распространения звука [16], просачивания [17], в полимерах [18], композитах [19], молекулярных [20], микромеханических [21] и наномеханических [22] системах. Кроме того, существуют проблемы, связанные со случайными колебаниями зданий, в т.ч., при землетрясениях [23], с оптимальным приемом сигналов [24], устойчивостью вибраций [25], флуктуационной электродинамикой [13], горением [26] и др.

При исследовании моделей, включающих СДУ в ЧП, применяют как точные, так и приближенные схемы. Известно, что методы анализа стохастических аналогов классических линейных уравнений математической физики (параболических, гиперболических, колебаний балок, пластин и др.) [3, 27] основаны на разложениях решений уравнений в ряды по неслучайным собственным функциям соответствующих задач. При решении несложных нелинейных задач нередко с достаточной точностью можно воспользоваться собственными функциями линеаризованных уравнений.

При изучении линейных систем с распределенными параметрами, как правило, ограничиваются использованием методов корреляционного анализа [3, 4]. При анализе же нелинейных объектов одним из наиболее часто применяемых подходов является трактовка движений таких систем, как векторных марковских процессов [28]. Но для использования теории таких процессов требуется предварительная дискретизация задачи (обычно используются различные конечно-разностные схемы, методы Фурье, Галеркина с разложением по формам свободных упругих колебаний и редукция к конечному числу степеней свободы), что вносит в задачу необратимую погрешность [28]. Другим инструментом решения подобных задач является использование характеристических функционалов (ХФ) для случайных полей [12, 13]. Такая методика состоит в выводе уравнений в функциональных производных для этих ХФ [29], а затем, после представления ХФ рядами по

моментным функциям, в построении бесконечной цепочки зацепляющихся уравнений для указанных моментных функций.

К настоящему времени разработано значительное число приближенных методов исследования СДУ в ЧП. Заметим, что, как правило, такие методы являются комбинациями двух или более различных схем, а некоторые основаны на прямом переносе процедур анализа соответствующих детерминированных моделей, а также схем оценки реакции систем, описываемых стохастическими обыкновенными дифференциальными уравнениями (СОДУ), на область исследования процессов в случайной сплошной среде, в частности, это методы возмущений [30], полиномиального хаотического разложения [30], регулярного ортогонального разложения [31], стохастические конечно-разностные методы [32], методы конечных элементов [30], Галеркина [33], конечных объемов [34], коллокаций [35], моментов [36], граничных элементов [37] и др. Методы статистического моделирования (Монте–Карло, ММК) [38] являются самыми общими, наиболее распространенными и применимыми для численного анализа любых стохастических систем с распределенными параметрами. Реализации различных схем метода Монте–Карло концептуально просты и удобны в использовании. Но основным недостатком ММК остается необходимость выполнения огромного объема расчетов. Поэтому ММК по-прежнему используются, в основном, для проверки других подходов.

Перечисленные методы и схемы обладают определенными положительными характеристиками, но не лишены и слабых мест, основным из которых, по нашему мнению, является потеря информации об исходных случайных полях при различных формах дискретизации. Поэтому несложно заметить, что следствием необходимости уменьшения таких потерь, а также появления новых задач, требующих оперативного решения, является то, что в настоящее время постоянно возрастает интенсивность разработок современных вычислительных процедур, предназначенных для применения в рассматриваемой научной области.

В рамках наших исследований ранее была предложена схема дискретизации общей системы СДУ в ЧП, которая комбинируется с последующим возвратом к непрерывной среде [39]. Такая схема позволяет решать различные задачи, связанные как с динамическим анализом поведения стохастических распределенных систем, так и с получением стационарных характеристик (функционалов плотности вероятности). Ниже рассматривается применение данной схемы для получения дифференциальных уравнений для старших моментных функций решения линейного стохастического дифференциального уравнения в частных производных параболического

типа с аддитивными и мультипликативными возмущениями. Данная процедура лишена одного из недостатков упомянутых выше методов, а именно, результирующие дифференциальные уравнения являются ДУ В ЧП для первичных статистических характеристик неизвестного поля, а не для вторичных – параметров предварительной дискретизации этого поля.

2. Постановка задачи

Рассмотрим следующее СДУ в ЧП (в смысле Стратоновича [40]):

$$\frac{\partial U(x, t)}{\partial t} + \alpha^2 U(x, t) = \beta^2 \frac{\partial^2 U(x, t)}{\partial x^2} + [\gamma_0 + \gamma V(x, t)] U(x, t), \quad (1)$$

$$U(x, 0) = U^0(x), \quad (2)$$

где $t > 0$ – время; $x \in \mathbb{R} = (-\infty, +\infty)$; $U(x, t)$ – неизвестное случайное поле; $U^0(x)$ – случайная функция с известными характеристиками; $V(x, t)$ – пространственно-временной белый шум, вероятностные характеристики которого определяются следующими соотношениями:

$$E[V(x, t)] = 0, \quad E[V(x_1, t_1) V(x_2, t_2)] = \delta(x_2 - x_1) \delta(t_2 - t_1);$$

$\alpha > 0, \beta > 0, \gamma, \gamma_0$ – постоянные; $E[\dots]$ – символ математического ожидания.

Задача исследования состоит в построении детерминированных ДУ в ЧП для старших моментов случайного поля $U(x, t)$, удовлетворяющего уравнению (1) и начальному условию (2).

3. Дискретизация СДУ в ЧП

Дискретизируем уравнения (1) по пространственной координате x . Для этого введем сетку $x_\ell = \ell h, \ell = 0, \pm 1, \pm 2, \dots, h$ – шаг сетки ($0 < h \ll 1$), и заменим вторую частную производную функции $U(x, t)$ по пространственной координате ее конечно-разностной аппроксимацией. В результате получим бесконечную систему СОДУ для счетного числа случайных функций $U_\ell(t)$ с сосредоточенными белыми шумами $V_\ell(t)$ на входе (аналог метода прямых для детерминированных систем):

$$\dot{U}_\ell(t) + \alpha^2 U_\ell(t) = \beta^2 \frac{U_{\ell-1}(t) - 2U_\ell(t) + U_{\ell+1}(t)}{h^2} + [\gamma_0 + \gamma V_\ell(t)] U_\ell(t). \quad (3)$$

По системе (3) можно формально построить уравнение Фоккера–Планка–Колмогорова (ФПК-уравнение) для совместной плотности вероятности $p = p(\mathbf{u}, t)$ распределения функций $U_\ell(t)$

$$\frac{\partial p}{\partial t} = \frac{1}{2} \sum_{r, q=-\infty}^{+\infty} \frac{\partial^2 (b_{rq} p)}{\partial u_r \partial u_q} - \sum_{r=-\infty}^{+\infty} \frac{\partial (a_r p)}{\partial u_r} \equiv \text{L}p, \quad (4)$$

$$p(\mathbf{u}, 0) = p^0(\mathbf{u}), \quad (5)$$

где коэффициенты сноса a_r и диффузии b_{rq} вычисляются по соотношениям, аналогичным формулам Стратоновича:

$$a_r = f_r + \frac{1}{2} \sum_{q=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \frac{\partial g_{rk}}{\partial u_q} g_{qk}, \quad b_{rq} = \sum_{k=-\infty}^{+\infty} g_{rk} g_{qk}.$$

Учитывая структуру уравнений (3), можем записать:

$$f_r = -\alpha^2 u_r + \beta^2 \frac{u_{r-1} - 2u_r + u_{r+1}}{h^2} + \gamma_0 u_r, \quad g_{rk} = \delta_{rk} \gamma u_r,$$

где δ_{rk} – символ Кронекера. Отсюда находим:

$$a_r = \mu u_r + \beta^2 \frac{u_{r-1} - 2u_r + u_{r+1}}{h^2}, \quad b_{rq} = \delta_{rq} \gamma^2 u_r^2,$$

где $\mu = \gamma_0 + \frac{1}{2} \gamma^2 - \alpha^2$. При этом ФПК-уравнение примет следующую форму:

$$\frac{\partial p}{\partial t} = \frac{\gamma^2}{2} \sum_{r=-\infty}^{+\infty} \frac{\partial^2 (u_r^2 p)}{\partial u_r^2} - \sum_{r=-\infty}^{+\infty} \frac{\partial (a_r p)}{\partial u_r}. \quad (6)$$

Теперь можно построить систему обыкновенных дифференциальных уравнений (ОДУ) для моментных функций различного порядка векторного случайного процесса

$$\mathbf{U}(t) = \{\dots, U_{-n}(t), \dots, U_{-2}(t), U_{-1}(t), U_0(t), U_1(t), U_2(t), \dots, U_n(t), \dots\}.$$

Для этого будем домножать обе части уравнения (6) на различные произведения компонент u_ℓ вектора $\mathbf{u} = \{\dots, u_{-n}, \dots, u_{-2}, u_{-1}, u_0, u_1, u_2, \dots, u_n, \dots\}$ вида u_{ℓ_1} , $u_{\ell_1}^2$, $u_{\ell_1} u_{\ell_2}$, $u_{\ell_1}^3$, $u_{\ell_1}^2 u_{\ell_2}$, $u_{\ell_1} u_{\ell_2} u_{\ell_3}$, $u_{\ell_1}^4$, $u_{\ell_1}^3 u_{\ell_2}$, $u_{\ell_1}^2 u_{\ell_2}^2$, $u_{\ell_1}^2 u_{\ell_2} u_{\ell_3}$, $u_{\ell_1} u_{\ell_2} u_{\ell_3} u_{\ell_4}$ ($\ell_1 \neq \ell_2 \neq \ell_3 \neq \ell_4$) и интегрировать по векторному пространству счетной размерности, получая ОДУ для моментных функций $M_{\ell_1}(t)$, $M_{\ell_1 \ell_1}(t)$, $M_{\ell_1 \ell_2}(t)$, $M_{\ell_1 \ell_1 \ell_1}(t)$, $M_{\ell_1 \ell_1 \ell_2}(t)$, $M_{\ell_1 \ell_2 \ell_3}(t)$, $M_{\ell_1 \ell_1 \ell_1 \ell_1}(t)$, $M_{\ell_1 \ell_1 \ell_1 \ell_2}(t)$, $M_{\ell_1 \ell_1 \ell_2 \ell_2}(t)$, $M_{\ell_1 \ell_1 \ell_2 \ell_3}(t)$, $M_{\ell_1 \ell_2 \ell_3 \ell_4}(t)$. В результате получим:

$$\begin{aligned} \dot{M}_{\ell_1}(t) &= \mu M_{\ell_1}(t) + \beta^2 \frac{M_{\ell_1-1}(t) - 2M_{\ell_1}(t) + M_{\ell_1+1}(t)}{h^2}, \\ \dot{M}_{\ell_1 \ell_1}(t) &= (\gamma^2 + 2\mu) M_{\ell_1 \ell_1}(t) + 2\beta^2 \frac{M_{\ell_1, \ell_1-1}(t) - 2M_{\ell_1 \ell_1}(t) + M_{\ell_1, \ell_1+1}(t)}{h^2}, \\ \dot{M}_{\ell_1 \ell_2}(t) &= 2\mu M_{\ell_1 \ell_2}(t) + \beta^2 \frac{M_{\ell_1-1, \ell_2}(t) - 2M_{\ell_1 \ell_2}(t) + M_{\ell_1+1, \ell_2}(t)}{h^2} + \\ &\quad + \beta^2 \frac{M_{\ell_1, \ell_2-1}(t) - 2M_{\ell_1 \ell_2}(t) + M_{\ell_1, \ell_2+1}(t)}{h^2}, \end{aligned}$$

$$\begin{aligned}\dot{M}_{\ell_1 \ell_1 \ell_1}(t) &= (3\gamma^2 + 3\mu) M_{\ell_1 \ell_1 \ell_1}(t) + \\ &\quad + 3\beta^2 \frac{M_{\ell_1 \ell_1, \ell_1-1}(t) - 2M_{\ell_1 \ell_1 \ell_1}(t) + M_{\ell_1 \ell_1, \ell_1+1}(t)}{h^2},\end{aligned}$$

$$\begin{aligned}\dot{M}_{\ell_1 \ell_1 \ell_2}(t) &= (\gamma^2 + 3\mu) M_{\ell_1 \ell_1 \ell_2}(t) + \\ &\quad + 2\beta^2 \frac{M_{\ell_1, \ell_1-1, \ell_2}(t) - 2M_{\ell_1 \ell_1 \ell_2}(t) + M_{\ell_1, \ell_1+1, \ell_2}(t)}{h^2} + \\ &\quad + \beta^2 \frac{M_{\ell_1 \ell_1, \ell_2-1}(t) - 2M_{\ell_1 \ell_1 \ell_2}(t) + M_{\ell_1 \ell_1, \ell_2+1}(t)}{h^2},\end{aligned}$$

$$\begin{aligned}\dot{M}_{\ell_1 \ell_2 \ell_3}(t) &= 3\mu M_{\ell_1 \ell_1 \ell_2}(t) + \beta^2 \frac{M_{\ell_1-1, \ell_2 \ell_3}(t) - 2M_{\ell_1 \ell_2 \ell_3}(t) + M_{\ell_1+1, \ell_2 \ell_3}(t)}{h^2} + \\ &\quad + \beta^2 \frac{M_{\ell_1, \ell_2-1, \ell_3}(t) - 2M_{\ell_1 \ell_2 \ell_3}(t) + M_{\ell_1, \ell_2+1, \ell_3}(t)}{h^2} + \\ &\quad + \beta^2 \frac{M_{\ell_1 \ell_2, \ell_3-1}(t) - 2M_{\ell_1 \ell_2 \ell_3}(t) + M_{\ell_1 \ell_2, \ell_3+1}(t)}{h^2},\end{aligned}$$

$$\begin{aligned}\dot{M}_{\ell_1 \ell_1 \ell_1 \ell_1}(t) &= (6\gamma^2 + 4\mu) M_{\ell_1 \ell_1 \ell_1 \ell_1}(t) + \\ &\quad + 4\beta^2 \frac{M_{\ell_1 \ell_1 \ell_1, \ell_1-1}(t) - 2M_{\ell_1 \ell_1 \ell_1 \ell_1}(t) + M_{\ell_1 \ell_1 \ell_1, \ell_1+1}(t)}{h^2},\end{aligned}$$

$$\begin{aligned}\dot{M}_{\ell_1 \ell_1 \ell_1 \ell_2}(t) &= (3\gamma^2 + 4\mu) M_{\ell_1 \ell_1 \ell_1 \ell_2}(t) + \\ &\quad + 3\beta^2 \frac{M_{\ell_1 \ell_1, \ell_1-1, \ell_2}(t) - 2M_{\ell_1 \ell_1 \ell_1 \ell_2}(t) + M_{\ell_1 \ell_1, \ell_1+1, \ell_2}(t)}{h^2} + \\ &\quad + \beta^2 \frac{M_{\ell_1 \ell_1 \ell_1, \ell_2-1}(t) - 2M_{\ell_1 \ell_1 \ell_1 \ell_2}(t) + M_{\ell_1 \ell_1 \ell_1, \ell_2+1}(t)}{h^2},\end{aligned}$$

$$\begin{aligned}\dot{M}_{\ell_1 \ell_1 \ell_2 \ell_2}(t) &= (2\gamma^2 + 4\mu) M_{\ell_1 \ell_1 \ell_2 \ell_2}(t) + \\ &\quad + 2\beta^2 \frac{M_{\ell_1, \ell_1-1, \ell_2 \ell_2}(t) - 2M_{\ell_1 \ell_1 \ell_2 \ell_2}(t) + M_{\ell_1 \ell_1+1, \ell_2 \ell_2}(t)}{h^2} + \\ &\quad + 2\beta^2 \frac{M_{\ell_1 \ell_1 \ell_2, \ell_2-1}(t) - 2M_{\ell_1 \ell_1 \ell_2 \ell_2}(t) + M_{\ell_1 \ell_1 \ell_2, \ell_2+1}(t)}{h^2},\end{aligned}$$

$$\begin{aligned}\dot{M}_{\ell_1 \ell_1 \ell_2 \ell_3}(t) &= (\gamma^2 + 4\mu) M_{\ell_1 \ell_1 \ell_2 \ell_3}(t) + \\ &\quad + 2\beta^2 \frac{M_{\ell_1, \ell_1-1, \ell_2 \ell_3}(t) - 2M_{\ell_1 \ell_1 \ell_2 \ell_3}(t) + M_{\ell_1 \ell_1+1, \ell_2 \ell_3}(t)}{h^2} + \\ &\quad + \beta^2 \frac{M_{\ell_1 \ell_1, \ell_2-1, \ell_3}(t) - 2M_{\ell_1 \ell_1 \ell_2 \ell_3}(t) + M_{\ell_1 \ell_1, \ell_2+1, \ell_3}(t)}{h^2} + \\ &\quad + \beta^2 \frac{M_{\ell_1 \ell_1 \ell_2, \ell_3-1}(t) - 2M_{\ell_1 \ell_1 \ell_2 \ell_3}(t) + M_{\ell_1 \ell_1 \ell_2, \ell_3+1}(t)}{h^2},\end{aligned}$$

$$\begin{aligned}\dot{M}_{\ell_1 \ell_2 \ell_3 \ell_4}(t) &= 4\mu M_{\ell_1 \ell_2 \ell_3 \ell_4}(t) + \\ &\quad + \beta^2 \frac{M_{\ell_1-1, \ell_2 \ell_3 \ell_4}(t) - 2M_{\ell_1 \ell_2 \ell_3 \ell_4}(t) + M_{\ell_1+1, \ell_2 \ell_3 \ell_4}(t)}{h^2} + \\ &\quad + \beta^2 \frac{M_{\ell_1, \ell_2-1, \ell_3 \ell_4}(t) - 2M_{\ell_1 \ell_2 \ell_3 \ell_4}(t) + M_{\ell_1, \ell_2+1, \ell_3 \ell_4}(t)}{h^2} +\end{aligned}$$

$$\begin{aligned}
& + \beta^2 \frac{M_{\ell_1 \ell_2, \ell_3-1, \ell_4}(t) - 2 M_{\ell_1 \ell_2 \ell_3 \ell_4}(t) + M_{\ell_1 \ell_2, \ell_3+1, \ell_4}(t)}{h^2} + \\
& + \beta^2 \frac{M_{\ell_1 \ell_2, \ell_3, \ell_4-1}(t) - 2 M_{\ell_1 \ell_2 \ell_3 \ell_4}(t) + M_{\ell_1 \ell_2 \ell_3, \ell_4+1}(t)}{h^2}.
\end{aligned}$$

Возвратимся теперь к непрерывной среде, трактуя полученные выше уравнения для моментов функций $U_\ell(t)$ как конечно-разностные аппроксимации соответствующих уравнений для моментов исходного случайного поля $U(x, t)$, а именно:

$$\begin{aligned}
& \left. \frac{\partial^2 m_1(x, t)}{\partial x^2} \right|_{x=x_{\ell_1}} = \left. \frac{\partial^2 \{ \mathbb{E}[U(x, t)] \}}{\partial x^2} \right|_{x=x_{\ell_1}} = \mathbb{E} \left[\left. \frac{\partial^2 U(x, t)}{\partial x^2} \right|_{x=x_{\ell_1}} \right] \approx \\
& \approx \mathbb{E} \left[\frac{U_{\ell_1-1}(t) - 2U_{\ell_1}(t) + U_{\ell_1+1}(t)}{h^2} \right] = \frac{M_{\ell_1-1}(t) - 2M_{\ell_1}(t) + M_{\ell_1+1}(t)}{h^2}, \\
& \left. \frac{\partial^2 m_{11}(x, y, t)}{\partial x^2} \right|_{x=x_{\ell_1}, y=x_{\ell_2}} = \left. \frac{\partial^2 \{ \mathbb{E}[U(x, t) U(y, t)] \}}{\partial x^2} \right|_{x=x_{\ell_1}, y=x_{\ell_2}} = \\
& = \mathbb{E} \left[\left. \frac{\partial^2 \{ U(x, t) U(y, t) \}}{\partial x^2} \right|_{x=x_{\ell_1}, y=x_{\ell_2}} \right] = \mathbb{E} \left[\left. \frac{\partial^2 U(x, t)}{\partial x^2} \right|_{x=x_{\ell_1}} U(y, t) \right|_{y=x_{\ell_2}} \right] \approx \\
& \approx \mathbb{E} \left[\frac{U_{\ell_1-1}(t) U_{\ell_2}(t) - 2U_{\ell_1}(t) U_{\ell_2}(t) + U_{\ell_1+1}(t) U_{\ell_2}(t)}{h^2} \right] = \\
& = \frac{M_{\ell_1-1, \ell_2}(t) - 2M_{\ell_1 \ell_2}(t) + M_{\ell_1+1, \ell_2}(t)}{h^2}, \\
& \left. \frac{\partial^2 m_{111}(x, y, z, t)}{\partial x^2} \right|_{x=x_{\ell_1}, y=x_{\ell_2}, z=x_{\ell_3}} = \\
& = \left. \frac{\partial^2 \{ \mathbb{E}[U(x, t) U(y, t) U(z, t)] \}}{\partial x^2} \right|_{x=x_{\ell_1}, y=x_{\ell_2}, z=x_{\ell_3}} = \\
& = \mathbb{E} \left[\left. \frac{\partial^2 \{ U(x, t) U(y, t) U(z, t) \}}{\partial x^2} \right|_{x=x_{\ell_1}, y=x_{\ell_2}, z=x_{\ell_3}} \right] = \\
& = \mathbb{E} \left[\left. \frac{\partial^2 U(x, t)}{\partial x^2} \right|_{x=x_{\ell_1}} U(y, t) U(z, t) \right|_{y=x_{\ell_2}, z=x_{\ell_3}} \right] \approx \\
& \approx \mathbb{E} \left[\frac{U_{\ell_1-1}(t) U_{\ell_2}(t) U_{\ell_3}(t) - 2U_{\ell_1}(t) U_{\ell_2}(t) U_{\ell_3}(t) + U_{\ell_1+1}(t) U_{\ell_2}(t) U_{\ell_3}(t)}{h^2} \right] = \\
& = \frac{M_{\ell_1-1, \ell_2 \ell_3}(t) - 2M_{\ell_1 \ell_2 \ell_3}(t) + M_{\ell_1+1, \ell_2 \ell_3}(t)}{h^2}
\end{aligned}$$

и т.д. Кроме того, с точностью порядка $O(h)$ можно получить соотношение

$$\frac{M_{\ell_1, \ell_1-1}(t) - 2M_{\ell_1 \ell_1}(t) + M_{\ell_1, \ell_1+1}(t)}{h^2} \approx$$

$$\approx \frac{M_{\ell_1-1, \ell_1-1}(t) - 2 M_{\ell_1 \ell_1}(t) + M_{\ell_1+1, \ell_1+1}(t)}{h^2} \approx \frac{\partial^2 m_2(x, t)}{\partial x^2} \Big|_{x=x_{\ell_1}},$$

а затем по аналогии соответствующие формулы и для других комбинаций индексов. Тогда будем иметь:

$$\begin{aligned} \frac{\partial m_1(x, t)}{\partial t} &= \mu m_1(x, t) + \beta^2 \frac{\partial^2 m_1(x, t)}{\partial x^2}, \\ \frac{\partial m_2(x, t)}{\partial t} &= (\gamma^2 + 2\mu) m_2(x, t) + 2\beta^2 \frac{\partial^2 m_2(x, t)}{\partial x^2}, \\ \frac{\partial m_{11}(x, y, t)}{\partial t} &= \mu m_{11}(x, y, t) + \beta^2 \left[\frac{\partial^2 m_{11}(x, y, t)}{\partial x^2} + \frac{\partial^2 m_{11}(x, y, t)}{\partial y^2} \right], \\ \frac{\partial m_3(x, t)}{\partial t} &= (3\gamma^2 + 3\mu) m_3(x, t) + 3\beta^2 \frac{\partial^2 m_3(x, t)}{\partial x^2}, \\ \frac{\partial m_{21}(x, y, t)}{\partial t} &= (\gamma^2 + 3\mu) m_{21}(x, y, t) + \\ &+ \beta^2 \left[2 \frac{\partial^2 m_{21}(x, y, t)}{\partial x^2} + \frac{\partial^2 m_{11}(x, y, t)}{\partial y^2} \right], \\ \frac{\partial m_{111}(x, y, z, t)}{\partial t} &= 3\mu m_{111}(x, y, z, t) + \\ &+ \beta^2 \left[\frac{\partial^2 m_{111}(x, y, z, t)}{\partial x^2} + \frac{\partial^2 m_{111}(x, y, z, t)}{\partial y^2} + \frac{\partial^2 m_{111}(x, y, z, t)}{\partial z^2} \right], \\ \frac{\partial m_4(x, t)}{\partial t} &= (6\gamma^2 + 4\mu) m_4(x, t) + 4\beta^2 \frac{\partial^2 m_4(x, t)}{\partial x^2}, \\ \frac{\partial m_{31}(x, y, t)}{\partial t} &= (3\gamma^2 + 4\mu) m_{31}(x, y, t) + \\ &+ \beta^2 \left[3 \frac{\partial^2 m_{31}(x, y, t)}{\partial x^2} + \frac{\partial^2 m_{31}(x, y, t)}{\partial y^2} \right], \\ \frac{\partial m_{22}(x, y, t)}{\partial t} &= (2\gamma^2 + 4\mu) m_{22}(x, y, t) + \\ &+ \beta^2 \left[2 \frac{\partial^2 m_{22}(x, y, t)}{\partial x^2} + 2 \frac{\partial^2 m_{22}(x, y, t)}{\partial y^2} \right], \\ \frac{\partial m_{211}(x, y, z, t)}{\partial t} &= (\gamma^2 + 4\mu) \mu m_{211}(x, y, z, t) + \\ &+ \beta^2 \left[2 \frac{\partial^2 m_{211}(x, y, z, t)}{\partial x^2} + \frac{\partial^2 m_{211}(x, y, z, t)}{\partial y^2} + \frac{\partial^2 m_{211}(x, y, z, t)}{\partial z^2} \right], \\ \frac{\partial m_{1111}(x, y, z, \rho, t)}{\partial t} &= 4\mu m_{1111}(x, y, z, \rho, t) + \beta^2 \left[\frac{\partial^2 m_{1111}(x, y, z, \rho, t)}{\partial x^2} + \right. \\ &+ \left. \frac{\partial^2 m_{1111}(x, y, z, \rho, t)}{\partial y^2} + \frac{\partial^2 m_{1111}(x, y, z, \rho, t)}{\partial z^2} + \frac{\partial^2 m_{1111}(x, y, z, \rho, t)}{\partial \rho^2} \right], \end{aligned} \tag{7}$$

где обозначено

$$\begin{aligned}
 m_i(x, t) &= \mathbb{E}[U^i(x, t)], \quad m_{ij}(x, y, t) = \mathbb{E}[U^i(x, t) U^j(y, t)], \\
 m_{ijk}(x, y, z, t) &= \mathbb{E}[U^i(x, t) U^j(y, t) U^k(z, t)], \\
 1 \leq i + j + k &\leq 4, \quad k \leq j \leq i, \\
 m_{1111}(x, y, z, \rho, t) &= \mathbb{E}[U(x, t) U(y, t) U(z, t) U(\rho, t)].
 \end{aligned}$$

Заметим, что определять начальную плотность $p^0(\mathbf{u})$ в данной схеме нет необходимости, т.к. ФПК-уравнение используется только как промежуточный инструмент получения уравнений для моментов.

4. Пример

Воспользуемся полученными уравнениями для вычисления $m_3(x, t)$ и $m_4(x, t)$ для следующих исходных параметров:

$$\alpha = 3/4, \quad \beta = 1/8, \quad \gamma = \gamma_0 = 1/5.$$

Предположим, что начальное поле $U^0(x) = \exp(-x^2)$, т.е. является неслучайным. Отсюда

$$m_1^0(x) = e^{-x^2}, \quad m_3^0(x) = [m_1^0(x)]^3 = e^{-3x^2}, \quad m_4^0(x) = [m_1^0(x)]^4 = e^{-4x^2}.$$

Результаты расчетов поведения $m_3(x, t)$ и $m_4(x, t)$ – профили этих функций при значениях времени t от 0 до 2 с шагом 0,2 приведены на рис. 1 и 2.

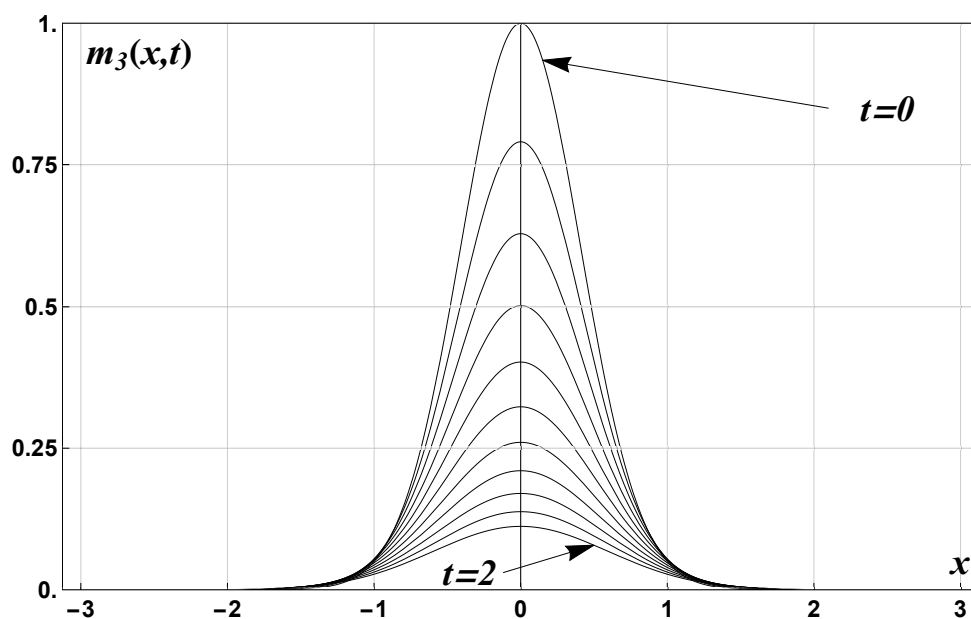


Рис. 1

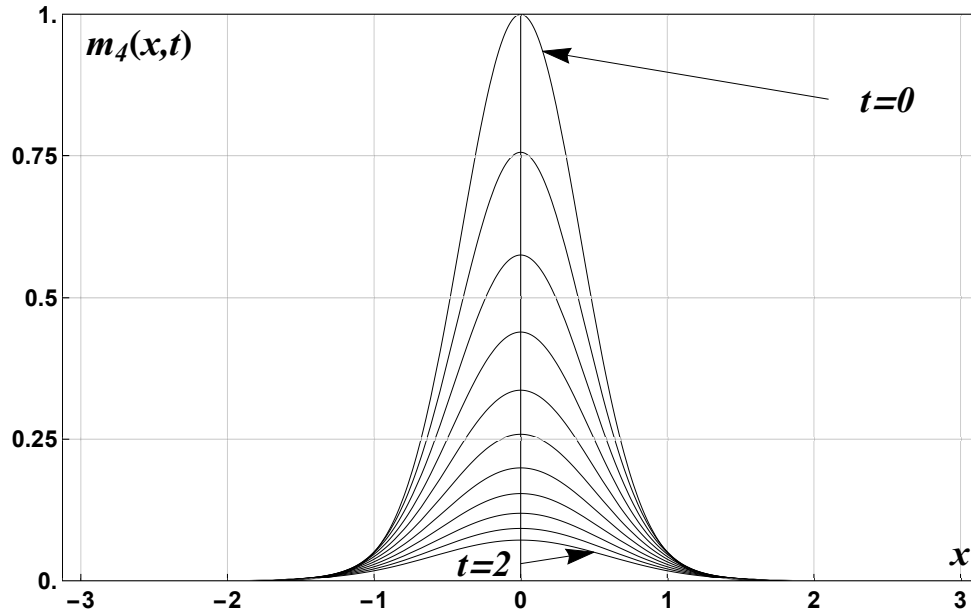


Рис. 2

5. Заключение

Анализируя представленные результаты, несложно установить, что изложенная схема может быть применена для расчета старших моментов решений линейных параметрических эволюционных СДУ в ЧП более высоких (по времени) порядков, а также систем таких уравнений.

Библиографический список

1. *Розовский Б.Л.* Эволюционные стохастические системы. Линейная теория и приложения к статистике случайных процессов. – М.: Наука, 1983. – 208 с.
2. *Lord G.J., Powell C.E., Shardlow T.* An introduction to computational stochastic PDEs. – New York: Cambridge Univ. Press, 2014. – XVII, 503 p.
3. *Болотин В.В.* Случайные колебания упругих систем. – М.: Наука, 1979. – 336 с.
4. *Светлицкий В.А.* Случайные колебания механических систем. – М.: Машиностроение, 1991. – 320 с.
5. *Николаенко Н.А., Ульянов С.В.* Статистическая динамика машиностроительных конструкций. – М.: Машиностроение, 1977. – 368 с.

6. *Коненков Ю.К., Давтян М.Д.* Случайные механические процессы в оборудовании машин. – М.: Машиностроение, 1988. – 272 с.
7. *Макеев В.П., Гриненко Н.И., Павлюк Ю.С.* Статистические задачи динамики упругих конструкций. – М.: Наука, 1984. – 232 с.
8. *Bucher C.* Computational analysis of randomness in structural mechanics. – London: Taylor & Francis, 2009. – XVIII, 230 p.
9. *Макаров Б.П.* Нелинейные задачи статистической динамики машин и приборов. – М.: Машиностроение, 1983. – 264 с.
10. *Ланда П.С.* Автоколебания в распределенных системах. – М.: Наука, 1983. – 320 с.
11. *Кафаров В.В., Дорохов И.Н., Липатов Л.Н.* Системный анализ процессов химической технологии. – М.: Наука, 1982. – 344 с.
12. *Кляцкин В.И.* Стохастические уравнения и волны в случайно неоднородных средах. – М.: Наука, 1980. – 336 с.
13. *Рытов С.М., Кравцов Ю.А., Татарский В.И.* Введение в статистическую радиофизику. – М.: Наука, 1978. – Ч. 2. Случайные поля. – 464 с.
14. *Монин А.С., Яглом А.М.* Статистическая гидромеханика. – М.: Наука, 1965. – Ч. 1. – 640 с.; 1967. – Ч. 2. – 720 с.
15. *Ortiz de Zàrate J.M., Sengers J.V.* Hydrodynamic fluctuations in fluids and fluid mixtures. – Amsterdam: Elsevier, 2006. – X, 309 p.
16. *Ohayon R., Soize C.* Structural acoustics and vibration mechanical models: Variational formulations and discretization. – London: Academic Press, 1998. – X, 424 p.
17. *Kulasiri D., Venvoerd W.* Stochastic dynamics: modeling solute transport in porous media. – Amsterdam: Elsevier, 2002. – XII, 239 p.
18. *den Hollander F.* Random polymers. – Berlin: Springer-Verlag, 2009. – XIII, 258 p.
19. *Kamiński M.M.* Computational mechanics of composite materials: sensitivity, randomness, and multiscale behaviour. – London: Springer-Verlag, 2005. – XVII, 418 p.
20. *Haile J.M.* Molecular dynamics simulation: elementary methods. – New York: John Wiley & Sons, 1992. – XVIII, 489 p.

21. *Ostoja-Starzewski M.* Microstructural randomness and scaling in mechanics of materials. – New York: Taylor & Francis, 2008. – XXV, 471 p.
22. *Radons G., Rumpf B., Schuster H.G.* (eds.) Nonlinear dynamics of nanosystems. – Weinheim: Wiley-VCH Verlag, 2010. – XX, 455 p.
23. *Balendra T.* Vibration of buildings to wind and earthquake loads. – Berlin: Springer-Verlag, 1993. – X, 149 p.
24. *Шмелев А.Б.* Основы марковской теории нелинейной обработки случайных полей. – М.: Изд-во МФТИ, 1998. – 208 с.
25. *Potapov V.D.* Stability of stochastic elastic and viscoelastic systems. – New York: John Wiley & Sons, 1999. – 275 p.
26. *Schwarz A., Janicka J.* (eds.) Combustion noise. – Berlin: Springer-Verlag, 2009. – XVII, 291 p.
27. *Elishakoff I.* Probabilistic theory of structures. – 2nd ed. – Mineola (NY): Dover, 1999. – XVII, 492 p.
28. *Гончаренко В.М.* О случайных колебаниях упругих тел и теория марковских процессов // ПМ. – 1991. – Т. 27, № 8. – С. 95–100.
29. *Hopf E.* Statistical hydrodynamics and functional calculus // Journal of Rational Mechanics and Analysis. – 1952. – Vol. 1, P. 87–123.
30. *Ghanem R., Spanos P.D.* Stochastic finite elements: a spectral approach. – New York: Springer-Verlag, 1991. – XI, 222 p.
31. *Holmes P., Lumley J.L., Berkooz G.* Turbulence, coherent structures, dynamical systems, and symmetry. – Cambridge: Cambridge Univ. Press, 1996. – XVIII, 420 p.
32. *McDonald S.* Finite difference approximation for linear stochastic partial differential equations with method of lines // MPRA Paper No. 3983. – 2007. – 18 p. – URL: mpra.ub.uni-muenchen.de/3983/1/MPRA_paper_3983.pdf
33. *Cao T., Yin L.* Spectral Galerkin method for stochastic wave equations driven by space-time white noise // Communications on Pure and Applied Analysis. – 2007. – Vol. 6, № 3. – P. 607–617.
34. *Tokareva S.* Stochastic finite volume methods for computational uncertainty quantification in hyperbolic conservation laws. – Diss. ... Doctor of Sciences, № 21498. – Zürich: ETH, 2013. – VI, 130 p.

35. *Baroth J., Bressollette Ph., Chauvière C., Fogli M.* An efficient SFE method using Lagrange polynomials: Application to nonlinear mechanical problems with uncertain parameters // *Computer Methods in Applied Mechanics and Engineering*. – 2007. – Vol. 196, № 45–48. – P. 4419–4429.
36. *Fox R.O.* Quadrature-based moment methods for polydisperse multiphase flows // *Stochastic Methods in Fluid Mechanics* / J.-P. Minier, S. Chibbaro (eds.). – Wien: Springer-Verlag; Udine: CISM, 2014. – P. 87–136.
37. *Burczyński T.* Stochastic boundary element methods: Computational methodology and applications // *Proc. of IUTAM Symp. "Probabilistic Structural Mechanics: Advances in Structural Reliability Methods"* (San Antonio, USA, 1993) / P.D. Spanos, Y.-T. Wu (eds.). – Berlin: Springer-Verlag, 1994. – P. 42–55.
38. *Kroese D.P., Taimre T., Botev Z.I.* Handbook of Monte Carlo methods. – Hoboken (NJ): John Wiley & Sons, 2011. – XIX, 743 p.
39. *Полосков И.Е.* Об одном подходе к анализу случайных процессов в распределенных системах // *Математическое моделирование*. – 2003. – Т. 15, № 4. – С. 85–100.
40. *Гардинер К.В.* Стохастические методы в естественных науках. – М.: Мир, 1986. – 528 с.

Equations for Higher Moment Characteristics of Solution of a Linear Stochastic Partial Differential Equation with Additive and Multiplicative Perturbations

Igor E. Poloskov¹

¹ Perm State National Research University;
polosk@psu.ru; +7(342) 2396560, Russia, Perm, Bukirev str., 15

Abstract. *This paper is devoted to consideration of the problem to obtain differential equations for the higher moment functions of solution for a linear stochastic partial differential equation (SPDE) of parabolic type with additive and multiplicative perturbations. The problem is resolved on the basis of a usage of our discretization scheme, previously developed, for the general SPDE systems with subsequent return to a continuous medium. This scheme allows to solve a number of problems related to both dynamic analysis of the behavior of stochastic distributed systems and obtaining stationary characteristics, in particular, probability density functionals.*

Key words: *stochastic partial differential equation, system with distributed parameters, random field, moment functions, parametric fluctuation, Fokker–Planck–Kolmogorov equation.*

Содержание

От редакционной коллегии	3
--------------------------------	---

Задачи статистической теории оценивания и проверки гипотез

Абдушукуров А. А., Рахимова Г. Г. Асимптотические свойства последовательного интервального оценивания функционалов от функции распределения	4–14
Королев В. Ю., Фомина Д. В., Щербинина А. А. Комбинированный сеточный метод оценивания параметров распределений Линника и Миттаг–Леффлера	15–34
Орлов А. И. Скорость сходимости ядерных оценок плотности в пространствах произвольной природы	35–45
Пагурова В. И. О некоторых свойствах порядковых статистик	46–52
Чичагов В. В. Сравнение UMVUE и MLE с помощью степенной функции риска в случае однопараметрического экспоненциального семейства непрерывных распределений	53–65

Применение и исследование вероятностно-статистических моделей

Дорофеева А. В. Неравенства типа Каца–Петрова для некоторых случайных сумм	66–75
Захарова Т. В. Применение EM-алгоритма для обработки циклических сигналов	76–90
Захарова Т. В., Меньших И. М. Об асимптотически оптимальных размещениях станций обслуживания на прямой	91–101
Макаров А. А., Симонова Г. И., Зверева Д. И. К методике кластеризации крупных городских школ России. Факторы, отвечающие за результаты образовательной деятельности	102–118
Полосков И. Е. Уравнения для старших моментных характеристик решения линейного стохастического дифференциального уравнения в частных производных с аддитивными и мультипликативными возмущениями	119–131