S©MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru

Том 12, стр. 766–776 (2015) УДК 519.21
DOI 10.17377/semi.2015.12.062 MSC 62G07

ОБ УСЛОВИЯХ ГАУССОВСКОЙ АППРОКСИМАЦИИ ЯДЕРНЫХ ОЦЕНОК ДЛЯ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ

А.С. КАРТАШОВ, А.И. САХАНЕНКО

Abstract. Recently E. Gine, V. Koltchinskii and L. Sakhanenko (2004) investigated asymptotical behavior of a random variable of the form $\sqrt{nh_n}\sup_{t\in\mathbf{R}}|\psi(t)(f_n(t)-\mathbf{E}f_n(t))|$ with some weight function $\psi(t)$, where f_n is a kernel density estimator. The proof of their limit theorems consists of a large number of technically difficult stages and uses more than ten bulky assumptions. In this work we show that under simpler and wider conditions the above stated problem is reduced to the study of asymptotics of a supremum of some special Gaussian process. The obtained result can be used in further investigation of functionals based on empirical processes and kernel density estimators. Our proof is based on the well-known approximation of Komlos, Major and Tusnady (1975).

Keywords: kernel density estimators, brownian motion, brownian bridge, KMT approximation, function of bounded variation.

1. Введение

Пусть $X, X_1, X_2, ...$ — независимые одинаково распределенные случайные величины с функцией распределения $F(x)$. Введем в рассмотрение случайную функцию, называемую ядерной оценкой плотности распределения случайной величины X:

$$f_n(t) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_n}\right),$$

Kartashov, A.S., Sakhanenko, A.I., About conditions of gaussian approximation of kernel estimates for distribution density.

Работа поддержана РФФИ (гранты РФФИ 15-01-07460а и 14-01-00220а).

Поступила 29 августа 2015 г., опубликована 5 ноября 2015 г.

где ширина полосы наблюдений $h_n > 0$ такова, что

$$h_n \to 0 \quad \text{и} \quad nh_n \to \infty \quad \text{при} \quad n \to \infty,$$

а K — функция ограниченной вариации. Если плотность $f(x)$ существует и $\int K(x)dx = 1$, то $f_n(t)$ можно рассматривать в качестве оценки неизвестной плотности $f(x)$.

В данной работе нас интересует асимптотика распределения случайной величины $\beta_n(\mathbf{R})$, где

$$\beta_n(T) := \sqrt{nh_n} \sup_{t \in T} \left|\psi(t)(f_n(t) - \mathbf{E}f_n(t))\right|, \tag{1}$$

а $\psi(t) \geq 0$ — некоторая весовая функция. В [1] было показано, что при очень серьезных дополнительных условиях, большинство из которых восходит к известной работе [2], имеет место следующая сходимость:

$$P\left(A_n\left(\frac{\beta_n(\mathbf{R})}{\|K\|_2} - A_n\right) \leq x\right) \longrightarrow e^{-e^{-x}} \quad \text{при} \quad n \to \infty,$$

где $\|K\|_2^2 = \int\limits_{-\infty}^{\infty} K^2(x)dx$, а A_n определяется как решение некоторого функционального уравнения. В частности (см. [3])

$$A_n \sim \sqrt{2\left|\log h_n\right|} \cdot \sup_{t \in \mathbf{R}} \left|\psi(t)\sqrt{f(t)}\right|.$$

В данной статье мы укажем способ изучения асимптотических свойств величины $\beta_n(\mathbf{R})$, который позволяет избежать введения большинства громоздких ограничений из [1] и [2]. Для этого мы аппроксимируем величину $\beta_n(T)$ из (1) аналогичным супремумом, но для специально подобранных гауссовских процессов, более сложно устроенных, чем использовались в [1]. Поскольку асимптотика распределений супремумов гауссовских процессов изучена достаточно хорошо, то мы предполагаем в дальнейшем установить эту асимптотику при достаточно простых условиях по аналогии с теоремой 2 в [1].

Условимся, что далее в работе все пределы берутся при $n \to \infty$.

2. Основные результаты

Теорема 1. *Пусть K — функция ограниченной вариации, $\psi \geq 0$ — некоторая весовая функция, а $h_n > 0, b_n > 0$ и $C_n > 0$ — некоторые числа такие, что*

$$C_n = o\left(\frac{\sqrt{nh_n}}{b_n \log n}\right) \quad \textit{при} \quad n \to \infty.$$

Тогда на одном вероятностном пространстве с наблюдениями $\{X_i\}$ можно построить такой броуновский мост $W_n(t)$, что

$$\Delta_{0n}(\boldsymbol{R}) := b_n\sqrt{nh_n} \sup_{t \in \boldsymbol{R}} \left|\overline{\psi}_n(t)\delta_{0n}(t)\right| \xrightarrow{\boldsymbol{P}} 0, \tag{2}$$

где $\overline{\psi}_n(t) = \min\{\psi(t), C_n\}$ — срезка функции $\psi(t)$, а

$$\delta_{0n}(t) = f_n(t) - \boldsymbol{E}f_n(t) - \gamma_{0n}(t) \quad \textit{и} \quad \gamma_{0n}(t) = \frac{1}{\sqrt{nh_n}} \int\limits_{-\infty}^{\infty} W_n(F(t - vh_n))dK(v).$$

С помощью непосредственной проверки можно убедиться, что $\gamma_{0n}(t)$ — это гауссовский процесс с той же самой ковариационной функцией, что и у процесса $f_n(t) - \mathbf{E}f_n(t)$.

Если дополнительно наложить еще два ограничения, то в теореме 1 броуновский мост можно заменить на винеровский процесс:

Теорема 2. *Пусть выполнены условия теоремы* 1, *функция* K *имеет носитель на отрезке* $[-\frac{1}{2}, \frac{1}{2}]$ *и*

$$\overline{D}_n := \sup_{t\in\boldsymbol{R}}\left|\overline{\psi}_n(t)\cdot b_n\frac{F(t+h_n/2)-F(t-h_n/2)}{\sqrt{h_n}}\right| \longrightarrow 0 \quad \text{при} \quad n\to\infty. \tag{3}$$

Тогда на одном вероятностном пространстве с наблюдениями $\{X_i\}$ *можно построить такое стандартное броуновское движение* $B_n(t)$, *что*

$$\Delta_n(\boldsymbol{R}) := b_n\sqrt{nh_n}\sup_{t\in\boldsymbol{R}}\left|\overline{\psi}_n(t)\delta_n(t)\right| \xrightarrow{\boldsymbol{P}} 0, \tag{4}$$

при

$$\delta_n(t) = f_n(t) - \boldsymbol{E}f_n(t) - \gamma_n(t) \quad \text{и} \quad \gamma_n(t) = \frac{1}{\sqrt{n}h_n}\int\limits_{-\infty}^{\infty} B_n(F(t-vh_n))dK(v).$$

Далее введем в рассмотрение обозначения

$$\widetilde{\beta}_n(T) := \sqrt{nh_n}\sup_{t\in T}\left|\psi(t)\gamma_n(t)\right|, \qquad T_n = \{t:\psi(t)\le C_n\}. \tag{5}$$

Следствие 1. *Пусть выполнены условия теоремы* 2. *Тогда для того чтобы имела место сходимость*

$$b_n(\beta_n(T_n) - a_n) \Rightarrow \zeta, \tag{6}$$

необходимо и достаточно, чтобы

$$b_n(\widetilde{\beta}_n(T_n) - a_n) \Rightarrow \zeta$$

при тех же самых ζ, $a_n > 0$ *и* $b_n > 0$.

Замечание 1. *Сравнивая утверждение приведенного выше следствия* 1 *с аналогичным утверждением из теоремы* 1 *в* [1] *нетрудно заметить, что наш результат имеет следующие достоинства. Во-первых, у нас меньше условий и они существенно более простые. Так мы не используем ни комплекс условий из* [2] *(которые в* [1] *называются* General assumptions*), ни дополнительные условия, накладываемые в* [1]. *Из ограничений, используемых в* [1], *нами задействованы лишь часть ограничений на функцию* K *(ограниченная вариация и ограниченный носитель). Более того, в нашем условии (3) не предполагается даже существования плотности. Во-вторых, наша аппроксимация* (4) *строится на значительно более широких множествах* T_n, *чем аналогичная аппроксимация в теореме* 1 *из* [1], *где используются множества*

$$D_a = \{|t|\le a, f(t)\ge 1/a\}.$$

В-третьих, наше доказательство значительно проще. Из него в частности видно, что T_n *— максимальное множество, на котором можно получить аппроксимацию вида* (4) *методом Комлоша — Майера — Тушнади. Единственный минус нашего следствия 1 по сравнению с теоремой 1 в* [1] *— немного более сложная структура аппроксимационного процесса* $\gamma_n(t)$. *Однако этот процесс* *гауссовский, то есть в дальнейшем нами дополнительно будут исследоваться свойства только гауссовских процессов.*

Введем обозначения

$$\psi_h(t) := \sup_{|t-\tau|\le h/2} \psi(\tau);$$

$$D_n(T) := \sup_{t\in T} \left| \psi(t) \cdot \frac{F(t+h_n/2) - F(t-h_n/2)}{h_n} \right|. \tag{7}$$

И будем далее через $\overline{T}_n$ обозначать дополнение к множеству T_n, введенному в формуле (5).

Теорема 3. *Пусть выполнены условия теоремы* 2, $a_n b_n \longrightarrow \infty$,

$$n\boldsymbol{P}(\psi_{h_n}(X_1) > C_n) \longrightarrow 0 \quad \textit{при} \quad n\to\infty, \tag{8}$$

$$\frac{\sqrt{nh_n}D_n(\overline{T}_n)}{a_n} \longrightarrow 0 \quad \textit{при} \quad n\to\infty. \tag{9}$$

Тогда для того чтобы имела место сходимость

$$b_n(\beta_n(\boldsymbol{R}) - a_n) \Rightarrow \zeta, \tag{10}$$

необходимо и достаточно, чтобы

$$b_n(\widetilde{\beta}_n(T_n) - a_n) \Rightarrow \zeta \tag{11}$$

при тех же самых $\zeta, a_n > 0, b_n > 0$.

Замечание 2. *Теорема 3 имеет два преимущества по сравнению с аналогичной теоремой* 6 *в* [1]. *Во-первых, у нее существенно меньше условий и они более простые. Во-вторых, у нее простое доказательство. Однако теорема* 3 *имеет также два недостатка. Во-первых, в ней более сложный аппроксимационный гауссовский процесс* $\gamma_n(t)$. *Во-вторых, ее условие* (8) *является чуть более жестким, чем аналогичное условие* (4.7) *в* [1].

Остальная часть работы будет посвящена доказательству сформулированных выше результатов.

3. Применение аппроксимации КМТ

Сначала займемся рассмотрением КМТ (Komlos, Major, Tusnady, 1975) аппроксимации, описанной в работе [4]. Она утверждает, что существует вероятностное пространство с последовательностью $\{\widetilde{\xi}_i\}$ независимых одинаково равномерно распределенных на отрезке $[0,1]$ случайных величин и последовательностью броуновских мостов $\widetilde{W}_n$ таких, что

$$\mathbf{P}\left(\left\|\widetilde{\alpha}_n - \widetilde{W}_n\right\|_\infty > \frac{x+k_1\log n}{\sqrt{n}}\right) \le k_0 e^{-k_2 x}, \quad \text{при всех} \ \ x\ge 0 \ \ \text{и} \ \ n\in\mathbf{N},$$

где k_0, k_1, k_2 — абсолютные константы, а $\widetilde{\alpha}_n(t) = \frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}\left(\mathbf{I}_{[0,t]}\left(\widetilde{\xi}_i\right) - t\right)$ — стандартный эмпирический процесс. В частности если взять $x = \frac{\log n}{k_2}$ и $k_3 = \frac{1}{k_2} + k_1$, то

$$\mathbf{P}\left(\left\|\widetilde{\alpha}_n - \widetilde{W}_n\right\|_\infty > \frac{k_3\log n}{\sqrt{n}}\right) \le \frac{k_0}{n} \underset{n\to\infty}{\longrightarrow} 0.$$

Отметим еще, что в этом случае $\widetilde{X}_i = F^{-1}(\widetilde{\xi}_i)$ — независимые случайные величины с функцией распределения $F(x)$.

Нам потребуется следующее утверждение, которое вытекает из теоремы 1 в работе А. В. Скорохода [5].

Лемма 1. *Пусть на том же вероятностном пространстве существует независимая от* $\{\widetilde{W}_n\}$ *и* $\{\widetilde{\xi}_n\}$ *(а значит, независимая и от* $\{\widetilde{X}_i\}$*) случайная величина* $\widetilde{\nu}$*, которая равномерно распределена на отрезке* $[0,1]$*. Тогда существует функция* $g(\cdot,\cdot)$ *и случайная величина* $\widetilde{\zeta}$ *такая, что*

(1) $\widetilde{\zeta} = \widetilde{\zeta}\left(\{\widetilde{\xi}_n\},\{\widetilde{W}_n\},\widetilde{\nu}\right)$ *и не зависит от* $\{\widetilde{X}_n\}$;

(2) $\widetilde{\zeta}$ *равномерно распределена на отрезке* $[0,1]$;

(3) $\left(\{\widetilde{\xi}_n\},\{\widetilde{W}_n\}\right) = g\left(\{\widetilde{X}_n\},\widetilde{\zeta}\right)$ *с вероятностью* 1.

Пусть теперь $\{X_i\}$ — случайные величины, о которых говорилось в начале работы, и которые одинаково распределены с $\{\widetilde{X}_i\}$, а ζ — случайная величина, имеющая равномерное на $[0,1]$ распределение и не зависящая от $\{X_i\}$. Тогда положим

$$(\{\xi_n\},\{W_n\}) = g\left(\{X_n\},\zeta\right),$$

и определим

$$\alpha_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\mathbf{I}_{[0,t]}(\xi_i) - t), \qquad X_i = F^{-1}(\xi_i).$$

В этом случае пятёрка случайных элементов $(\{\xi_n\},\{W_n\},\{\alpha_n\},\{X_n\},\zeta)$ очевидно одинаково распределена с пятёркой $\left(\{\widetilde{\xi}_n\},\{\widetilde{W}_n\},\{\widetilde{\alpha}_n\},\{\widetilde{X}_n\},\widetilde{\zeta}\right)$, поскольку пары $(\{X_n\},\zeta)$ и $\left(\{\widetilde{X}_n\},\widetilde{\zeta}\right)$ одинаково распределены по определению, а тройки $(\{\xi_n\},\{W_n\},\{\alpha_n\})$ и $\left(\{\widetilde{\xi}_n\},\{\widetilde{W}_n\},\{\widetilde{\alpha}_n\}\right)$ получены из упомянутых пар при помощи одних и тех же функций. В частности, в этом случае

$$\mathbf{P}\left(\|\alpha_n - W_n\|_\infty > \frac{k_3\log n}{\sqrt{n}}\right) \le \frac{k_0}{n} \underset{n\to\infty}{\to} 0. \tag{12}$$

Еще зададим на этом пространстве последовательность случайных величин $\{\eta_n\}$, имеющих стандартное нормальное распределение и не зависящих от $\{W_n\}$. И введем стандартные винеровские процессы $B_n(t) = W_n(t) + t\eta_n$.

4. Вспомогательные леммы

Нам понадобится следующая очевидная

Лемма 2. *Пусть* K *— функция ограниченной вариации. Тогда*

$$K\left(\frac{t-X_i}{h_n}\right) = \int\limits_{-\infty}^{\infty} \boldsymbol{I}(X_i \le t - vh_n)\,dK(v)$$

$$= \int\limits_{-\infty}^{\infty} \boldsymbol{I}\big(F^{-1}(\xi_i) \le t - vh_n\big)\,dK(v) = \int\limits_{-\infty}^{\infty} \boldsymbol{I}(\xi_i \le F(t-vh_n))\,dK(v).$$

Лемма 3. *Пусть K — функция ограниченной вариации. Тогда справедливо следующее равенство:*

$$f_n(t) - \boldsymbol{E} f_n(t) = \frac{1}{\sqrt{n} h_n} \int\limits_{-\infty}^{\infty} \alpha_n(F(t - vh_n)) dK(v). \tag{13}$$

Доказательство. Из леммы 2 следует, что имеет место следующая цепочка равенств

$$\mathbf{E} f_n(t) = \frac{1}{nh_n} \sum_{j=1}^{n} \int\limits_{-\infty}^{\infty} \mathbf{P}\left(X_i \leq t - vh_n\right) dK(v)$$

$$= \frac{1}{nh_n} \sum_{j=1}^{n} \int\limits_{-\infty}^{\infty} F(t - vh_n) dK(v) = \frac{1}{h_n} \int\limits_{-\infty}^{\infty} F(t - vh_n) dK(v).$$

Тогда получим

$$f_n(t) = \frac{1}{nh_n} \sum_{j=1}^{n} \int\limits_{-\infty}^{\infty} \mathbf{I}\left(X_i \leq t - vh_n\right) dK(v)$$

$$= \frac{1}{nh_n} \sum_{j=1}^{n} \int\limits_{-\infty}^{\infty} \mathbf{I}\left(\xi_i \leq F(t - vh_n)\right) dK(v)$$

$$= \frac{1}{nh_n} \sum_{j=1}^{n} \int\limits_{-\infty}^{\infty} \mathbf{I}_{[0, F(t - vh_n)]}(\xi_i) dK(v).$$

В итоге

$$f_n(t) - \mathbf{E} f_n(t) = \frac{1}{nh_n} \sum_{j=1}^{n} \int\limits_{-\infty}^{\infty} (\mathbf{I}_{[0, F(t - vh_n)]}(\xi_i) - F(t - vh_n)) dK(v)$$

$$= \frac{1}{\sqrt{n} h_n} \int\limits_{-\infty}^{\infty} \alpha_n(F(t - vh_n)) dK(v).$$

□

Напомним, что

$$\gamma_{n0}(t) = \frac{1}{\sqrt{n} h_n} \int\limits_{-\infty}^{\infty} W_n(F(t - vh_n)) dK(v), \quad \gamma_n(t) = \frac{1}{\sqrt{n} h_n} \int\limits_{-\infty}^{\infty} B_n(F(t - vh_n)) dK(v),$$

$$\delta_{0n}(t) = f_n(t) - \mathbf{E} f_n(t) - \gamma_{n0}(t), \quad \delta_n(t) = f_n(t) - \mathbf{E} f_n(t) - \gamma_n(t).$$

Лемма 4. *Имеет место следующее неравенство*

$$\left\| \overline{\psi}_n \delta_{0n} \right\|_\infty \leq \frac{C_n}{\sqrt{n} h_n} \left\| \alpha_n - W_n \right\|_\infty \cdot varK. \tag{14}$$

Доказательство. Учитывая результат леммы 3 очевидно имеют место следующие равенства:

$$\delta_{0n}(t) = f_n(t) - \mathbf{E}f_n(t) - \frac{1}{\sqrt{n}h_n}\int\limits_{-\infty}^{\infty} W_n(F(t-vh_n))dK(v)$$

$$= \frac{1}{\sqrt{n}h_n}\int\limits_{-\infty}^{\infty} \left(\alpha_n(F(t-vh_n)) - W_n(F(t-vh_n))\right) dK(v).$$

Отсюда следует, что

$$|\delta_{0n}(t)| \le \frac{1}{\sqrt{n}h_n}\int\limits_{-\infty}^{\infty} |\alpha_n(F(t-vh_n)) - W_n(F(t-vh_n))|\,|dK(v)|$$

$$\le \frac{1}{\sqrt{n}h_n}\|\alpha_n - W_n\|_\infty \int\limits_{-\infty}^{\infty} |dK(v)| = \frac{varK}{\sqrt{n}h_n}\|\alpha_n - W_n\|_\infty .$$

В итоге:

$$|\delta_{0n}(t)| \le \frac{varK}{\sqrt{n}h_n}\|\alpha_n - W_n\|_\infty, \qquad \forall t \in \mathbf{R}.$$

А поскольку $\overline{\psi}_n \le C_n$, то

$$\left\|\overline{\psi}_n\delta_{0n}\right\|_\infty \le \frac{C_n}{\sqrt{n}h_n}\|\alpha_n - W_n\|_\infty \cdot varK.$$

Тем самым мы доказали формулу (14). □

Отметим, что поскольку K — это функция ограниченной вариации, то

(15) $$\kappa := \sup_{t\in\mathbf{R}}|K(t)| \le varK/2 < \infty.$$

Лемма 5. *Пусть K - функция ограниченной вариации с носителем на отрезке $[-\frac{1}{2},\frac{1}{2}]$. Тогда для любого множества $T \subset \boldsymbol{R}$*

(16) $$\sup_{t\in T}|\psi(t)\boldsymbol{E}f_n(t)| \le \kappa\cdot D_n(T),$$

(17) $$b_n\sqrt{h_n}\cdot\sup_{t\in\boldsymbol{R}}|\psi(t)\boldsymbol{E}f_n(t)| \le \kappa\cdot\overline{D}_n.$$

Доказательство. Для начала заметим, что

$$\mathbf{E}f_n(t) = \frac{1}{h_n}\int\limits_{-\infty}^{\infty} K\left(\frac{t-x}{h_n}\right)dF(x) = \frac{1}{h_n}\int\limits_{t-h_n/2}^{t+h_n/2} K\left(\frac{t-x}{h_n}\right)dF(x).$$

Тогда при всех t

$$|\mathbf{E}f_n(t)| \le \frac{1}{h_n}\int\limits_{t-h_n/2}^{t+h_n/2}\left|K\left(\frac{t-x}{h_n}\right)\right|dF(x) \le \kappa\cdot\frac{F(t+h_n/2)-F(t-h_n/2)}{h_n}.$$

Чтобы из последнего неравенства извлечь (16) и (17), достаточно воспользоваться определениями (3) и (7) величин $\overline{D}_n$ и $D_n(T)$. □

Лемма 6. *Пусть выполнены условия леммы* 5. *Тогда имеет место следующее неравенство:*

$$\left\|\overline{\psi}_n \delta_n\right\|_\infty \le \left\|\overline{\psi}_n \delta_{0n}\right\|_\infty + \frac{\kappa \cdot D_n(\boldsymbol{R})}{\sqrt{n}} \cdot |\eta_n|\,. \tag{18}$$

Доказательство. Так как $B_n(t) = W_n(t) + t\eta_n$, то справедливо следующее равенство:

$$\delta_n(t) = \delta_{0n}(t) - \nu_n(t),$$

где

$$\nu_n(t) = \frac{1}{\sqrt{n}h_n} \int\limits_{-\infty}^{\infty} F(t - vh_n) dK(v) \cdot \eta_n.$$

В силу справедливости (16) и равенства

$$\frac{1}{\sqrt{n}h_n} \int\limits_{-\infty}^{\infty} F(t - vh_n) dK(v) = \frac{1}{\sqrt{n}} \mathbf{E} f_n(t)$$

получим, что

$$\left\|\overline{\psi}_n \cdot \nu_n\right\|_\infty \le \frac{\kappa \cdot D_n(\mathbf{R})}{\sqrt{n}} \cdot |\eta_n|.$$

Тогда

$$\left\|\overline{\psi}_n \delta_n\right\|_\infty \le \left\|\overline{\psi}_n \delta_{0n}\right\|_\infty + \frac{\kappa \cdot D_n(\mathbf{R})}{\sqrt{n}} \cdot |\eta_n|\,.$$

□

5. Доказательства основных результатов

Доказательство теоремы 1. Пусть

$$B_n = \left\{ \|\alpha_n - W_n\|_\infty \le \frac{k_3 \log n}{\sqrt{n}} \right\}.$$

Тогда при любом $\varepsilon > 0$

$$\begin{aligned} \mathbf{P}(\Delta_{0n}(\mathbf{R}) > \varepsilon) &= \mathbf{P}(\Delta_{0n}(\mathbf{R}) > \varepsilon, B_n) + \mathbf{P}(\Delta_{0n}(\mathbf{R}) > \varepsilon, \overline{B}_n) \\ &\le \mathbf{P}(b_n \sqrt{nh_n} C_n \frac{varK}{\sqrt{n}h_n} \cdot \|\alpha_n - W_n\|_\infty > \varepsilon, B_n) + \mathbf{P}(\overline{B}_n)e \\ &\le \mathbf{P}(\rho_n > \varepsilon) + \mathbf{P}(\overline{B}_n) \quad \text{при} \quad \rho_n = k_3 \cdot C_n b_n \log n \frac{varK}{\sqrt{nh_n}}. \end{aligned}$$

Здесь $\mathbf{P}(\overline{B}_n) \longrightarrow 0$ при $n \to \infty$ ввиду свойства (12) KMT аппроксимации. А так как $C_n = o\left(\frac{\sqrt{nh_n}}{b_n \log n}\right)$, то $\rho_n \longrightarrow 0$ и $\mathbf{P}(\rho_n > \varepsilon) = 0$ начиная с некоторого номера $n = n(\varepsilon) < \infty$.

Окончательно получаем, что

$$\forall \varepsilon > 0 \quad \mathbf{P}(\Delta_{0n}(\mathbf{R}) > \varepsilon) \to 0 \quad \text{при} \quad n \to \infty.$$

□

Доказательство теоремы 2. Из леммы 6, с учетом определений из (2) и (4) имеем

$$\Delta_n(\mathbf{R}) \le \Delta_{0n}(\mathbf{R}) + \kappa \overline{D}_n \,|\eta_n|\,. \tag{19}$$

Поскольку $\overline{D}_n \longrightarrow 0$ в силу (3), а все величины $\{\eta_n\}$ — одинаково распределены, то

$$\forall\varepsilon > 0 \quad \mathbf{P}(\overline{D}_n\,|\eta_n| > \varepsilon) = \mathbf{P}(\overline{D}_n\,|\eta_1| > \varepsilon) \longrightarrow 0.$$

Значит $\overline{D}_n\,|\eta_n| \xrightarrow{\mathbf{P}} 0$ и $\Delta_{0n}(\mathbf{R}) + \kappa\overline{D}_n\,|\eta_n| \xrightarrow{\mathbf{P}} 0$ в силу (2).
Из последней сходимости и (19) вытекает требуемая сходимость (4). □

Доказательство следствия 1. Величины

$$\mu_n = b_n(\beta_n(T_n) - a_n), \quad \widetilde{\mu}_n = b_n(\widetilde{\beta}_n(T_n) - a_n)$$

если и имеют предел, то один и тот же. Этот факт вытекает из следующего соотношения:

$$|\widetilde{\mu}_n - \mu_n| \le \Delta_n(T_n) \le \Delta_n(\mathbf{R}) \xrightarrow{\mathbf{P}} 0.$$

Здесь мы использовали при

$$x(t) = \sqrt{nh_n}\psi(t)(f_n(t) - \mathbf{E}f_n(t)), \quad y(t) = \sqrt{nh_n}\psi(t)\gamma_n(t)$$

тот факт, что для любых функций и множеств

$$\left|\sup_{t\in T}|x(t)| - \sup_{t\in T}|y(t)|\right| \le \sup_{t\in T}|x(t) - y(t)|\,.$$

□

Для доказательства теоремы 3 нам понадобится

Лемма 7. *Если выполнены условия теоремы* 3*, то*

$$\frac{\beta_n(\overline{T}_n)}{a_n} \xrightarrow{\boldsymbol{P}} 0. \tag{20}$$

Доказательство. Для любого $\varepsilon > 0$ имеет место

$$\mathbf{P}\left(\frac{\beta_n(\overline{T}_n)}{a_n} > \varepsilon\right) \le Q_n(\varepsilon) + \overline{Q}_n(\varepsilon), \quad \text{где}$$

$$Q_n(\varepsilon) = \mathbf{P}\left(\frac{\beta_n(\overline{T}_n)}{a_n} > \varepsilon, \max_{i\le n}\psi_{h_n}(X_i) \le C_n\right),$$

$$\overline{Q}_n(\varepsilon) = \mathbf{P}\left(\frac{\beta_n(\overline{T}_n)}{a_n} > \varepsilon, \max_{i\le n}\psi_{h_n}(X_i) > C_n\right)$$

$$\le \sum_{i=1}^{n}\mathbf{P}(\psi_{h_n}(X_i) > C_n) = n\mathbf{P}(\psi_{h_n}(X_1) > C_n) \longrightarrow 0.$$

Последняя сходимость справедлива в силу условия (8).

Заметим теперь, что

$$C_n < \psi(t) \le \psi_{h_n}(t) \quad \text{при} \quad t \in \overline{T}_n.$$

Следовательно, если $\max\limits_{i\le n}\psi_{h_n}(X_i) \le C_n$, то $f_n(t) = 0$ при $t \in \overline{T}_n$. Значит, в этом случае имеем:

$$|\psi(t)(f_n(t) - \mathbf{E}f_n(t))| = |\psi(t)\mathbf{E}f_n(t)| \le \kappa D_n(\overline{T}_n)$$

в силу леммы 5. Значит из условия (9) вытекает, что

$$\frac{\beta_n(\overline{T}_n)}{a_n} \le \kappa\frac{\sqrt{nh_n}D_n(\overline{T}_n)}{a_n} \longrightarrow 0 \quad \text{при} \quad \max_{i\le n}\psi_{h_n}(X_i) \le C_n.$$

Но отсюда немедленно следует, что

$$Q_n(\varepsilon) \le \mathbf{P}\left(\kappa\frac{\sqrt{nh_n}D_n(\overline{T}_n)}{a_n} > \varepsilon, \max_{i\le n}\psi_{h_n}(X_i) \le C_n\right)$$
$$\le \mathbf{P}\left(\frac{\sqrt{nh_n}D_n(\overline{T}_n)}{a_n} > \frac{\varepsilon}{\kappa}\right) \longrightarrow 0.$$

□

Доказательство теоремы 3. В силу следствия 1 осталось доказать, что

$$\mathbf{P}\left(\beta_n(T_n) \ne \beta_n(\mathbf{R})\right) \longrightarrow 0, \text{ где } \beta_n(\mathbf{R}) = \max\{\beta_n(T_n), \beta_n(\overline{T}_n)\}. \tag{21}$$

Причем нам нужно показать, что (21) имеет место как при выполнении (10), так и при справедливости (11).

Пусть сначала выполнено (10). В этом случае при любом $\varepsilon > 0$

$$\mathbf{P}\left(\left|\frac{\beta_n(\mathbf{R})}{a_n} - 1\right| \le \varepsilon\right) = \mathbf{P}(-a_nb_n\varepsilon \le b_n(\beta_n(\mathbf{R}) - a_n) \le a_nb_n\varepsilon),$$
$$\longrightarrow \mathbf{P}(-\infty < \zeta < \infty) = 1,$$

так как $\varepsilon a_nb_n \longrightarrow \infty$ в силу соответствующего условия теоремы 3. Значит $\frac{\beta_n(\mathbf{R})}{a_n} \xrightarrow{\mathbf{P}} 1$. Но из этого факта, с учетом (20), имеем:

$$\mathbf{P}\left(\beta_n(T_n) \ne \beta_n(\mathbf{R})\right) = \mathbf{P}\left(\frac{\beta_n(\overline{T}_n)}{a_n} = \frac{\beta_n(\mathbf{R})}{a_n}\right) \longrightarrow \mathbf{P}(0 = 1) = 0,$$

что доказывает (21) при выполнении (10).

Пусть теперь верно (11). В этом случае, ввиду следствия 1, справедливо также (6). Значит при всех $\varepsilon > 0$

$$\mathbf{P}\left(\left|\frac{\beta_n(T_n)}{a_n} - 1\right| \le \varepsilon\right) = \mathbf{P}(-a_nb_n\varepsilon \le b_n(\beta_n(T_n) - a_n) \le a_nb_n\varepsilon),$$
$$\longrightarrow \mathbf{P}(-\infty < \zeta < \infty) = 1,$$

поскольку $\varepsilon a_nb_n \longrightarrow \infty$. Следовательно, $\frac{\beta_n(T_n)}{a_n} \xrightarrow{\mathbf{P}} 1$. Но отсюда, с учетом (20), находим

$$\mathbf{P}\left(\beta_n(T_n) \ne \beta_n(\mathbf{R})\right) = \mathbf{P}\left(\frac{\beta_n(T_n)}{a_n} < \frac{\beta_n(\overline{T}_n)}{a_n}\right) \longrightarrow \mathbf{P}(1 < 0) = 0.$$

Значит сходимость (21) установлена и при выполнении (11). □

Таким образом, все утверждения из п.2 доказаны полностью.

References

[1] E. Gine, V. Koltchinskii, L. Sakhanenko, *Kernel density estimators: convergence in distribution for weighted sup-norms*, Probability Theory and Related Fields, **130**:2 (2004), 167–198. MR2093761

[2] E. Gine, V. Koltchinskii, J. Zinn, *Weighted uniform consistency of kernel density estimators*, Annals of Probability, **32**:3B (2004), 2570—2605. MR2078551

[3] L. Sakhanenko, *Asymptotics of suprema of the weighed Gaussian fields with applications to kernel density estimators*, Theory of Probability and its Applications, **59**:3 (2015), 415–451. Zbl 06477542

[4] J. Komlos, P. Major, G. Tusnady, *An Approximation of Partial Sums of Independent RV'-s, and the Sample DF. I*, Z.Wahrsch. verw. Gebiete **32** (1975), 111—131. MR0375412

[5] A.V. Skorokhod, *On a Representation of Random Variables*, Theory of Probability and its Applications, **21**:3 (1977), 628–632. Zbl 0362.60004

Andrey Sergeevich Kartashov
Novosibirsk State University,
st. Pirogova, 2,
630090, Novosibirsk, Russia
E-mail address: andrewkartashov@mail.ru

Alexander Ivanovich Sakhanenko
Sobolev Institute of Mathematics,
pr. Koptyuga, 4,
630090, Novosibirsk, Russia
E-mail address: aisakh@mail.ru