©2005 г. С. ЖИРАР, д-р наук,
А. ЮДИЦКИЙ, д-р наук, профессор
(Университет Гренобль I, Франция),
А. В. НАЗИН[1], д-р физ.-мат. наук
(Институт проблем управления им. В. А. Трапезникова РАН, Москва)

L_1-ОПТИМАЛЬНОЕ НЕПАРАМЕТРИЧЕСКОЕ ОЦЕНИВАНИЕ ГРАНИЦЫ НОСИТЕЛЯ ПОСРЕДСТВОМ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Предложен метод оценивания границы множества точек на плоскости, оптимальный в L_1-норме на заданном классе β-липшицевых граничных функций, при $\beta \in (0,1]$. Оценка определяется как достаточно регулярная линейная комбинация ядерных функций, центрированных в точках выборки, которая покрывает все эти точки и порождает носитель минимальной площади. Веса линейной комбинации вычисляются посредством решения соответствующей задачи линейного программирования. Показано, что L_1-норма ошибки оценивания сходится к нулю с вероятностью 1 с оптимальной скоростью.

1. Введение

Задачи и методы оценивания множества $S \subset \mathbb{R}^2$ по конечной случайной выборке его точек неоднократно рассматривались в литературе. Так, проблема оценивания границы или носителя возникает при классификации [1], в кластерном и дискриминантном анализе [2, 3] и при обнаружении больших выбросов [4]. Приложения можно также найти в медицинской диагностике [5] и в области мониторинга машин [4]. При анализе изображений задача сегментации может рассматриваться с позиции оценки носителя — выпуклого ограниченного множества в $\mathbb{R}^2$ [6]. Известны также применения в эконометрике, (см., например, [7]). В этих и во всех других случаях, когда существует возможность описания неизвестного носителя в виде

$$S \triangleq \{(x,y) \;:\; 0 \le x \le 1,\; 0 \le y \le f(x)\}, \tag{1}$$

где $f : [0,1] \to (0,+\infty)$ — неизвестная ограниченная функция, задача сводится к оцениванию граничной функции f (см., например, [8]). Данные состоят из пар (X,Y), где X представляет собой, например, производственные затраты (количество затраченного труда, энергии или капитала), порождающие доход Y данной фирмы. При таком рассмотрении значение $f(x)$ может интерпретироваться как максимальный уровень дохода, достигаемого при затратах x.

[1] Исследования проведены во время визитов в LMC-IMAG (университет Гренобль–I) и INRIA Rhône-Alpes, проект MISTIS, Франция, в 2004 г.

Одна из первых статей по этой проблеме [9] посвящена независимым одинаково распределенным наблюдениям с плотностью ϕ. Предложенная в ней оценка гистограммного типа также основана на экстремальных значениях выборки. Впоследствии эта оценка была обобщена в двух основных направлениях.

С одной стороны, были предложены кусочно-полиномиальные оценки, которые определяются локально, на данном участке, как наименьший полином фиксированной степени, покрывающий все точки этого участка. Их оптимальность в асимптотически минимаксном смысле доказана в [6, 10] при слабых предположениях о скорости убывания α плотности ϕ к нулю. Методы экстремальных значений предлагались впоследствии в [11, 12] для оценки параметра α. Отметим, что оценивание f также может рассматриваться как регрессионная задача с $Y = f(X) + \varepsilon$ при отрицательном шуме ε. В таком контексте также предлагались локально-полиномиальные оценки, см., например, [13, 14].

С другой стороны, были предложены различные способы сглаживания оценки Жеффруа при рассмотрении точечного пуассоновского процесса. В [15] изучались оценки, основанные на ядерных регрессиях и на методе ортогональных разложений [16, 17]. Аналогично в [18] предлагалось использовать оценку Фабера–Шаудера. В [19] описан общий подход к изучению методов подобного типа и их распространию на множества вида

$$S = \{(x, y) : \ \ x \in E \ \ ; \ \ \ 0 \leq \ y \ \leq \ f(x)\},$$

где f — неизвестная функция, а E — произвольное заданное множество.

Отметим, что во всех этих работах было установлено предельное распределение оценок. Следует также отметить [20, 21], где развивался аналогичный подход для пуассоновского процесса без использования экстремальных значений.

Оценка, предложенная в [22, 23], может рассматриваться как попытка одновременного использования идей этих двух направлений. С практической точки зрения она определяется как ядерная оценка, получаемая сглаживанием ряда отобранных точек выборки. Эти точки выбираются автоматически посредством решения некоторой задачи линейного программирования с целью получить оценку носителя, покрывающую все точки выборки и выделяющую множество наименьшей площади. С теоретической точки зрения, доказывается состоятельность этой оценки в L_1-норме.

В настоящей статье (см. также [24]) оценка из [22, 23] существенно модифицируется. Во-первых, используется скорректированное ядро, обеспечивающее нулевое среднее для смещения на всем интервале оценивания. Во-вторых, в задачу оптимизации вводятся дополнительные ограничения, гарантирующие необходимую регулярность оценки. В результате получающаяся оценка достигает оптимальную в L_1-норме минимаксную скорость сходимости (с точностью до логарифмического множителя). Отметим также, что здесь рассматривается более общий по сравнению с [22, 23] класс липшицевых функций, зависящий не только от константы Липшица L, но и от параметра $\beta \in (0, 1]$; при $\beta = 1$ получаем случай, изученный в [22, 23].

Структура статьи следующая: задача и метод оценивания описаны в разделе 2, а некоторые предварительные свойства оценки сформулированы в разделе 3 и доказаны в Приложении; в последующем разделе описаны основные результаты, доказательство которых также изложено в Приложении.

2. Задача и метод оценивания границы

Пусть все случайные величины определены на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Рассмотрим задачу оценивания неизвестной положительной функции $f : [0,1] \to (0, \infty)$ по наблюдениям $Z_N = (X_i, Y_i)_{i=1,\dots,N}$, которые представляют собой последовательность независимых равномерно распределенных на S пар (X_i, Y_i), где множество S имеет вид (1). Для простоты, причем без ограничения общности, далее будем рассматривать функции f, определенные на всей числовой оси $\mathbb{R}$, полагая $f(x) = 0$ для всех $x \notin [0,1]$. Введем функционал

$$C_f \triangleq \int_0^1 f(u)\, du = \int_{\mathbb{R}} f(u)\, du\,. \tag{2}$$

Таким образом, случайные величины X_i распределены на отрезке $[0,1]$ с плотностью $f(\cdot)/C_f$, а Y_i имеют равномерное условное распределение относительно X_i на отрезке $[0, f(X_i)]$.

Будем далее предполагать, что $f \in \Sigma_{[0,1]}(\beta, L_{f,\beta})$, $0 < \beta \le 1$, т.е. функция $f : [0,1] \to (0, \infty)$ является β-липшицевой с константой $L_{f,\beta}$:

$$|f(x) - f(u)| \le L_{f,\beta}\, |x-u|^\beta \quad \forall x, u \in [0,1]\,. \tag{3}$$

Рассматриваемая далее оценка выбирается из следующего семейства функций:

$$\begin{cases} \widehat{f}_N(x) = \displaystyle\sum_{i=1}^N \alpha_i\, K_h(x, X_i)\,, & K_h(x,t) = \dfrac{g(x)}{h}\, K\left(\dfrac{x-t}{h}\right), \\ \alpha_i \ge 0, \qquad i = 1, \dots, N, & \end{cases} \tag{4}$$

где $K(\cdot)$ — достаточно гладкая основная ядерная функция $K : \mathbb{R} \to [0, +\infty)$, удовлетворяющая условию нормировки

$$\int_{\mathbb{R}} K(u)\, du = 1$$

и имеющая конечный носитель $[-1, 1]$; параметр ширины окна $h \in (0, 1/2)$ зависит от размера выборки N так, что $h \to 0$ при $N \to \infty$; функция

$$g(x) = \left(\int_{(x-1)/h}^{x/h} K(t)\, dt \right)^{-1}, \quad x \in [0,1]\,, \tag{5}$$

корректирует основное ядро K вблизи краев интервала $[0,1]$, т.е. при $x \in [0, h)$ или $x \in (1-h, 1]$. Действительно, $g(x) \equiv 1$ при $x \in [h, 1-h]$, и $g(x) > 1$ при $x \in [0, h)$ и/или при $x \in (1-h, 1]$.

Нетрудно убедиться, что

$$\int_0^1 K_h(x, u)\, du = 1 \qquad \forall\, x \in [0,1] \tag{6}$$

и, вследствие возможности поменять очередность интегрирования и дифференцирования,

$$\int_0^1 \frac{\partial}{\partial x} K_h(x, u)\, du = 0 \qquad \forall\, x \in [0,1]\,. \tag{7}$$

Заметим, что равенство (7) можно проверить и непосредственно, см. [24].

Обозначим $K_{\max} \triangleq \max K(t)$, $g_{\max} \triangleq \sup\limits_{x,h} g(x)$ и введем функционалы

$$C_\beta(\varphi) \triangleq \int_{-1}^{1} |t|^\beta |\varphi(t)|\, dt\,, \quad \varphi \in C^0([-1,1])\,, \tag{8}$$

$$C_\beta(K,K') \triangleq g_{\max} K_{\max} C_\beta(K) + C_\beta(K')\,. \tag{9}$$

Через L_φ обозначим константу Липшица функции $\varphi : \mathbb{R} \to \mathbb{R}$, т.е.

$$|\varphi(s) - \varphi(t)| \le L_\varphi |s-t|\,, \tag{10}$$

причем $L_\varphi < \infty$. Индикатор обозначаем через $\mathbf{1}\{\cdot\}$; его значение равно 1, если аргументное условие верно, и 0 в противном случае.

Как показано ниже в лемме 1, площадь оценки носителя

$$\widehat{S}_N \triangleq \{(x,y) :\ 0 \le x \le 1\,,\ 0 \le y \le \widehat{f}_N(x)\} \tag{11}$$

может быть аппроксимирована следующим образом:

$$\int_0^1 \widehat{f}_N(x)\, dx = \sum_{i=1}^{N} \alpha_i + O(h)\,. \tag{12}$$

Отсюда естественно вытекает определение вектора параметров $\alpha = (\alpha_1, \dots, \alpha_N)^{\mathrm{T}}$ как решение следующей задачи оптимизации:

$$J_P^* \triangleq \min_\alpha \sum_{i=1}^{N} \alpha_i \tag{13}$$

при ограничениях

$$\widehat{f}_N(X_i) \ge Y_i\,, \quad i = 1, \dots, N\,, \tag{14}$$

$$|\widehat{f}_N'(X_i)| \le L_{f,\beta}\, g_{\max} C_\beta(K,K') \frac{\log N}{Nh^2}\,, \quad i = 0, \dots, N+1\,, \tag{15}$$

$$C_\alpha h \ge \sum_{i=1}^{N} \alpha_i\, \mathbf{1}\{(j-1)/m_h \le X_i < j/m_h\}\,, \quad j = 1, \dots, m_h\,, \tag{16}$$

$$0 \le \alpha_i\,, \quad i = 1, \dots, N, \tag{17}$$

где параметр $m_h = \lfloor h^{-1} \rfloor$ — целая часть числа $1/h$. Эта задача оптимизации может формально быть записана как задача линейного программирования (ЛП)

$$J_P^* \triangleq \min_\alpha \mathbf{1}_N^{\mathrm{T}} \alpha \tag{18}$$

при ограничениях

$$Y \le A\alpha\,, \tag{19}$$

$$\mathrm{abs}\,(B\alpha) \le L_{f,\beta}\, g_{\max} C_\beta(K,K')\, \frac{\log N}{Nh^2}\, \mathbf{1}_N\,, \tag{20}$$

$$D^{\mathrm{T}}\alpha \le C_\alpha h\, \mathbf{1}_{m_h}\,, \tag{21}$$

$$0 \le \alpha\,. \tag{22}$$

Здесь неравенства между векторами и функция $\mathrm{abs}(\cdot)$ от векторного аргумента понимаются покомпонентно, причем i-я компонента вектора $\mathrm{abs}(x)$ равна $|x_i|$. Ограничения (16) и (21) содержат один положительный параметр C_α, значение которого обсуждается ниже в разделе 4. Кроме того, нами используются следующие обозначения:

$$X_0 \triangleq 0, \quad X_{N+1} \triangleq 1, \tag{23}$$

$$\mathbf{1}_N \triangleq (1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^N, \tag{24}$$

$$A \triangleq \|K_h(X_i\,, X_j)\|_{i,j=1,\ldots,N}\ , \tag{25}$$

$$B \triangleq \left\| \frac{d}{dx} K_h(x, X_j)\Big|_{x=X_i} \right\|_{i,j=1,\ldots,N}\ , \tag{26}$$

$$D \triangleq \|\mathbf{1}\{(j-1)/m_h \le X_i < j/m_h\}\|_{i=1,\ldots,N;\, j=1,\ldots,m_h}\ , \tag{27}$$

$$Y \triangleq (Y_1, \ldots, Y_N)^{\mathrm{T}}. \tag{28}$$

Суть ограничений (14)–(17) и/или (19)–(22) проясняется в следующем разделе после введения ряда предположений и установления предварительных результатов.

3. Предположения и предварительные результаты

Основные предположения относительно граничной функции следующие:

A1. $0 < f_{\min} \le f(x) \le f_{\max} < \infty$ для всех $x \in [0, 1]$,

A2. $|f(x) - f(y)| \le L_{f,\beta}\,|x-y|^\beta$ для всех $x, y \in [0, 1]$, где постоянные $L_{f,\beta} < \infty$ и $\beta \in (0, 1]$ фиксированы.

Введем следующие предположения относительно основной ядерной функции:

B1. ядро $K : \mathbb{R} \to [0, \infty)$ имеет конечный носитель, а именно: $\mathrm{supp}_{t\in\mathbb{R}}\, K(t) = [-1, 1]$,

B2. $\displaystyle\int_{-1}^{1} K(t)\,dt = 1$,

B3. функция K трижды непрерывно дифференцируема.

Заметим, что $g_{\max} = \left(\min\left\{\int\limits_0^1 K(t)\,dt,\ \int\limits_{-1}^0 K(t)\,dt\right\}\right)^{-1}$; в частности, $g_{\max} = 2$ для любого четного ядра $K(\cdot)$, удовлетворяющего условиям B1–B2. Приведем два предварительных результата относительно оценки $\widehat{f}_N$. Первый — об аппроксимации (12) площади $\widehat{S}_N$ (11), а второй — о липшицевости функции $\widehat{f}_N$. Их доказательства приводятся в Приложении.

Л е м м а 1. *Допустим, что выполнены условия B1, B2 и $0 < h < 1/4$. Кроме того, пусть справедливы условия (16) и (17) с $m_h = \lfloor h^{-1} \rfloor$. Тогда площадь оцененного носителя (11) удовлетворяет следующим неравенствам:*

$$-2C_\alpha K_{\max} h \le \widehat{S}_N - \sum_{i=1}^{N} \alpha_i \le 4C_\alpha (g_{\max} - 1) K_{\max} h\,. \tag{29}$$

З а м е ч а н и е 1. *Ниже фактически используется только одна часть леммы 1, определяющая верхнюю границу площади $\widehat{S}_N$, т.е. правая часть неравенства (29).*

З а м е ч а н и е 2. *Лемма 1 и все последующие результаты могут быть легко распространены на основные ядра $K(\cdot)$, имеющие также и отрицательные значения: тогда $K_{\max} \triangleq \max|K(t)|$ и дополнительно следует предположить $g(x) > 0\ \forall\, x \in [0,1]$.*

Л е м м а 2. *Предположим,что выполнены условия A1 и B1–B3. Пусть оценка $\widehat{f}_N$ определяется решением ЛП-задачи (18)–(22). Кроме того, пусть $h \to 0$ при $N \to \infty$ так, что*

$$\lim_{N\to\infty} \frac{\log N}{Nh} = 0\,. \tag{30}$$

Тогда почти наверное существует конечное $N_4 = N_4(\omega)$ такое, что для любых $N \geq N_4$ константа Липшица оценки $\widehat{f}_N$ на интервале $[0,1]$ ограничена, а именно:

$$L_{\widehat{f}_N} \triangleq \max_{x\in[0,1]} |\widehat{f}_N'(x)| \tag{31}$$

$$\leq 2L_{f,\beta}\, g_{\max} C_\beta(K,K') \frac{\log N}{Nh^2}\,. \tag{32}$$

З а м е ч а н и е 3. *Как можно увидеть из доказательства леммы 2, а именно из (П.11)–(П.12), можно было бы несколько уменьшить количество ограничений (15) на производную оценки (13)–(17). Фактически, можно было бы наложить ограничения того же вида не во всех точках X_i, $i = 1, \ldots, N$: было бы достаточно взять для этого точки, отстоящие каждая от своих соседних на расстояние $O\left((h\log N/N)^{1/2}\right)$ или по крайней мере $o\left((h\log N/N)^{1/2}\right)$; при этом константа Липшица для $\widehat{f}_N$ сохранится той же, что дана в лемме 2.*

Таким образом, оценка $\widehat{f}_N$, определяемая решением задачи оптимизации (13)–(17) или ее эквивалентной ЛП-версией (18)–(22), представляет собой ядерную оценку носителя, покрывающего все точки $(X_i, Y_i)_{i=1,\ldots,N}$ и приблизительно имеющего наименьшую площадь, с точностью до члена $O(h)$, указанного в лемме 1. Кроме того, ограничения (15)–(16) или (20)–(21) обеспечивают $\widehat{f}_N \in \Sigma_{[0,1]}(1, L_{\widehat{f}_N})$ с константой Липшица $L_{\widehat{f}_N}$, указанной в лемме 2. Ограничение $\alpha_i \geq 0$ для всех $i = 1, \ldots, N$ обеспечивает $\widehat{f}_N(x) \geq 0$ для всех $x \in [0,1]$, поскольку ядерная функция K неотрицательна; это представляется естественным, так как функция $f(\cdot)$ положительна. Отметим также, что описанная выше оценка (4), (18)–(22) может рассматриваться как аппроксимация оценки максимального правдоподобия, соответствующей семейству (4); подробнее см. [23].

Следующие две леммы дают верхнюю границу функционала J_P^* и поточечную нижнюю границу самой оценки $\widehat{f}_N$.

Л е м м а 3. *Пусть справедливы предположения теоремы 1. Тогда для всякого конечного*

$$\gamma > L_{f,\beta}\, g_{\max} C_\beta(K) \tag{33}$$

и почти всех $\omega \in \Omega$ существуют конечные числа $N_1 = N_1(\omega, \gamma)$ такие, что для всех $N \geq N_1$ ЛП-задача (18)–(22) разрешима и

$$J_P^* \leq C_f + \gamma h^\beta \,. \tag{34}$$

Л е м м а 4. *В условиях теоремы 1 для почти всех $\omega \in \Omega$ существуют конечные числа $N_2(\omega)$ такие, что для любого $x \in [0,1]$ и произвольного $N \geq N_2(\omega)$*

$$\widehat{f}_N(x) \geq f(x) - \frac{C_4(\beta)}{h^2} \left(\frac{\log N}{N} \right)^{\frac{2+\beta}{1+\beta}} \tag{35}$$

с константой $C_4(\beta)$, определенной соотношением (39).

4. Основные результаты

Следующая теорема устанавливает состоятельность и скорость сходимости описанной выше оценки в смысле L_1-нормы ошибки на интервале оценивания $[0,1]$.

Т е о р е м а 1. *Пусть выполнены описанные выше предположения A1, A2 и B1–B3, а параметр оценки $C_\alpha > 6 f_{\max}$. Кроме того, пусть $h \to 0$ при $N \to \infty$ так, что*

$$\liminf_{N\to\infty} \frac{\log N}{N h^{1+\beta}} > \rho > 0 \,, \quad \lim_{N\to\infty} \frac{\log N}{N h^{1+\beta/2}} = 0 \,. \tag{36}$$

Тогда оценка $\widehat{f}_N$, определяемая соотношениями (4)–(5) и решением ЛП-задачи (18)–(22), имеет следующие асимптотические свойства:

$$\|\widehat{f}_N - f\|_1 \leq \left(C_{12}(\beta) h^\beta + 2 C_4(\beta) h^{-2} (\log N/N)^{\frac{2+\beta}{1+\beta}} \right) (1 + o(1)) \quad \textit{п.н.} \tag{37}$$

с константами

$$C_{12}(\beta) \triangleq 2 L_{f,\beta}\; g_{\max} C_\beta(K, K') + 4 C_\alpha (g_{\max} - 1) K_{\max} \mathbf{1}\{\beta = 1\} \tag{38}$$

и

$$C_4(\beta) \triangleq 2 L_{f,\beta} \left[\left(\frac{2C_f}{L_{f,\beta}} \right)^{\frac{\beta}{1+\beta}} \left(\frac{1}{\rho} \right)^{\frac{2}{1+\beta}} + g_{\max} C_\beta(K, K') \left(\frac{2C_f}{L_{f,\beta}} \right)^{\frac{1}{1+\beta}} \right] . \tag{39}$$

С л е д с т в и е 1. *Наибольшая скорость сходимости, гарантируемая теоремой 1,*

$$\|\widehat{f}_N - f\|_1 = O\left((\log N/N)^{\frac{\beta}{1+\beta}} \right) \quad \textit{п.н.},$$

достигается при ширине окна h, имеющей асимптотику

$$h \sim \widetilde{\rho} \left(\frac{\log N}{N} \right)^{\frac{1}{1+\beta}} , \quad 0 < \widetilde{\rho} < \rho^{-\frac{1}{1+\beta}} \,, \tag{40}$$

что совпадает с верхней границей (37), а именно:

$$\limsup_{N\to\infty}\left(\frac{\log N}{N}\right)^{-\frac{\beta}{1+\beta}}\|\widehat{f}_N - f\|_1 \le C_{12}(\beta)\widetilde{\rho}^{\beta} + 2C_4(\beta)\widetilde{\rho}^{-2} \quad \text{п.н.} \tag{41}$$

Подчеркнем, что (41) показывает, что $\widehat{f}_N$ достигает (с точностью до логарифмического множителя) оптимальную минимаксную точность в L_1-норме для β-липшицевых граничных функций f; см. [7, теорема 4.1.1].

З а м е ч а н и е 4. Второе условие в (36) может быть ослаблено до

$$\lim_{N\to\infty}\frac{\log N}{Nh^{1+\beta/2}} < \infty\,, \tag{42}$$

что ведет к другой, более общей формуле для константы в (37)–(39).

Приведенное в Приложении доказательство теоремы 1 основано как на верхней, так и на нижней границах, полученных в леммах 3 и 4 соответственно.

5. Заключение

Предложенный метод оценивания границы носителя порождает оценку как линейную комбинацию гладких ядерных функций, центрированных в точках выборки, которая покрывает все эти точки и порождает носитель минимальной площади. Метод по существу достигает L_1-оптимальную скорость сходимости $O\left((\log N/N)^{\beta/(1+\beta)}\right)$ на классе β-липшицевых граничных функций при выборе ширины окна h порядка $(\log N/N)^{1/(1+\beta)}$. Оптимальность стала возможной за счет усовершенствования метода из [22, 23] путем введения в соответствующую ЛП-задачу дополнительных линейных ограничений (15), (20) на регулярность оценки. Таким образом, с вычислительной точки зрения метод не претерпел существенных усложнений, так как по-прежнему сводится к задаче линейного программирования.

П Р И Л О Ж Е Н И Е

Д о к а з а т е л ь с т в о л е м м ы 1. Заметим, что определения (4)–(5) влекут следующее разложение:

$$\int_0^1 \widehat{f}_N(x)\,dx = \left(\int_0^h + \int_h^{1-h} + \int_{1-h}^1\right)\widehat{f}_N(x)\,dx = \tag{П.1}$$

$$= \int_0^1 \sum_{i=1}^N \alpha_i \frac{1}{h} K\left(\frac{x-X_i}{h}\right)dx + \tag{П.2}$$

$$+\sum_{i=1}^N \alpha_i \left(\int_0^h + \int_{1-h}^1\right)\frac{g(x)-1}{h}K\left(\frac{x-X_i}{h}\right)dx\,. \tag{П.3}$$

Поскольку все α_i и ядро K неотрицательны, получаем

$$\text{(П.4)}\qquad \int\limits_0^1 \sum_{i=1}^N \alpha_i \frac{1}{h} K\left(\frac{x-X_i}{h}\right) dx \le \sum_{i=1}^N \alpha_i \int\limits_{\mathbb{R}} \frac{1}{h} K\left(\frac{x-X_i}{h}\right) dx = \sum_{i=1}^N \alpha_i$$

и, следовательно,

$$\begin{aligned}
&\text{(П.5)} & \int\limits_0^1 \widehat{f}_N(x)\,dx - \sum_{i=1}^N \alpha_i &\le \frac{g_{\max}-1}{h} K_{\max} \sum_{i=1}^N \alpha_i \left(\int\limits_0^h + \int\limits_{1-h}^1\right) \mathbf{1}\{|x-X_i| \le h\} dx \le \\
&\text{(П.6)} & &\le (g_{\max}-1) K_{\max} \left(\sum_{i=1}^N \alpha_i\, \mathbf{1}\{0 \le X_i \le 2h\} + \right. \\
&\text{(П.7)} & &\qquad \left. + \sum_{i=1}^N \alpha_i\, \mathbf{1}\{1-2h \le X_i \le 1\}\right) \le \\
&\text{(П.8)} & &\le (g_{\max}-1) K_{\max} 4 C_\alpha h\,.
\end{aligned}$$

Неравенство (П.8) вытекает из (16), поскольку оба интервала в (П.6)–(П.7) имеют длину $2h$ и, следовательно, могут быть покрыты двумя соответствующими интервалами вида $[(j-1)/m_h, j/m_h]$ в (16). Следовательно, доказана верхняя граница для разности из левой части (П.5). Нижняя граница доказывается таким же образом. Действительно, из разложения (П.1)–(П.3) вытекает, благодаря неотрицательности члена (П.3),

$$\begin{aligned}
\int\limits_0^1 \widehat{f}_N(x)\,dx - \sum_{i=1}^N \alpha_i &\ge -\sum_{i=1}^N \alpha_i \left(\int\limits_{-h}^0 + \int\limits_1^{1+h}\right) \frac{1}{h} K\left(\frac{x-X_i}{h}\right) dx \ge \\
&\ge -K_{\max} \left(\sum_{i=1}^N \alpha_i\, \mathbf{1}\{0 \le X_i \le h\} + \right. \\
&\qquad \left. + \sum_{i=1}^N \alpha_i\, \mathbf{1}\{1-h \le X_i \le 1\}\right) \ge \\
&\ge -K_{\max}\, 2 C_\alpha h\,.
\end{aligned}$$

Это завершает доказательство леммы 1. ∎

При доказательстве следующих двух лемм предполагается, что последовательность X-точек выборки $(X_i)_{i=1,\ldots,N}$ уже упорядочена по возрастанию без изменения обозначения с X_i на $X_{(i)}$, для упрощения обозначений, т.е.

$$\text{(П.9)}\qquad X_i \le X_{i+1}\,, \quad \forall\, i\,.$$

Существенным является применение равномерной асимптотической границы порядка $O(\log N/N)$ для $\Delta X_i \triangleq X_i - X_{i-1}$, доказанной в лемме П.2.

Д о к а з а т е л ь с т в о л е м м ы 2. Предполагаем выполненным соотношение (П.9). Применяя вспомогательные леммы П.2 и П.4, приходим сначала к

$$\text{(П.10)}\qquad \max_{x\in[0,1]} |\widehat{f}_N'(x)| = \max_{1\le i\le N+1}\ \max_{x\in[X_{i-1},X_i]} |\widehat{f}_N'(x)| \le$$

$$
\text{(П.11)}\quad \le\; L_{f,\beta}\, g_{\max} C_\beta(K,K') \frac{\log N}{Nh^2} + \frac{1}{8} \max_{1\le i\le N+1} \left[(X_i - X_{i-1})^2 \max_{x\in[X_{i-1},X_i]} |\widehat{f}_N'''(x)| \right] \le
$$

$$
\text{(П.12)}\quad \le\; L_{f,\beta}\, g_{\max} C_\beta(K,K') \frac{\log N}{Nh^2} + \frac{1}{8} \left(C_X \frac{\log N}{N} \right)^2 \max_{x\in[0,1]} |\widehat{f}_N'''(x)|
$$

с $C_X > 4C_f/f_{\min}$. Максимум в (П.12) ограничивается сверху следующим образом: для любого $x \in [0,1]$

$$
\text{(П.13)}\qquad |\widehat{f}_N'''(x)| \;\le\; \sum_{i=1}^{N} \alpha_i \left| \frac{d^3}{dx^3} K_h(x, X_i) \right| \le
$$

$$
\text{(П.14)}\qquad \le\; \sup_{u,v} \left| \frac{\partial^3}{\partial v^3} K_h(v,u) \right| \sum_{i=1}^{N} \alpha_i \, \mathbf{1}\{|x - X_i| \le h\} \le
$$

$$
\text{(П.15)}\qquad \le\; g_{\max} L_{\widetilde{K}''} h^{-4} 3 C_\alpha h \,,
$$

поскольку (см. лемму П.1)

$$
\text{(П.16)}\qquad \sup_{u,v} \left| \frac{\partial^3}{\partial v^3} K_h(v,u) \right| \le g_{\max} L_{\widetilde{K}''} h^{-4} \,,
$$

где

$$
\text{(П.17)}\quad L_{\widetilde{K}''} \triangleq L_{K''} + 3 L_{K'} K_{\max} g_{\max} + L_K g_{\max} \left(3L_K + 10 K_{\max}^2 g_{\max} \right) + 6 K_{\max}^4 g_{\max}^3 \,.
$$

Подстановка (П.13)–(П.15) в (П.12) приводит к

$$
\begin{aligned}
\max_{x\in[0,1]} |\widehat{f}_N'(x)| \;&\le\; L_{f,\beta}\, g_{\max} C_\beta(K,K') \frac{\log N}{Nh^2} + \frac{3}{8}\, g_{\max} L_{\widetilde{K}''} C_\alpha \left(C_X \frac{\log N}{Nh^2} \right)^2 h \le \\
&\le\; 2 L_{f,\beta}\, g_{\max} C_\beta(K,K') \frac{\log N}{Nh^2}
\end{aligned}
$$

при следующем дополнительном предположении (которое справедливо для всякого достаточно большого N):

$$
\text{(П.18)}\qquad h \ge \frac{3 C_X^2 C_\alpha\, L_{\widetilde{K}''} \log N}{8 L_{f,\beta}\, C_\beta(K,K')\, N} \,.
$$

Отсюда следует требуемый результат. ■

Доказательство леммы 3. Рассмотрим произвольное $N \ge N_0(\omega)$ с $N_0(\omega)$ из леммы П.2. Введем функцию $f_\gamma(u) = f(u) + \gamma h^\beta$ и псевдооценки

$$
\text{(П.19)}\qquad \widetilde{\alpha}_i = \frac{1+\delta_{i1}}{2} \int_{X_{i-1}}^{X_i} f_\gamma(u)\, du + \frac{1+\delta_{iN}}{2} \int_{X_i}^{X_{i+1}} f_\gamma(u)\, du \,, \quad i = 1, \ldots, N \,,
$$

где δ_{ij} обозначает символ Кронекера. Ниже показывается, что условие (33) приводит к тому, что вектор псевдооценок $\widetilde{\alpha} = (\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_N)^{\mathrm{T}}$ представляет собой допустимую точку в ЛП-задаче (18)–(22) для любого достаточно большого N. Отсюда вытекает разрешимость ЛП-задачи (18)–(22) и

$$J_P^* \leq \sum_{i=1}^{N} \widetilde{\alpha}_i = \int_0^1 (f(u) + \gamma h^\beta)\, du = C_f + \gamma h^\beta \,. \tag{П.20}$$

Положим $C_X > 4C_f/f_{\min}$. Для простоты введем дополнительные предположения

$$h^\beta \leq \frac{\log N}{\rho N h} \leq \min\left\{\frac{f_{\max}}{\gamma}, \frac{1}{\rho C_X}\right\}, \tag{П.21}$$

которые справедливы для всякого достаточно большого N.

1. Сначала покажем, что при $\alpha_i = \widetilde{\alpha}_i$, $i = 1, \ldots, N$ ограничения (14) выполняются. Для произвольного $x \in [0, 1]$

$$\widetilde{f}_N(x) \triangleq \sum_{i=1}^{N} \widetilde{\alpha}_i K_h(x, X_i) = \tag{П.22}$$

$$= \sum_{i=1}^{N+1} \int_{X_{i-1}}^{X_i} f_\gamma(u)\, du\, \frac{K_h(x, X_i) + K_h(x, X_{i-1})}{2} + \tag{П.23}$$

$$+ \frac{1}{2} \int_0^{X_1} f_\gamma(u)\, du\, (K_h(x, X_1) - K_h(x, 0)) + \tag{П.24}$$

$$+ \frac{1}{2} \int_{X_N}^{1} f_\gamma(u)\, du\, (K_h(x, X_N) - K_h(x, 1)) = \tag{П.25}$$

$$= \int_0^1 f_\gamma(u)\, K_h(x, u)\, du + \tag{П.26}$$

$$+ \sum_{i=1}^{N+1} \int_{X_{i-1}}^{X_i} f_\gamma(u) \left(\frac{K_h(x, X_i) + K_h(x, X_{i-1})}{2} - K_h(x, u)\right) du + \tag{П.27}$$

$$+ \frac{1}{2} \int_0^{X_1} f_\gamma(u)\, du\, (K_h(x, X_1) - K_h(x, 0)) + \tag{П.28}$$

$$+ \frac{1}{2} \int_{X_N}^{1} f_\gamma(u)\, du\, (K_h(x, X_N) - K_h(x, 1)) \,. \tag{П.29}$$

Теперь ограничим снизу каждое слагаемое в (П.26)–(П.29). Благодаря (6) основной член (П.26) ограничивается следующим образом:

$$\int_0^1 f_\gamma(u)\, K_h(x, u)\, du = f(x) + \gamma h^\beta + \int_0^1 (f(u) - f(x))\, K_h(x, u)\, du \geq \tag{П.30}$$

$$\geq f(x) + (\gamma - L_{f,\beta}\, g_{\max} C_\beta(K)) h^\beta \,. \tag{П.31}$$

Далее, i-е слагаемое из (П.27) разлагается и затем ограничивается на основе формулы трапеции:

$$\int_{X_{i-1}}^{X_i} f_\gamma(u) \left(\frac{K_h(x, X_i) + K_h(x, X_{i-1})}{2} - K_h(x, u)\right) du \geq \tag{П.32}$$

$$
\begin{aligned}
&\geq \; f_\gamma(x)\int_{X_{i-1}}^{X_i}\left(\frac{K_h(x,X_i)+K_h(x,X_{i-1})}{2}-K_h(x,u)\right)du-\\
&\quad -\int\limits_{X_{i-1}}^{X_i}|f_\gamma(u)-f_\gamma(x)|\left|\frac{K_h(x,X_i)+K_h(x,X_{i-1})}{2}-K_h(x,u)\right|du\geq
\end{aligned}
$$

$$
\geq \; -(f_{\max}+\gamma h^\beta)\,\frac{(X_i-X_{i-1})^3}{12}\max_{u\in[0,1]}\left|\frac{\partial^2 K_h(x,u)}{\partial u^2}\right|\mathbf{1}\{|x-X_i|\leq 2h\}- \tag{П.33}
$$

$$
-L_{f,\beta}\int_{X_{i-1}}^{X_i}|u-x|^\beta\,\mathbf{1}\{|x-X_i|\leq 2h\}\frac{g_{\max}L_K}{2h^2}\left[(u-X_{i-1})+(X_i-u)\right]du. \tag{П.34}
$$

Применяя лемму П.2, первый член оцениваем следующим образом:

$$
(\text{П.33})\geq-\left(C_X\frac{\log N}{N}\right)^2\frac{f_{\max}\,g_{\max}L_{K'}}{6h^3}\,(X_i-X_{i-1})\,\mathbf{1}\{|x-X_i|\leq 2h\}; \tag{П.35}
$$

второй член оцениваем следующим образом:

$$
(\text{П.34})\geq-\frac{g_{\max}L_{f,\beta}\,L_K}{2h^2}C_X\frac{\log N}{N}\mathbf{1}\{|x-X_i|\leq 2h\}\int_{X_{i-1}}^{X_i}|u-x|^\beta du\,. \tag{П.36}
$$

Кроме того, исходя из леммы П.2 можно, во-первых, показать, что

$$
\sum_{i=1}^{N+1}\mathbf{1}\{|x-X_i|\leq 2h\}(X_i-X_{i-1})\leq 4h+\frac{C_X\log N}{N},
$$

а во-вторых, что

$$
\begin{aligned}
&\sum_{i=1}^{N+1}\mathbf{1}\{|x-X_i|\leq 2h\}\int_{X_{i-1}}^{X_i}|u-x|^\beta du\leq\\
&\leq\int_{x-2h-C_X(\log N)/N}^{x+2h}|u-x|^\beta du\leq\\
&\leq\left(4h+C_X\frac{\log N}{N}\right)\max_{v\in[-2h-C_X(\log N)/N,2h]}|v|^\beta\leq\\
&\leq\left(4h+C_X\frac{\log N}{N}\right)\left(2h+C_X\frac{\log N}{N}\right)^\beta.
\end{aligned}
$$

Таким образом, приходим к следующей границе для суммы (П.27):

$$
\begin{aligned}
&\sum_{i=1}^{N+1}\int_{X_{i-1}}^{X_i}f_\gamma(u)\left(\frac{K_h(x,X_i)+K_h(x,X_{i-1})}{2}-K_h(x,u)\right)du\geq\\
&\geq\; -g_{\max}C_X\frac{\log N}{Nh}\left(4+C_X\frac{\log N}{Nh}\right)\left(\frac{C_Xf_{\max}L_{K'}}{6}\frac{\log N}{Nh}+\right.\\
&\qquad\qquad\left.+\frac{L_{f,\beta}\,L_K}{2}h^\beta\left(2+C_X\frac{\log N}{Nh}\right)^\beta\right)\geq\\
&\geq\; -\frac{5}{6}g_{\max}C_X\left(\frac{\log N}{Nh}\right)^2\left(C_Xf_{\max}L_{K'}+3^{\beta+1}\rho^{-1}L_{f,\beta}\,L_K\right).
\end{aligned}
$$

Наконец, аналогичным образом показываем, что оба слагаемых (П.28) и (П.29) ограничены сверху членом $O((\log N/(Nh))^2)$. Например, для (П.28) получаем

$$\begin{aligned}\left|\int_0^{X_1} f_\gamma(u)\,du\,(K_h(x,X_1)-K_h(x,0))\right| &\le (f_{\max}+\gamma h^\beta)X_1\,|K_h(x,X_1)-K_h(x,0)| \le \\ &\le 2f_{\max}\,g_{\max}L_K\left(C_X\frac{\log N}{Nh}\right)^2 . \qquad (\Pi.37)\end{aligned}$$

Таким образом, из (П.22)–(П.37) следует для каждого $j=1,\ldots,N$, что

$$\widetilde{f}_N(X_j) \ge f(X_j) + (\gamma - L_{f,\beta}\;g_{\max}C_\beta(K))h^\beta + O\left(\left(\frac{\log N}{Nh}\right)^2\right) \ge Y_j \qquad (\Pi.38)$$

для достаточно больших $N \ge N_0(\omega)$, когда выполнены как неравенства (П.21), так и следующее условие:

$$\gamma - L_{f,\beta}\;g_{\max}C_\beta(K) \ge \frac{5}{6}\,g_{\max}C_X\left(\frac{\log N}{Nh^{1+\beta/2}}\right)^2\left(C_Xf_{\max}\left(L_{K'}+\frac{12L_K}{5}\right)+3^{\beta+1}\frac{L_{f,\beta}\,L_K}{\rho}\right).$$

2. Аналогично, ограничения (15) выполнены при $\alpha_i=\widetilde{\alpha}_i$, $i=1,\ldots,N$. Действительно, для всякого $x\in[0,1]$ теперь нужно ограничить абсолютное значение суммы

$$\widetilde{f}'_N(x) = \sum_{i=1}^N \widetilde{\alpha}_i\,\frac{d}{dx}\,K_h(x,X_i) = \sum_{i=1}^N \widetilde{\alpha}_i\,\widetilde{K}_h(x,X_i) \qquad (\Pi.39)$$

вместо (П.22). Здесь

$$\widetilde{K}_h(x,u) \triangleq \frac{\partial}{\partial x}\,K_h(x,u) \qquad (\Pi.40)$$

со следующей верхней гранью

$$\left|\widetilde{K}_h(x,u)\right| \le h^{-2}g_{\max}\left\{\left|K'\left(\frac{x-u}{h}\right)\right| + g_{\max}\,K_{\max}\left|K\left(\frac{x-u}{h}\right)\right|\right\} . \qquad (\Pi.41)$$

Следовательно, можно повторить рассуждения (П.23)–(П.29), заменив K_h на $\widetilde{K}_h$. Все скорости из (П.32)–(П.38) следует разделить теперь на h, а абсолютное значение основного члена разложения благодаря (7) ограничивается следующим образом:

$$\begin{aligned}\left|\int_0^1 f_\gamma(u)\,\widetilde{K}_h(x,u)\,du\right| &= \left|\int_0^1 (f(u)-f(x))\,\frac{\partial}{\partial x}\,K_h(x,u)\,du\right| \le \\ &\le L_{f,\beta}\;g_{\max}C_\beta(K,K')h^{\beta-1} ,\end{aligned}$$

вместо (П.30)–(П.31). Напомним определение (8)–(9) для $C_\beta(K,K')$, которое вытекает из (П.41). Таким образом, для достаточно больших $N \ge N_0(\omega)$ и для каждого X_j приходим к неравенству

$$\left|\widetilde{f}'_N(X_j)\right| \le L_{f,\beta}\;g_{\max}C_\beta(K,K')h^{\beta-1} + O\left(\frac{\log^2 N}{N^2h^3}\right) \le L_{f,\beta}\;g_{\max}C_\beta(K,K')\frac{\log N}{\rho Nh^2}. \qquad (\Pi.42)$$

Именно, неравенство (П.42) справедливо почти наверное (п.н.) для всех таких $N \geq N_0(\omega)$, что справедливо (П.21) и

$$L_{f,\beta}\, C_\beta(K,K')\left(\frac{\log N}{h^{\beta+1}N}-1\right) \;\geq\; \frac{5}{6}\, g_{\max} C_X\left(\frac{\log N}{Nh^{1+\beta/2}}\right)^2 \times \tag{П.43}$$

$$\times\left(C_X f_{\max}\left(L_{\widetilde{K}'}+\frac{12L_{\widetilde{K}}}{5}\right)+3^{\beta+1}\frac{L_{f,\beta}\, L_{\widetilde{K}}}{\rho}\right), \tag{П.44}$$

где (подробнее см. доказательство леммы П.1)

$$L_{\widetilde{K}} \triangleq L_{K'} + L_K\, g_{\max} K_{\max}\,, \quad L_{\widetilde{K}'} \triangleq L_{K''} + L_{K'}\, g_{\max} K_{\max}\,. \tag{П.45}$$

3. Наконец, ограничения (16) с

$$C_\alpha \geq 6 f_{\max} \tag{П.46}$$

также выполнены при $\alpha_i = \widetilde{\alpha}_i$, $i = 1, \ldots, N$. Действительно, по лемме П.2, п.н. для всякого $N \geq N_0(\omega)$ и любого $j = 1, \ldots, m_h$ (где $m_h = \lfloor h^{-1} \rfloor$) имеет место следующее неравенство:

$$\sum_{i=1}^{N} \widetilde{\alpha}_i\, \mathbf{1}\{(j-1)/m_h \leq X_i < j/m_h\} \;\leq\; (f_{\max} + \gamma h^\beta)\left(1/m_h + 2C_X\frac{\log N}{N}\right) \leq \tag{П.47}$$

$$\leq\; 6 f_{\max} h, \tag{П.48}$$

при дополнительных предположениях (П.21). Таким образом, ограничения (16) выполнены при (П.46) п.н., для достаточно больших N.

4. Поскольку все $\widetilde{\alpha}_i \geq 0$, ограничения (17) выполнены, и лемма 3 доказана. ∎

Доказательство леммы 4. Воспользуемся леммой П.3 и ее следствием П.1. Положим

$$\delta_y = L_{f,\beta}\, \delta_x^\beta, \quad \delta_x \triangleq \left(\frac{2C_f \log N}{f_{\min} L_{f,\beta}\, N}\right)^{\frac{1}{1+\beta}}. \tag{П.49}$$

Таким образом, с вероятностью 1 для любого $N \geq N_6(\omega)$ и произвольного $x \in [0,1]$ существуют целые числа $i_k \in \{1, \ldots, N\}$ такие, что

$$|x - X_{i_k}| \leq \delta_x \tag{П.50}$$

и

$$Y_{i_k} \geq f(X_{i_k}) - \delta_y\,. \tag{П.51}$$

Теперь, ошибка оценивания в точке x может быть представлена следующим разложением:

$$f(x) - \widehat{f}_N(x) \;=\; [f(x) - f(X_{i_k})] + \tag{П.52}$$

$$+\; \left[f(X_{i_k}) - \widehat{f}_N(X_{i_k})\right] + \tag{П.53}$$

$$+\; \left[\widehat{f}_N(X_{i_k}) - \widehat{f}_N(x)\right]. \tag{П.54}$$

Член в правой части (П.52) может быть ограничен следующим образом:

$$|f(x) - f(X_{i_k})| \le L_{f,\beta}\, |x - X_{i_k}|^\beta \le L_{f,\beta}\, \delta_x^\beta\,; \tag{П.55}$$

аналогично ограничивается член (П.54):

$$\left|\widehat{f}_N(X_{i_k}) - \widehat{f}_N(x)\right| \le L_{\widehat{f}_N}\, |x - X_{i_k}| \le L_{\widehat{f}_N}\, \delta_x \tag{П.56}$$

с константой Липшица $L_{\widehat{f}_N}$ для функции-оценки $\widehat{f}_N(x)$. Напомним, что $\widehat{f}_N(X_{i_k}) \ge Y_{i_k}$ вследствие (14) или (19). Таким образом, (П.51) влечет

$$f(X_{i_k}) - \widehat{f}_N(X_{i_k}) \le (Y_{i_k} + \delta_y) - Y_{i_k} = \delta_y\,. \tag{П.57}$$

Объединяя все эти оценки, получаем из (П.52), что для всякого $N \ge N_6(\omega)$

$$f(x) - \widehat{f}_N(x) \le \delta_y + L_{f,\beta}\, \delta_x^\beta + L_{\widehat{f}_N}\delta_x\,. \tag{П.58}$$

Следовательно, применение леммы 2 и подстановка выражения (П.49) для δ_x и δ_y в (П.58) приводит к нижней границе

$$\widehat{f}_N(x) \;\ge\; f(x) - \left(2L_{f,\beta}\, \delta_x^\beta + L_{\widehat{f}_N}\delta_x\right) \ge \tag{П.59}$$

$$\ge\; f(x) - \frac{C_4(\beta)}{h^2}\left(\frac{\log N}{N}\right)^{\frac{2+\beta}{1+\beta}} \tag{П.60}$$

для любого достаточно большого N (начиная со случайного, конечного п.н., целого числа, которое не зависит от x). Здесь было использовано первое неравенство в (36), что позволило упростить выражение нижней границы. Лемма 4 доказана. ■

Доказательство теоремы 1. Поскольку $|u| = u - 2u\mathbf{1}\{u < 0\}$, L_1-норма ошибки оценивания может быть разложена следующим образом:

$$\|\widehat{f}_N - f\|_1 \;=\; \int_0^1 \left[\widehat{f}_N(x) - f(x)\right]\, dx + \tag{П.61}$$

$$+\;\; 2\int_0^1 \left[f(x) - \widehat{f}_N(x)\right] \mathbf{1}\left\{\widehat{f}_N(x) < f(x)\right\}\, dx. \tag{П.62}$$

Применение лемм 1 и 3 к правой части (П.61) дает

$$\limsup_{N\to\infty} h^{-\beta}\left(\int_0^1 \left[\widehat{f}_N(x) - f(x)\right]\, dx\right) \le \gamma + 4C_\alpha(g_{\max} - 1)K_{\max}\mathbf{1}\{\beta = 1\} \quad \text{п.н.} \tag{П.63}$$

Заметим, что можно, например, зафиксировать $\gamma = 2L_{f,\beta}\, g_{\max}\, C_\beta(K)$. Для получения аналогичного результата для члена (П.62) заметим, что из леммы 4 следует

$$\zeta_N(x, \omega) \triangleq \varepsilon_{LB}^{-1}(N)\left[f(x) - \widehat{f}_N(x)\right] \le C_4(\beta) < \infty \quad \text{п.н.}$$

равномерно относительно как $x \in [0,1]$, так и $N \geq N_2(\omega)$, с

$$\varepsilon_{LB}(N) \triangleq \frac{1}{h^2}\left(\frac{\log N}{N}\right)^{\frac{2+\beta}{1+\beta}} . \tag{П.64}$$

Следовательно, можно применить лемму Фату, учитывая непрерывность и монотонность функции $\psi(u) = u\mathbf{1}\{u > 0\}$:

$$\begin{aligned} &\limsup_{N\to\infty} \varepsilon_{LB}^{-1}(N) \int_0^1 \left[f(x) - \widehat{f}_N(x)\right] \mathbf{1}\left\{\widehat{f}_N(x) < f(x)\right\} dx \leq \\ \leq\ & \int_0^1 \limsup_{N\to\infty} \zeta_N(x,\omega)\, \mathbf{1}\{\zeta_N(x,\omega) > 0\}\, dx \leq \\ \leq\ & C_4(\beta) < \infty \quad \text{п.н.} \end{aligned}$$

Таким образом, полученные соотношения вместе с (П.61) и (П.62) приводят к (37). Теорема 1 доказана. ■

В следующей лемме получена константа Липшица из (П.45) и (П.16)–(П.17). Ее доказательство, детально изложенное в [24], носит технический характер и здесь не приводится.

Л е м м а П.1. *Пусть ядро K_h, определенное в (4)–(5), удовлетворяет предположениям B1–B3 и ширина окна $h \in (0, 1/2)$. Пусть $\widetilde{K}_n$ определена соотношением (П.40). Тогда справедливы следующие верхние границы:*

$$\left|\widetilde{K}_h(x,u)\right| \leq g_{\max} h^{-2}\left(L_K + g_{\max} K_{\max}^2\right) , \tag{П.65}$$

$$\left|\frac{\partial}{\partial u}\,\widetilde{K}_h(x,u)\right| \leq g_{\max} h^{-3} L_{\widetilde{K}} , \quad \left|\frac{\partial^2}{\partial u^2}\,\widetilde{K}_h(x,u)\right| \leq g_{\max} h^{-4} L_{\widetilde{K}'} , \tag{П.66}$$

где $L_{\widetilde{K}} = L_{K'} + L_K\, g_{\max} K_{\max}$ и $L_{\widetilde{K}'} = L_{K''} + L_{K'}\, g_{\max} K_{\max}$. Кроме того,

$$\left|\frac{\partial^3}{\partial x^3}\, K_h(x,u)\right| \leq g_{\max} L_{\widetilde{K}''} h^{-4}, \tag{П.67}$$

где

$$L_{\widetilde{K}''} = g_{\max}\left[L_{K''} + 3L_{K'}K_{\max}g_{\max} + 3L_K g_{\max} K_{\max}^2 (1 + 3g_{\max}) + \right. \tag{П.68}$$

$$\left. + (L_K^2 + 2g_{\max}^2 K_{\max}^4)(1 + 2g_{\max})\right] . \tag{П.69}$$

Для полноты изложения ниже кратко доказываются следующие вспомогательные результаты.

Л е м м а П.2. *Пусть функция $f : [0,1] \to \mathbb{R}$ удовлетворяет предположению A1, а последовательность $(X_i)_{i=1,\dots,N}$ получена из независимой выборки с плотностью распределения $f(x)/C_f$ посредством упорядочения по возрастанию (П.9), где C_f определена в (2). Обозначим $X_0 = 0$ и $X_{N+1} = 1$. Тогда для произвольной конечной постоянной $C_X > 4C_f/f_{\min}$ существует почти наверное постоянное число $N_0 = N_0(\omega)$ такое, что*

$$\max_{i=1,\dots,N+1} \Delta X_i \leq C_X \frac{\log N}{N} \quad \forall\ N \geq N_0 \tag{П.70}$$

с вероятностью 1. Например, можно взять следующее значение константы C_X:

$$C_X = 5 f_{\max}/f_{\min}\,. \tag{П.71}$$

Доказательство леммы П.2. Введем эквидистантное разбиение интервала $[0,1]$ на m_N подынтервалов Δ_k с (одинаковой) лебеговой мерой

$$\ell(\Delta_k) \triangleq 1/m_N \le C_X \log N/(2N)\,, \quad k = 1, \ldots, m_N\,, \tag{П.72}$$

где размер разбиения

$$m_N \triangleq \min\{\text{целое}\, m \,:\, m \ge 2N/(C_X \log N)\} \le \tag{П.73}$$

$$\le 1 + \frac{2N}{C_X \log N} \le \frac{(2+\varepsilon)N}{C_X \log N} \tag{П.74}$$

для произвольного $\varepsilon > 0$ и для любого достаточно большого N. Следовательно, событие

$$A_N \triangleq \{\omega \,:\, \max_{i=1,\ldots,N+1} \Delta X_i \le C_X \log N/N\} \supseteq \bigcap_{k=1}^{m_N} \left[\bigcup_{i=1}^{N} \{X_i \in \Delta_k\} \right].$$

По лемме Бореля–Кантелли доказываем, что дополнительное событие $A_N^c \triangleq \Omega \setminus A_N$ может произойти только конечное число раз (с вероятностью 1): очевидно,

$$P\left(A_N^c\right) \le \sum_{k=1}^{m_N} P\left(\bigcap_{i=1}^{N} \{X_i \notin \Delta_k\} \right) = \tag{П.75}$$

$$= \sum_{k=1}^{m_N} \prod_{i=1}^{N} \left(1 - \int_{\Delta_k} C_f^{-1} f(u)\, du \right) \le \tag{П.76}$$

$$\le m_N \left(1 - \frac{f_{\min}}{C_f}\, \ell(\Delta_1) \right)^N \le \tag{П.77}$$

$$\le \frac{(2+\varepsilon)N}{C_X \log N} \exp\left\{ -\frac{f_{\min} C_X}{(2+\varepsilon) C_f} \log N \right\} = \tag{П.78}$$

$$= O\left(N^{1 - f_{\min} C_X/((2+\varepsilon)C_f)} \right),$$

условие $C_X > 4 C_f/f_{\min}$ обеспечивает существование положительного ε, гарантирующего сходимость ряда

$$\sum_{N=1}^{\infty} P\left(A_N^c\right) < \infty\,, \tag{П.79}$$

и можно применить лемму Бореля–Кантелли. Заметим, что события $\bigcap_{i=1}^N \{X_i \notin \Delta_k\}$ не зависят от перенумеровки точек $(X_i)_{i=1,\ldots,N}$, что приводит к (П.76) из (П.75); кроме того, в (П.77)–(П.78) было использовано как соотношение (П.73), так и неравенство $1 - x \le e^{-x}$. Лемма П.2 доказана. ■

Л е м м а П.3. Пусть случайная выборка $\{(X_i, Y_i) \mid i = 1, \ldots, N\}$ *определена в разделе 2, а последовательность* $\delta_x = \delta_x(N)$ *положительна, и для некоторого* $\varepsilon > 0$

$$\liminf_{N\to\infty} N^{1-\varepsilon}\delta_x > 0\,. \tag{П.80}$$

Определим

$$m_\delta \triangleq \min\{\text{\textit{целое} } m \,:\, m \ge \delta_x^{-1}\} \tag{П.81}$$

и предположим, что положительная последовательность $\delta_y = \delta_y(N) < f_{\min}$ *для достаточно больших* N *удовлетворяет неравенству*

$$\delta_y \ge \kappa\, m_\delta \frac{\log N}{N}\,, \quad \text{\textit{где}} \quad \kappa > \frac{(2-\varepsilon)C_f}{f_{\min}}\,. \tag{П.82}$$

Тогда в условиях леммы П.2 почти наверное существует конечное $N_6(\omega)$ *такое, что для всякого* $N \ge N_6(\omega)$ *найдется подмножество точек* $\{(X_{i_k}, Y_{i_k})\,,\, k = 1, \ldots, m_\delta\}$ *в выборке* $\{(X_i, Y_i)\,,\ i = 1, \ldots, N\}$, *для которого имеют место следующие неравенства:*

$$(k-1)/m_\delta \le X_{i_k} < k/m_\delta\,, \quad f(X_{i_k}) - \delta_y \le Y_{i_k} \le f(X_{i_k})\,. \tag{П.83}$$

Д о к а з а т е л ь с т в о л е м м ы П.3 аналогично доказательству леммы П.2. Введем эквидистантное разбиение интервала $[0,1]$ на подынтервалы $[(k-1)/m_\delta, k/m_\delta]$, $k = 1, \ldots, m_\delta$. Кроме того, введем соответствующие множества в $\mathbb{R}^2$:

$$\Delta_k \triangleq \{(u,v) : (k-1)/m_\delta \le u \le k/m_\delta\,,\ f(u) - \delta_y \le v \le f(u)\}\,, \quad k = 1, \ldots, m_\delta\,.$$

Рассмотрим событие

$$\begin{aligned} A_N &\triangleq \{\omega \,:\, \forall\, k = 1, \ldots, m_\delta\ \exists\, i = 1, \ldots, N : (X_i, Y_i) \in \Delta_k\} = \\ &= \bigcap_{k=1}^{m_\delta}\left[\bigcup_{i=1}^{N}\{(X_i, Y_i) \in \Delta_k\}\right]. \end{aligned}$$

Оценим вероятность дополнительного события:

$$\begin{aligned} P\left(A_N^c\right) &\le \sum_{k=1}^{m_\delta} P\left(\bigcap_{i=1}^{N}\{(X_i, Y_i) \notin \Delta_k\}\right) = \sum_{k=1}^{m_\delta}\prod_{i=1}^{N}\left(1 - \int_{\Delta_k} C_f^{-1} f(u)\, du\right) \le \\ &\le m_\delta \left(1 - \frac{f_{\min}\,\delta_y}{C_f\, m_\delta}\right)^N \le \left(1 + \delta_x^{-1}\right)\, \exp\left\{-\frac{f_{\min}\kappa}{C_f}\,\log N\right\} = \\ &= O\left(N^{1-\varepsilon - f_{\min}\kappa/C_f}\right). \end{aligned}$$

Значит, условие $\kappa > (2-\varepsilon)C_f/f_{\min}$ обеспечивает $\sum_{N=1}^{\infty} P\left(A_N^c\right) < \infty$ и можно применить лемму Бореля–Кантелли. Лемма П.3 доказана. ■

Следствие П.1. Пусть δ_x и δ_y удовлетворяют условиям леммы П.3. Тогда с вероятностью 1 для всякого $N \geq N_6(\omega)$ и произвольного $x \in [0,1]$ существует целое число $i_k \in \{1, \dots, N\}$ такое, что $|x - X_{i_k}| \leq \delta_x$ и $f(X_{i_k}) - \delta_y \leq Y_{i_k} \leq f(X_{i_k})$.

Лемма П.4. Пусть функция $g : [0, \Delta] \to \mathbb{R}$ дважды непрерывно дифференцируема, $\Delta > 0$. Тогда

$$\max_{x\in[0,\Delta]} |g(x)| \leq \max\{|g(0)|, |g(\Delta)|\} + \frac{\Delta^2}{8} \max_{x\in[0,\Delta]} |g''(x)| \,. \tag{П.84}$$

Доказательство леммы П.4. Обозначим $\bar{g}_b = \max\{|g(0)|, |g(\Delta)|\}$. Достаточно доказать случай, когда существует точка $x_1 \in (0, \Delta)$ с

$$|g(x_1)| = \max_{x\in[0,\Delta]} |g(x)| > \bar{g}_b \,. \tag{П.85}$$

Тогда $g'(x_1) = 0$ и для любого $x \in [0, \Delta]$

$$g(x_1) = g(x) - \int_{x_1}^{x} dt \int_{x_1}^{\mathrm{T}} g''(u)\, du \,. \tag{П.86}$$

Следовательно, положив $x = \Delta$, получаем из (П.86)

$$|g(x_1)| \leq |g(\Delta)| + \int_{x_1}^{\Delta} dt \int_{x_1}^{\mathrm{T}} |g''(u)|\, du \leq \bar{g}_b + \frac{(\Delta - x_1)^2}{2} \max_{x\in[0,\Delta]} |g''(x)| \,. \tag{П.87}$$

Аналогично, положив $x = 0$ в (П.86), приходим к

$$|g(x_1)| \leq |g(0)| + \int_{0}^{x_1} dt \int_{x_1}^{\mathrm{T}} |g''(u)|\, du \leq \bar{g}_b + \frac{x_1^2}{2} \max_{x\in[0,\Delta]} |g''(x)| \,. \tag{П.88}$$

Таким образом, объединив (П.87) и (П.88), приходим к

$$|g(x_1)| \leq \bar{g}_b + \frac{1}{2} \min\{(\Delta - x_1)^2, x_1^2\} \max_{x\in[0,\Delta]} |g''(x)| \,. \tag{П.89}$$

Поскольку

$$\max_{x\in[0,\Delta]} \min\{(\Delta - x)^2, x^2\} = \frac{\Delta^2}{4} \,, \tag{П.90}$$

сразу получаем желаемое неравенство (П.84) из (П.85), (П.89)–(П.90). ■

СПИСОК ЛИТЕРАТУРЫ

1. *Hardy A., Rasson J.P.* Une nouvelle approche des problèmes de classification automatique // Stat. Anal. Données. 1982. V.7. P. 41–56.

2. *Hartigan J.A.* Clustering Algorithm. Chichester: Wiley, 1975.

3. *Baufays P., Rasson J.P.* A new geometric discriminant rule // Comput. Stat. Quaterly. 1985. V.2. P. 15–30.

4. *Devroye L.P., Wise G.L.* Detection of abnormal behavior via non parametric estimation of the support // SIAM J. Appl. Math. 1980. V.38. P. 448–480.

5. *Tarassenko L., Hayton P., Cerneaz N., Brady M.* Novelty detection for the identification of masses in mammograms // Proc. fourth IEE Int. Conf. Artificial Neural Networks. Cambridge, 1995. P. 442–447.

6. *Korostelev A.P., Tsybakov A.B.* Minimax theory of image reconstruction. Lect. Notes Statist. N.Y.: Springer-Verlag, 1993. V.82.

7. *Deprins D., Simar L., Tulkens H.* Measuring Labor Efficiency in Post Offices / "The Performance of Public Enterprises: Concepts and Measurements," M. Marchand, P. Pestieau, and H. Tulkens, eds. P. 243-267. Amsterdam: North Holland, 1984.

8. *Härdle W., Hall P., Simar L.* Iterated bootstrap with application to frontier models // J. Productiv. Anal. 1995. V.6. P. 63–76.

9. *Geffroy J.* Sur un problème d'estimation géométrique. Publications de l'Institut de Stat., Université de Paris, 1964. V.XIII. P. 191–200.

10. *Härdle W., Park B.U., Tsybakov A. B.* Estimation of a non sharp support boundaries // J. Multivariate Anal. 1995. V.43. P. 205–218.

11. *Hall P., Nussbaum M., Stern S.E.* On the estimation of a support curve of indeterminate sharpness // J. Multivariate Anal. 1997. V.62. P. 204–232.

12. *Gijbels I., Peng L.* Estimation of a support curve via order statistics // Extremes. 2000. V.3. P. 251–277.

13. *Knight K.* Limiting distributions of linear programming estimators // Extremes. 2001. V.4. No.2. P. 87–103.

14. *Hall P., Park B.U., Stern S.E.* On polynomial estimators of frontiers and boundaries // J. Multivariate Anal. 1998. V.66. P. 71–98.

15. *Girard S., Jacob P.* Extreme values and kernel estimates of point processes boundaries // ESAIM: Prob. Stat. 2004. V.8. P. 150–168.

16. *Girard S., Jacob P.* Extreme values and Haar series estimates of point processes boundaries // Scandinav. J. Stat. 2003. V.30. No.2. P. 369–384.

17. *Girard S., Jacob P.* Projection estimates of point processes boundaries // J. Stat. Planning and Inference. 2003. V.116. No.1. P. 1–15.

18. *Gardes L.* Estimating the support of a Poisson process via the Faber-Shauder basis and extreme values // Publications de l'Institut de Stat., Université de Paris, 2002. V.XXXXVI. P. 43–72.

19. *Girard S., Menneteau L.* Central limit theorems for smoothed extreme value estimates of Poisson point processes boundaries // J. Stat. Planning and Inference. 2005. V.135. No.2. P. 433–460.

20. *Abbar H.* Un estimateur spline du contour d'une répartition ponctuelle aléatoire // Stat. Anal. Données. 1990. V.15. No.3. P. 1–19.

21. *Jacob P., Suquet P.* Estimating the edge of a Poisson process by orthogonal series // J. Stat. Planning and Inference. 1995. V.46. P. 215–234.

22. *Bouchard G., Girard S., Iouditski A., Nazin A.* Linear programming problems for frontier estimation. Technical Report INRIA RR-4717, 2003. `http://www.inria.fr/rrrt/rr-4717.html`, and/or Technical Report IAP RT-0304, `http://www.stat.ucl.ac.be/Iapdp/tr2003/TR0304.ps`.

23. *Бушар Г., Жирар С., Юдицкий А.Б., Назин А.В.* Непараметрическое оценивание границы носителя методом линейного программирования // АиТ. 2004. N 1. С. 66–73.

24. *Girard S., Iouditski A., Nazin A.* Linear programming problems for L_1-frontier estimation. Technical Report INRIA RR-5466, 2005. `http://www.inria.fr/rrrt/rr-5466.html`, and/or Technical Report IAP RT-0506, `http://www.stat.ucl.ac.be/ISpub/tr/2005/TR0506.pdf`.