ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ЕЕ ПРИМЕНЕНИЯ

Том XXIII 1978 Выпуск 3

ПОЛНЫЕ АСИМПТОТИЧЕСКИЕ РАЗЛОЖЕНИЯ ДЛЯ МАКСИМАЛЬНОГО УКЛОНЕНИЯ ЭМПИРИЧЕСКОЙ ФУНКЦИИ ПЛОТНОСТИ

В. Д. КОНАКОВ

В настоящей работе исследуется предельное поведение распределения максимального уклонения эмпирической функции плотности $f_n(\mathrm{x})$ p-мерного случайного вектора X. В отличие от известных результатов ([5], [7]), в данной работе кроме предельного распределения найдено полное асимптотическое разложение. Основной результат работы содержится в теореме 3, в которой приведено асимптотическое разложение для распределения случайной величины

$$\max_{1\leqslant i\leqslant s_n^p}\left|\frac{f_n(\mathrm{x}_i)-\mathrm{M}f_n(\mathrm{x}_i)}{\sqrt{\mathrm{D}f_n^*(\mathrm{x}_i)}}\right|,\quad s_n=o(n).$$

Для получения наших результатов мы используем ряд приемов. Первый — это дополнительная рандомизация, состоящая в переходе к пуассоновскому числу наблюдений. Прием этот принадлежит, по-видимому, Бикелу и Розенблатту [3] и состоит в том, что вместо n наблюдений берут случайное число наблюдений, распределенное по закону Пуассона со средним n и независимое от наблюдений. При этом оценки $f_n^*(\mathrm{x})$ и $f_n^*(\mathrm{y})$, построенные по пуассоновскому числу наблюдений для различных точек x, $\mathrm{y}\in R^p$, будут независимы, если они строятся с помощью наблюдений, попавших в непересекающиеся области, одна из которых содержит точку x, другая — y. Второй прием — это использование асимптотических разложений для вероятностей больших уклонений для сумм независимых слагаемых в схеме серий. Эти разложения получены в теореме 1. Доказательство теоремы 1 существенно использует лемму о больших уклонениях, принадлежащую В. А. Статулявичусу (лемма 1). Третий прием — это применение пуассоновской аппроксимации к биномиальному распределению (лемма 3). Комбинируя указанные приемы, мы получаем асимптотические разложения для максимального уклонения оценок $f_n^*(\mathrm{x})$, построенных по пуассоновскому числу наблюдений (теорема 2). Для перехода от оценки $f_n^*(\mathrm{x})$ к $f_n(\mathrm{x})$ (теорема 3) мы даем оценку $\mathbf{P}(|f_n^*(\mathrm{x})-f_n(\mathrm{x})|>\varepsilon)$ (лемма 4).

В заключительной части работы мы улучшаем аппроксимацию изучаемого распределения (теорема 4), доводя ее до $O(n^{-1/4+\varepsilon})$ для любого $\varepsilon>0$, в то время как предельное распределение отличается от изучаемого на величину порядка $O(1/\ln n)$. Отмечается, что в одномерном случае оценку остаточного члена в теореме 4 можно довести до $O(n^{-1/4}\log^{3/2} n)$, если использовать результаты работы [10] о скорости сходимости в сильных принципах инвариантности. Полученные результаты сравниваются с предшествующими результатами Н. В. Смирнова ([5], [7]).

Пусть $\mathrm{X}_1, \mathrm{X}_2, \ldots$ — независимые p-мерные случайные векторы, одинаково распределенные с плотностью $f(\mathrm{x})$, которую мы будем предполагать непрерывной на p-мерном интервале $\mathcal{J} = \{a_i \leqslant x_i \leqslant b_i,\ i = \overline{1, p}\}$, причем $\min\limits_{\mathcal{J}} f(\mathrm{x}) > 0$. Пусть $\mathcal{N}(n)$ — пуассоновская случайная величина со средним n, $\mathcal{N}(n)$ не зависит от X_i, $i = 1, 2, 3, \ldots$ В качестве оценок $f(\mathrm{x})$ в точке x рассмотрим последовательность [3]

$$f_n^*(\mathrm{x}) = n^{-1}h_n^{-p}\sum_{j=1}^{\mathcal{N}(n)} K\left(\frac{\mathrm{x}-\mathrm{X}_j}{h_n}\right),$$

где $h_n \to 0$, $nh_n^p \to \infty$ при $n \to \infty$, $K(\mathrm{x})$ — симметричное ядро, равное нулю вне куба $[-1/2, 1/2]^p$, $\int K(\mathrm{x})\,d\mathrm{x} = 1$. Прежде всего нас будет интересовать асимптотика вероятности

$$p_n(\mathrm{x}) = \mathbf{P}\left(\left|\frac{f_n^*(\mathrm{x}) - \mathbf{M}f_n^*(\mathrm{x})}{\sqrt{\mathbf{D}f_n^*(\mathrm{x})}}\right| > \lambda_n\right),$$

где λ_n — некоторая возрастающая с ростом n функция n. Введем обозначения

$$p_n^+(x) = \mathbf{P}\left(\frac{f_n^*(\mathrm{x}) - \mathbf{M}f_n^*(\mathrm{x})}{\sqrt{\mathbf{D}f_n^*(\mathrm{x})}} > \lambda_n\right), \quad p_n^-(\mathrm{x}) = \mathbf{P}\left(\frac{f_n^*(\mathrm{x}) - \mathbf{M}f_n^*(\mathrm{x})}{\sqrt{\mathbf{D}f_n^*(\mathrm{x})}} < -\lambda_n\right).$$

Имеем

$$p_n^+(\mathrm{x}) = \sum_{m=0}^{\infty} \mathbf{P}\left(\frac{f_m(\mathrm{x}) - \mathbf{M}f_m(\mathrm{x})}{\sqrt{\mathbf{D}f_m(\mathrm{x})}} > \lambda_{nm}^+\right)\mathbf{P}(\mathcal{N}(n) = m), \tag{1}$$

где

$$\lambda_{nm}^+ = \left(\lambda_n - \frac{\mathbf{M}f_m(\mathrm{x}) - \mathbf{M}f_n^*(\mathrm{x})}{\sqrt{\mathbf{D}f_n^*(\mathrm{x})}}\right)(\mathbf{D}f_n^*(\mathrm{x}))^{1/2}(\mathbf{D}f_m(\mathrm{x}))^{-1/2},$$

$$f_m(\mathrm{x}) = n^{-1}h_n^{-p}\sum_{j=1}^{m} K\left(\frac{\mathrm{x}-\mathrm{X}_j}{h_n}\right).$$

Величина $p_n^-(\mathrm{x})$ представляется в виде ряда, аналогичного (1). Разобьем сумму (1) на три суммы: от $m = 0$ до $m = m_1(n)$, от $m = m_1(n) + 1$ до $m = m_2(n)$ и от $m = m_2(n) + 1$ до $m = +\infty$ — и займемся сначала изучением вероятности

$$p_{nm}(\mathrm{x}) = \mathbf{P}\left(\left|\frac{f_m(\mathrm{x}) - \mathbf{M}f_m(\mathrm{x})}{\sqrt{\mathbf{D}f_m(\mathrm{x})}}\right| > \lambda_{nm}\right)$$

при $m_1(n) < m \leqslant m_2(n)$, $\lim\limits_{n\to\infty}(m_2(n) - m_1(n)) = \lim\limits_{n\to\infty} m_1(n) = +\infty$.

Нетрудно убедиться в справедливости соотношений

$$\mathbf{D}f_m(\mathbf{x}) = mn^{-2}h_n^{-p}\left(h_n^{-p}\int K^2\left(\frac{\mathbf{x}-\mathbf{u}}{h_n}\right)f(\mathbf{u})\,d\mathbf{u} - n^2h_n^pm^{-2}\mathbf{M}^2f_m(\mathbf{x})\right),$$

$$f_m(\mathbf{x}) - \mathbf{M}f_m(\mathbf{x}) = n^{-1}h_n^{-p/2}\sum_{j=1}^{m}\left(h_n^{-p/2}K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h_n}\right) - \frac{n}{m}\,h_n^{p/2}\mathbf{M}f_m(\mathbf{x})\right),$$

откуда следует представление

$$\frac{f_m(\mathbf{x}) - \mathbf{M}\,f_m(\mathbf{x})}{\sqrt{\mathbf{D}f_m(\mathbf{x})}} = \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\xi_n^{(j)}(\mathbf{x}),$$

где $\xi_n^{(j)}(\mathbf{x})$ $(j = 1, 2, \ldots, m)$ — независимые одинаково распределенные случайные величины, $\mathbf{M}\xi_n^{(j)}(\mathbf{x}) = 0$, $\mathbf{D}\xi_n^{(j)}(\mathbf{x}) = 1$,

$$\xi_n^{(j)}(\mathbf{x}) = \left(h_n^{-p/2}K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h_n}\right) - h_n^{p/2}\mathbf{M}f_n(\mathbf{x})\right)\Big/\left(h_n^{-p}\int K^2\left(\frac{\mathbf{x}-\mathbf{u}}{h_n}\right)f(\mathbf{u})\,d\mathbf{u} - h_n^p\mathbf{M}^2f_n(\mathbf{x})\right)^{1/2}. \quad (2)$$

В дальнейшем нам понадобится лемма, принадлежащая В. А. Статулявичусу [1], которую мы для удобства ссылок приводим ниже.

Лемма 1. *Пусть* ξ *— случайная величина с функцией распределения* F, *со средним* $m = \mathbf{M}\xi$, *дисперсией* $\sigma^2 = \mathbf{D}\xi$ *и конечными моментами любого порядка* $\mathbf{M}|\xi|^k < \infty$, $k = 1, 2, \ldots$ *Пусть* $\mathfrak{S}_k$ *— k-й кумулянт* ξ *и*

$$\Delta = \sigma\inf_k\left(\frac{k!\,H\sigma^2}{|\mathfrak{S}_k|}\right)^{1/(k-2)},$$

где $H > 0$. *Тогда в интервале* $1 \leqslant x \leqslant \bar{\delta}\Delta$, $\bar{\delta} < \bar{\delta}_H$, *имеют место соотношения*

$$\begin{aligned}\frac{1-F(m+x\sigma)}{1-\Phi(x)} &= \exp\left\{\frac{x^3}{\Delta}\mu\left(\frac{x}{\Delta}\right)\right\}\left(1 + f_1(\bar{\delta}, H)\frac{x}{\Delta}\right),\\ \frac{F(m-x\sigma)}{\Phi(-x)} &= \exp\left\{-\frac{x^3}{\Delta}\mu\left(-\frac{x}{\Delta}\right)\right\}\left(1 + f_2(\bar{\delta}, H)\frac{x}{\Delta}\right).\end{aligned} \quad (3)$$

Здесь

$$|f_i(\bar{\delta}, H)| < \frac{8H\{1 + 7{,}2(1 + 2\delta + \min\{1/3(1-\bar{\delta})^3H^{-1}, 1/2H^{-1/4}\})^4\}}{(1-\delta)^4(1-\rho)^{3/2}},$$

$i = 1, 2$, *числа* δ, ρ *и* $\bar{\delta}_H$ *в неравенстве* $\bar{\delta} < \bar{\delta}_H$ *определяются из соотношений*

$$\bar{\delta} = \frac{\delta(1+\delta)}{2}, \quad \rho = \frac{6H\delta}{(1-\delta)^3}, \quad \bar{\delta}_H = \frac{\delta_H(1+\delta_H)}{2},$$

δ_H *— вещественный корень уравнения* $\rho = 1$ *и* $\mu(t) = \sum_{k=0}^{\infty}\mu_k t^k$ *— ряд Крамера, который сходится при* $|t| < \bar{\delta}_H$, *причем*

$$|\mu_k| \leqslant \frac{\delta_H}{(k+3)\,\bar{\delta}_H^{k+2}}, \quad k = 0, 1, \ldots$$

Обозначим через $F_n^{(m)}(y)$ функцию распределения случайной величины $S_{nm}(\mathbf{x}) = m^{-1/2}\sum_{j=1}^{m}\xi_n^{(j)}(\mathbf{x})$ и сформулируем следующую теорему.

Теорема 1. *Пусть*

$$\varliminf_{n\to\infty} \min_{m_1(n)<m\leqslant m_2(n)} \lambda_{nm} > 1, \quad \lim_{n\to\infty} \max_{m_1(n)<m\leqslant m_2(n)} \frac{\lambda_{nm}}{\sqrt{m}\, h_n^{p/2}} = 0,$$
$$\sup_{[-1/2,\, 1/2]^p} |K(\mathrm{x})| < \infty,$$

тогда при $n\to\infty$

$$\text{а)}\quad \frac{1-F_n^{(m)}(\lambda_{nm})}{1-\Phi(\lambda_{nm})} = \exp\left(\frac{\lambda_{nm}^3}{\sqrt{m}}\,\mu_n\left(\frac{\lambda_{nm}}{\sqrt{m}}\right)\right)\left(1+O\left(\frac{\lambda_{nm}}{\sqrt{m}\, h_n^{p/2}}\right)\right),$$
$$\frac{F_n^{(m)}(-\lambda_{nm})}{\Phi(-\lambda_{nm})} = \exp\left(-\frac{\lambda_{nm}^3}{\sqrt{m}}\,\mu_n\left(-\frac{\lambda_{nm}}{\sqrt{m}}\right)\right)\left(1+O\left(\frac{\lambda_{nm}}{\sqrt{n}\, h_n^{p/2}}\right)\right),$$

где $\mu_n(w) = \sum_{k=0}^{\infty} c_{nk} w^k$ *функция, регулярная в окрестности* $|w| < h_n^{p/2}$, *коэффициент ряда* c_{nk} *зависит от кумулянтов* γ_{ln} *случайной величины* $\xi_n^{(1)}(\mathrm{x})$ *до порядка* $k+3$ *включительно.*

Если $\varlimsup_{n\to\infty} \max_{m_1(n)<m\leqslant m_2(n)} \frac{\lambda_{nm}^3}{\sqrt{m}\, h_n^{p/2}} < \infty$, *то при* $n\to\infty$

$$\text{б)}\quad \frac{1-F_n^{(m)}(\lambda_{nm})}{1-\Phi(\lambda_{nm})} = \exp\left(\frac{\lambda_{nm}^3\gamma_{3n}}{6\sqrt{m}}\right)\left(1+O\left(\frac{\lambda_{nm}}{\sqrt{m}\, h_n^{p/2}}\right)\right),$$
$$\frac{F_n^{(m)}(-\lambda_{nm})}{\Phi(-\lambda_{nm})} = \exp\left(-\frac{\lambda_{nm}^3\gamma_{3n}}{6\sqrt{m}}\right)\left(1+O\left(\frac{\lambda_{nm}}{\sqrt{m}\, h_n^{p/2}}\right)\right).$$

Наконец, если $\lim_{n\to\infty} \max_{m_1(n)<m\leqslant m_2(n)} \frac{\lambda_{nm}^3}{\sqrt{m}\, h_n^{p/2}} = 0$, *то при* $n\to\infty$

$$\text{в)}\quad \frac{1-F_n^{(m)}(\lambda_{nm})}{1-\Phi(\lambda_{nm})} = 1+O\left(\frac{\lambda_{nm}^3}{\sqrt{m}\, h_n^{p/2}}\right), \quad \frac{F_n^{(m)}(-\lambda_{nm})}{\Phi(-\lambda_{nm})} = 1+O\left(\frac{\lambda_{nm}^3}{\sqrt{m}\, h_n^{p/2}}\right).$$

Доказательство. Из свойств $K(\mathrm{u})$ и $f(\mathrm{x})$, а также из (2) следует, что равномерно по $\mathrm{x} \in \mathcal{I}_n = \left\{a_i + \frac{h_n}{2} \leqslant x_i \leqslant b_i - \frac{h_n}{2},\ i = \overline{1, p}\right\}$ случайные величины $m^{-1/2}\xi_n^{(j)}(\mathrm{x})$ удовлетворяют условиям С. Н. Бернштейна

$$\left|\mathrm{M}\left(\frac{\xi_n^{(j)}(\mathrm{x})}{\sqrt{m}}\right)^l\right| \leqslant Cl!\,(m^{-1/2}h_n^{-p/2})^{l-2}\, m^{-1}, \quad j=1,2,\ldots,m,\ l\geqslant 3,$$

где C — абсолютная постоянная. Тогда, как показано в [1] (теорема 2), соотношения (3) леммы 1 имеют место с $m=0$, $F(y) = F_n^{(m)}(y)$, $\sigma = 1$, $H = 3/2$ и

$$\Delta = \{\max\,(h_n^{-p/2}m^{-1/2}(1+2C), \sqrt{2})\}^{-1} = C_1 h_n^{p/2} m^{1/2}.$$

Положим $\overline{\delta} = \overline{\delta}_n = \max_{m_1(n)<m\leqslant m_2(n)} \frac{\lambda_{nm}}{C_1 h_n^{p/2}\sqrt{m}}$. Итак, в интервале $1 \leqslant \lambda_{nm} \leqslant \overline{\delta}_n\Delta$, $\overline{\delta}_n < \overline{\delta}_{3/2}$, имеют место соотношения (3). Для того чтобы показатель экспоненты в (3) привести к нужному нам виду, заметим, что

$$\frac{\lambda_{nm}^3}{C_1 h_n^{p/2}\sqrt{m}}\,\mu\left(\frac{\lambda_{nm}}{C_1 h_n^{p/2}\sqrt{m}}\right) = \frac{\lambda_{nm}^3}{\sqrt{m}}\,\mu_n\left(\frac{\lambda_{nm}}{\sqrt{m}}\right),$$

где ряд $\mu_n(w)$ сходится при $|w| < C_1 h_n^{p/2} \overline{\delta}_{3/2}$. Из [1] (соотношение (43)) и теоремы Лагранжа об обращении [2] следует, что l-й коэффициент c_{nl} ряда $\mu_n(w)$ дается формулой

$$c_{nl} = \mu_l (C_1 h_n^{p/2})^{-(l+1)} = -\frac{a_{l+2}}{l+3} m^{(l+1)/2},$$

где

$$a_l = \frac{1}{l!} \lim_{z\to 0} \frac{d^{l-1}}{dz^{l-1}} \left\{ z^k \left(z + \sum_{j=3}^{\infty} \frac{\Gamma_{jnm}}{(j-1)!} z^{j-1} \right)^{-k} \right\}. \tag{4}$$

Здесь Γ_{jnm} — l-й кумулянт суммы S_{nm} (аргумент x мы, во избежание громоздких обозначений, опускаем). Пусть $\tilde{a}_l$ получаются из a_l заменой Γ_{jnm} на γ_{jn} в соотношении (4). Тогда из (4) и равенства $\Gamma_{jnm} = \gamma_{jn} m^{-(j/2-1)}$ следует, что $a_l = m^{-(l-1)/2} \tilde{a}_l$, $c_{nl} = -\tilde{a}_{l+2}/(l+3)$. Легко проверить, что

$$c_{n0} = \frac{\gamma_{3n}}{6}, \quad c_{n1} = -\frac{\gamma_{4n} - 3\gamma_{3n}^3}{24}, \quad c_{n2} = \frac{\gamma_{5n} - 10\gamma_{4n}\gamma_{3n} + 15\gamma_{3n}^3}{120}, \ldots$$

Вообще, c_{nl} зависит от кумулянтов γ_{ln} случайной величины $\xi_n^{(1)}(\mathrm{x})$ до порядка $l+3$ включительно. Утверждение а) теоремы доказано. Для доказательства б) достаточно заметить, что

$$\left| \sum_{k=1}^{\infty} c_{nk} \left(\frac{\lambda_{nm}}{\sqrt{m}} \right)^k \frac{\lambda_{nm}^3}{\sqrt{m}} \right| \leqslant$$

$$\leqslant \frac{\lambda_{nm}^3 \delta_{3/2}}{C_1 \sqrt{m} h_n^{p/2} \delta_{3/2}^2} \sum_{k=1}^{\infty} (k+3)^{-1} \left(\frac{\lambda_{nm}}{C_1 \delta_{3/2} \sqrt{m} h_n^{p/2}} \right)^k \leqslant C_2 \frac{\lambda_{nm}}{\sqrt{m} h_n^{p/2}}.$$

Наконец, в) очевидным образом следует из б), если учесть, что $|\gamma_{3n}| \leqslant \leqslant 9/C_1 h_n^{p/2}$. Теорема доказана.

З а м е ч а н и е 1. Мы получили явную оценку постоянных в остаточных членах соотношений а), б), в) теоремы 1 как следствие подобных оценок в соотношениях (3). Утверждение теоремы останется справедливым, если первое из условий теоремы заменить условием

$$\varliminf_{n\to\infty} \min_{m_1(n) < m \leqslant m_2(n)} \lambda_{nm} > 0.$$

З а м е ч а н и е 2. Оценки остаточных членов в теореме 1 равномерны по $\mathrm{x} \in \mathcal{J}_n$.

Пусть $a_0, a_1, a_2, \ldots$ последовательность вещественных чисел. Под переходом от этой последовательности к последовательности ее частичных сумм будем понимать переход к последовательности

$$a_0, \; a_0 + a_1, \; a_0 + a_1 + a_2, \ldots$$

Разделим каждое ребро p-мерного интервала $\mathcal{J}$ на s_n равных частей и образуем s_n^p p-мерных интервалов Δ_{ni}, $i = 1, 2, \ldots, s_n^p$. Выберем s_n так, чтобы $s_n \to \infty$, $s_n^p n^{-1} \to 0$ при $n \to \infty$, и проведем через точки деления гиперплоскости, параллельные соответствующим граням интервала $\mathcal{J}$. Интервалы Δ_{ni} не пересекаются и их объединение есть $\mathcal{J}$. Пусть x_i $(i = 1, 2, \ldots, s_n^p)$ — центры интервалов Δ_{ni}.

Лемма 2. *Положим* $\lambda_n = l_n + y l_n^{-1}$, *где* l_n — *корень уравнения*

$$s_n^{-p} = (2\pi)^{-1/2} x^{-1} \exp(-x^2/2),$$

а $y \in [-t_n, t_n]$, $\varliminf_{n\to\infty} (l_n^2 - t_n)\, l_n^{-1} > 0$. *Предположим, что существует такое* $\delta > 0$, *что*

$$\lim_{n\to\infty} n^{\delta} h_n^p = 0, \quad \lim_{n\to\infty} n^{1-\delta} h_n^p = \infty. \tag{5}$$

Тогда имеет место асимптотическое разложение

$$\sum_{i=1}^{s_n^p} p_n(\mathbf{x}_i) \sim 2e^{-y} \sum_{j=0}^{\infty} \frac{(-1)^j}{l_n^{2j}} \left(\sum_{i=0}^{j} \frac{\Gamma((i+1)/2)\, 2^i P_{2j-2i}^{(2i+1)}(y)}{\Gamma(1/2)} \right), \tag{6}$$

где

$$P_0^{(2i+1)}(y) \equiv 1, \quad P_{2l}^{(2i+1)}(y) = \sum_{k=0}^{l} \frac{b_k^{(2i+1)} y^{2l-k}}{2^{l-k}(l-k)!}, \quad l = 1, 2, \ldots,$$

а $\{b_k^{(2i+1)}\}$ — *последовательность, полученная* $2i$-*кратным применением перехода от* $\{1, 1, 1, \ldots\}$ *к последовательности частичных сумм.*

Доказательство. Фиксируем α, $1/2 < \alpha < \min\{1, 1/2 + \delta\}$, и рассмотрим

$$\sum_{m=n-[n^{\alpha}]}^{n+[n^{\alpha}]} \left\{ \mathbf{P}\left(\frac{f_m(\mathbf{x}) - \mathbf{M} f_m(\mathbf{x})}{\sqrt{\mathbf{D} f_m(\mathbf{x})}} > \lambda_{nm}^{+} \right) + \mathbf{P}\left(\frac{f_m(\mathbf{x}) - \mathbf{M} f_m(\mathbf{x})}{\sqrt{\mathbf{D} f_m(\mathbf{x})}} < -\lambda_{nm}^{-} \right) \right\} \times$$
$$\times \mathbf{P}(\mathcal{N}(n) = m), \tag{7}$$

где

$$\lambda_{nm}^{\pm} = \left(\lambda_n \mp \frac{\mathbf{M} f_m(\mathbf{x}) - \mathbf{M} f_n^*(\mathbf{x})}{\sqrt{\mathbf{D} f_n^*(\mathbf{x})}} \right) (\mathbf{D} f_n^*(\mathbf{x}))^{1/2} (\mathbf{D} f_m(\mathbf{x}))^{-1/2}.$$

Простой подсчет показывает, что

$$\mathbf{M} f_n^*(\mathbf{x}) = \mathbf{M} f_n(\mathbf{x}), \quad \mathbf{D} f_n^*(\mathbf{x}) = \mathbf{D} f_n(\mathbf{x}) + n^{-1} (\mathbf{M} f_n(\mathbf{x}))^2,$$

и для $\lambda_{nm}^{\pm}$ получаем

$$\lambda_{nm}^{\pm} = (\lambda_n \pm (1 - m/n) \sqrt{n h_n^p}\, \Theta_1(\mathbf{x}))(1 + h_n^p \Theta_2(\mathbf{x})) \sqrt{n/m}, \tag{8}$$

где $0 < c < \Theta_i(x) < C < \infty$, $i = 1, 2$, $\mathbf{x} \in \mathcal{I}$, при некоторых c и C. Для l_n верно соотношение ([8], стр. 410)

$$l_n = \sqrt{2p \ln s_n} - \frac{\ln(p \ln s_n) + \ln 4\pi}{2\sqrt{2p \ln s_n}} + O\left(\frac{1}{\log s_n} \right). \tag{9}$$

Отсюда и из условий леммы получим, что при $y \in [-t_n, t_n]$

$$\varliminf_{n\to\infty} \lambda_n > 0, \quad \lim_{n\to\infty} \frac{\lambda_n^3}{\sqrt{n}\, h_n^{p/2}} = 0. \tag{10}$$

Из (8) и (10) следует, что при $m \in [n - n^{\alpha}, n + n^{\alpha}]$, $n \to \infty$

$$\lambda_{nm} = \lambda_n \, (1 + O(n^{\alpha - 1/2} h_n^{p/2})), \tag{11}$$

где постоянная в $O(\cdot)$ единая для всех $m \in [n - n^\alpha,\ n + n^\alpha]$ и $\mathrm{x} \in \mathcal{J}$; для $\lambda_{nm}^{\pm}$ имеют место соотношения в) теоремы 1, поэтому (7) приводится к виду

$$\left(1 + O\left(\frac{\lambda_n^3}{\sqrt{n}\, h_n^{p/2}}\right)\right) \sum_{m=n-[n^\alpha]}^{n+[n^\alpha]} [(1 - \Phi(\lambda_{nm}^{+})) + (1 - \Phi(\lambda_{nm}^{-}))]\ \mathbf{P}\ (\mathcal{N}(n) = m). \tag{12}$$

Используя асимптотическое разложение функции ошибок ([4], стр. 113), получим

$$1 - \Phi(\lambda_{nm}^{\pm}) \sim \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\lambda_{nm}^{\pm})^2}{2}\right) \sum_{i=0}^{\infty} \frac{(-1)^i\, \Gamma((i+1)/2)\, 2^i}{\Gamma(^1/_2)\, (\lambda_{nm}^{\pm})^{2i+1}}. \tag{13}$$

При нашем выборе α и λ_n при любом k имеем $n^{\alpha-1/2} h_n^{p/2} = o\,(\lambda_n^{-k})$, $n \to \infty$, поэтому, согласно (11), при $m \in [n - n^\alpha,\ n + n^\alpha]$ в разложении (13) последовательности

$$\{\varphi_i^{\pm}\} = \left\{\exp\left(-\frac{(\lambda_{nm}^{\pm})^2}{2}\right)(\lambda_{nm}^{\pm})^{-2i-1}\right\}$$

можно заменить последовательностью

$$\{\varphi_i\} = \left\{\exp\left(-\frac{\lambda_n^2}{2}\right) \lambda_n^{-2i-1}\right\}. \tag{14}$$

Простой подсчет с использованием формулы Стирлинга показывает, что

$$\sum_{m=n-[n^\alpha]}^{n+[n^\alpha]} \mathbf{P}\,\{\mathcal{N}(n) = m\} = 1 - o\,(\exp(-qn^{2\alpha-1})), \quad q > 0, n \to \infty, \tag{15}$$

и из (12), (13), (14), (15) и определения l_n следует

$$\sum_{i=1}^{s_n^p} p_n(\mathrm{x}_i) \sim 2e^{-y} e^{-y^2/2l_n^2} \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma((i+1)/2)\, 2^i}{\Gamma(^1/_2)\, l_n^{2i}\, (1 + y/l_n^2)^{2i+1}}. \tag{16}$$

Разлагая в ряд экспоненту и перемножая геометрические прогрессии, нетрудно убедиться по индукции, что

$$\exp\left(-\frac{y}{2l_n^2}\right)\left(1 + \frac{y}{l_n^2}\right)^{-2i-1} = \sum_{j=0}^{\infty} \frac{(-1)^j P_{2j}^{(2i+1)}(y)}{l_n^{2j}}. \tag{17}$$

Подставляя (17) в (16), получим (6), что завершает доказательство леммы.

Замечание 3. Условия (5) можно несколько ослабить, например, вместо второго из этих условий можно потребовать, чтобы для любого k

$$\frac{\ln^k n}{\sqrt{n}\, h_n^{p/2}} = o(1), \quad n \to \infty.$$

Следующая лемма относится к аппроксимации биномиального распределения пуассоновским. Пусть имеется последовательность независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ таких, что

$$\mathbf{P}\ (\xi_k = 1) = p_k, \quad \mathbf{P}\ (\xi_k = 0) = 1 - p_k.$$

Введем

$$S_n = \sum_{j=1}^{n} \xi_j, \quad \mu = \sum_{j=1}^{n} p_j, \quad \pi_k = \frac{\mu^k e^{-\mu}}{k!}.$$

Лемма 3. *Для всех множеств* $B \subseteq \{0, 1, 2, \ldots\}$

$$\left| \mathbf{P}(S_n \in B) - \sum_{k \in B} \pi_k \right| \leqslant \sum_{j=1}^{n} p_j^2 \leqslant \mu \max_{1 \leqslant j \leqslant n} p_j.$$

Доказательство леммы 3 содержится в книге А. А. Боровкова ([9], стр. 117). Лемма полезна для оценки $\mathbf{P}(S_n = k)$, когда p_j малы, а μ «сравнимо с 1».

Для дальнейшего удобно ввести обозначения

$$M_{1n}^* = \max_{1 \leqslant i \leqslant s_n^p} \left| \frac{f_n^*(\mathbf{x}_i) - \mathbf{M} f_n^*(\mathbf{x}_i)}{\sqrt{\mathbf{D} f_n^*(\mathbf{x}_i)}} \right|, \quad M_{2n} = \max_{1 \leqslant i \leqslant s_n^p} \left| \frac{f_n(\mathbf{x}_i) - \mathbf{M} f_n(\mathbf{x}_i)}{\sqrt{\mathbf{D} f_n^*(\mathbf{x}_i)}} \right|,$$

$$\eta_n(\mathbf{x}_i) = \left| \frac{f_n^*(\mathbf{x}_i) - \mathbf{M} f_n^*(\mathbf{x}_i)}{\sqrt{\mathbf{D} f_n^*(\mathbf{x}_i)}} \right| - \left| \frac{f_n(\mathbf{x}_i) - \mathbf{M} f_n(\mathbf{x}_i)}{\sqrt{\mathbf{D} f_n^*(\mathbf{x}_i)}} \right|.$$

Теорема 2. *Пусть выполнены условия леммы 2, а также*

$$\tau_{ni} \subseteq \Delta_{ni}, \quad i = 1, 2, \ldots, s_n^p, \tag{18}$$

где τ_{ni} *— p-мерный куб с ребром* h_n *и центром* $\mathbf{x}_i$. *Тогда*

$$\mathbf{P}\left(M_{1n}^* < l_n + \frac{y}{l_n}\right) \quad \text{и} \quad \exp\left(-\sum_{i=1}^{s_n^p} p_n(\mathbf{x}_i)\right)$$

как функции y представляются одним и тем же асимптотическим рядом при $n \to \infty$

$$\exp(-2e^{-y}) \left(1 + \sum_{j=1}^{\infty} \sum_{k=1}^{j} \frac{(-1)^{j-k} e^{-ky} Q_{j-k+1}^{(k)}(y)}{k!} l_n^{-2j}\right), \tag{19}$$

где $Q_j^{(1)}(y)$ *— полином степени* $2j$:

$$Q_j^{(1)}(y) = \sum_{i=0}^{j} \sum_{k=0}^{j-i} \frac{2^{2i-j+k+1} \Gamma((i+1)/2)\, b_k^{(2i+1)} y^{2j-2i-k}}{\Gamma(1/2)(j-i-k)!},$$

а полиномы $Q_j^{(k)}(y)$, $k > 1$, *определяются соотношениями*

$$\left(\sum_{j=1}^{\infty} \frac{Q_j^{(1)}(y)}{l_n^{2j}}\right)^k = \sum_{j=1}^{\infty} \frac{Q_j^{(k)}(y)}{l_n^{2(k+j-1)}}.$$

Доказательство. Из элементарных свойств асимптотических рядов ([2], стр. 26) и (6) следует, что

$$\exp\left(-\sum_{i=1}^{s_n^p} p_n(\mathbf{x}_i)\right) \sim \exp(-2e^{-y}) \left(1 + \sum_{k=0}^{\infty} \frac{\Delta_n^k}{k!}\right),$$

где $\Delta_n = e^{-y}\sum\limits_{j=1}^{\infty}(-1)^{j+1}Q_j^{(1)}(y)/l_n^{2j}$. Собирая подобные члены, получим асимптотическое разложение (19). Рассмотрим независимые, в силу (18), случайные величины $\xi_1, \ldots, \xi_{s_n^p}$:

$$\xi_i = \left| \frac{f_n^*(\mathbf{x}_i) - \mathbf{M}f_n^*(\mathbf{x}_i)}{\sqrt{\mathbf{D}f_n^*(\mathbf{x}_i)}} \right|,$$

и связанные с ними независимые случайные величины $\eta_1, \eta_2, \ldots, \eta_{s_n^p}$, где $\eta_i(\omega)$ — индикатор множества $\{\omega: \xi_i \geqslant l_n + y/l_n\}$. Из того, что остаточные члены разложений леммы 2 оценивались равномерно по $\mathbf{x} \in \mathcal{I}_n$, и из (6) получим при $n \to \infty$

$$\sum_{i=1}^{s_n^p} p_n(\mathbf{x}_i) = 2e^{-y}(1 + O(l_n^{-2})), \qquad \max_{1 \leqslant i \leqslant s_n^p} p_n(\mathbf{x}_i) = \frac{2e^{-y}}{s_n^p}(1 + O(l_n^{-2})).$$

Таким образом, по лемме 3 при $n \to \infty$

$$\left|\mathbf{P}\left(\sum_{i=1}^{s_n^p}\eta_i = 0\right) - \exp\left(-\sum_{i=1}^{s_n^p}p_n(\mathbf{x}_i)\right)\right| = O(s_n^{-p}e^{-2y}). \tag{20}$$

Разность (20) есть $o(l_n^{-k})$ для любого k, поэтому асимптотические разложения $\mathbf{P}\left(\sum\limits_{i=1}^{s_n^p}\eta_i = 0\right)$ и $\exp\left(-\sum\limits_{i=1}^{s_n^p}p_n(\mathbf{x}_i)\right)$ одинаковы. Теорема доказана.

Покажем теперь, как с помощью теоремы 2 получить асимптотическое разложение в случае выборки неслучайного объема n, $n \to \infty$. Для этого нам понадобится лемма, в которой дана оценка близости $f_n^*(\mathbf{x})$ к $f_n(\mathbf{x})$. Отметим, что в [3] при выводе порядка убывания $\mathbf{D}(f_n^*(\mathbf{x}) - f_n(\mathbf{x}))$ допущены ошибки.

Лемма 4. *Существует постоянная $C(q)$ такая, что при* $\mathbf{x} \in \mathcal{I}$, $\varepsilon > 0$ *и* $q = 1, 2, \ldots$:

$$\mathbf{P}(|f_n(\mathbf{x}) - f_n^*(\mathbf{x})| > \varepsilon) \leqslant C(q)\varepsilon^{-2q}\left(\sum_{l=1}^{q} n^{l/2-2q}h_n^{p(l-2q)} + n^{-q}\right).$$

Доказательство. Центрируя $f_n(\mathbf{x})$ математическим ожиданием $\mathbf{M}f_n(\mathbf{x})$, а $f_n^*(\mathbf{x})$ — условным математическим ожиданием $\mathbf{M}(f_n^*(\mathbf{x}) \mid \mathcal{N})$, получим неравенство

$$\mathbf{P}(|f_n(\mathbf{x}) - f_n^*(\mathbf{x})| > \varepsilon) \leqslant$$

$$\leqslant \mathbf{P}\left(n^{-1}h_n^{-p}\left|{\sum_{i=\mathcal{N}}^{n}}' \alpha_i(\mathbf{x})\right| > \frac{\varepsilon}{2}\right) + \mathbf{P}\left(n^{-1}h_n^{-p}|\mathcal{N} - n|\,\mathbf{M}K\left(\frac{\mathbf{x}-\mathbf{X}}{h_n}\right) > \frac{\varepsilon}{2}\right),$$

где под ${\sum\limits_{i=\mathcal{N}}^{n}}'$ мы понимаем $\sum\limits_{i=1}^{n} - \sum\limits_{i=1}^{\mathcal{N}}$, а

$$\alpha_i(\mathbf{x}) = K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right) - \mathbf{M}K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right).$$

Используя разложение для характеристической функции $\mathcal{N}-n$

$$\mathrm{M}\exp[it(\mathcal{N}-n)]=\exp\left[\frac{(it)^2}{2!}n+\frac{(it)^3}{3!}n+\ldots\right]$$

и связь между моментами и семиинвариантами случайной величины $\mathcal{N}-n$ ([6]), получим оценки моментов случайной величины $\mathcal{N}-n$:

$$\mathrm{M}\,|\,\mathcal{N}-n\,|^k\leqslant C\,(k)\;n^{k/2}. \tag{21}$$

Рассмотрим сначала случай $q\leqslant|\,\mathcal{N}-n\,|$. Тогда получим

$$\mathrm{M}\,|\,n^{-1}h_n^{-p}\sum_{i=m}^{n}{}'\alpha_i(\mathbf{x})\,|^{2q}=$$

$$=n^{-2q}h_n^{-2pq}\sum_{j_1+\ldots+j_{n-m}=2q}\frac{2q!}{j_1!\ldots j_{n-m}!}\mathrm{M}\alpha_1^{j_1}(\mathbf{x})\ldots\mathrm{M}\alpha_1^{j_{n-m}}(\mathbf{x})=$$

$$=n^{-2q}h_n^{-2pq}\sum_{l=1}^{q}C_{n-m}^{l}\sum_{l;\,j_1+\ldots+j_l=2q}\frac{2q!}{j_1!\ldots j_l!}\mathrm{M}\alpha_1^{j_1}(\mathbf{x})\ldots\mathrm{M}\alpha_1^{j_l}(\mathbf{x}),$$

где через $\sum\limits_{l;\,j_1+\ldots+j_l=2q}$ обозначено суммирование по $j_k\neq0,1,\ \sum\limits_{k=1}^{l}j_k=2q$. При $\mathcal{N}=m>n$ имеет место то же соотношение с заменой $n-m$ на $m-n$. Учитывая, что при $n\to\infty$

$$\mathrm{M}\alpha_1^{j_1}(\mathbf{x})=h_n^pf(\mathbf{x})\int K^{j_1}(\mathbf{u})\,d\mathbf{u}\,(1+o(1)),$$

имеем

$$\mathrm{M}\left|n^{-1}h_n^{-p}\sum_{i=m}^{n}{}'\alpha_i(\mathbf{x})\right|^{2q}\leqslant C\,(q)\,n^{-2q}h_n^{-2pq}\sum_{l=1}^{q}|\,m-n\,|^l\,h_n^{pl}. \tag{22}$$

Очевидно, и в случае $|\,\mathcal{N}-n\,|<q$ неравенство (22) имеет место, ибо его левая часть не превосходит $C\,(q)\;n^{-2q}\;h_n^{-p(2q-1)}$.

Переходя к безусловным математическим ожиданиям, получим из (21) и (22)

$$\mathrm{M}\left|n^{-1}h_n^{-p}\sum_{i=\mathcal{N}}^{n}{}'\alpha_i(\mathbf{x})\right|^{2q}\leqslant C\,(q)\sum_{l=1}^{q}n^{l/2-2q}h_n^{p(l-2q)},$$

$$\mathrm{M}\left|n^{-1}h_n^{-p}\,|\,\mathcal{N}-n\,|\,\mathrm{M}K\left(\frac{\mathbf{x}-\mathbf{X}}{h_n}\right)\right|^{2q}\leqslant C\,(q)\,n^{-q}.$$

Применением неравенств Маркова завершается доказательство леммы.

Следующая лемма проста, и ее доказательство мы опускаем.

Лемма 5. *Пусть* $\mathbf{P}\,(|\,\xi-\eta\,|>\varepsilon)<\delta$. *Тогда*

$$\mathbf{P}\,(\eta<y-\varepsilon)-\delta\leqslant\mathbf{P}\,(\xi<y)\leqslant\mathbf{P}\,(\eta<y+\varepsilon)+\delta.$$

Теорема 3. *Пусть выполнены условия леммы* 2 *и условие* (18). *Тогда для* $\mathbf{P}\,(M_{2n}<l_n+yl_n^{-1})$ *справедливо асимптотическое разложение* (19).

Доказательство. Предположим, что мы нашли $\delta_1>0$ и $\delta_2>0$, для которых

$$\mathbf{P}\,(l_n\,|\,M_{1n}^*-M_{2n}\,|>n^{-\delta_1})\leqslant n^{-\delta_2}. \tag{23}$$

Тогда по лемме 5

$$\mathbf{P}\,(l_nM_{1n}^*-l_n^2<y-n^{-\delta_1})-n^{-\delta_2}\leqslant\mathbf{P}\,(l_nM_{2n}-l_n^2\leqslant y)\leqslant$$

$$\leqslant\mathbf{P}\,(l_nM_{1n}^*-l_n^2<y+n^{-\delta_1})+n^{-\delta_2}. \tag{24}$$

Применим теорему 2 к левой и правой частям неравенств (24). Поскольку при любом m имеем $n^{-\delta_1} = o\,(l_n^{-m})$, $n \to \infty$, то разложение (19) в точках $y \mp n^{-\delta_1}$ можно заменить разложением в точке y, так как погрешность уйдет в остаточный член. То же касается и добавок $\mp n^{-\delta_2}$. Мы показали, что слева и справа в (24) стоят функции, допускающие разложение (19), тогда то же верно и для

$$\mathbf{P}\,(l_n M_{2n} - l_n^2 < y).$$

Осталось показать, что имеет место (23). Поскольку

$$\mathbf{P}\,(l_n\,|\,M_{1n}^* - M_{2n}\,| > n^{-\delta_1}) \leqslant \sum_{i=1}^{s_n^p} \mathbf{P}\Big(|\,\eta_n(\mathrm{x}_i)\,| > \frac{1}{2}\, l_n^{-1} n^{-\delta_1}\Big) \leqslant$$

$$\leqslant \sum_{l=1}^{s_n^p} \mathbf{P}\,\Big(|\,f_n^*(\mathrm{x}_i) - f_n(\mathrm{x}_i)\,| > \frac{\sigma}{2}\, l_n^{-1} n^{-(1/2+\delta_1)} h_n^{-p/2}\Big),$$

где $\sigma = \inf\limits_{\mathscr{Y}}\,(nh_n^p\,\mathrm{D}f_n^*(\mathrm{x}))^{1/2}$, то можно воспользоваться оценкой леммы 4 с

$$\varepsilon_n = \frac{\sigma}{2}\, l_n^{-1} n^{-(1/2+\delta_1)} h_n^{-p/2}.$$

Из условия (18) следуют неравенства

$$\begin{gathered} s_n^p \varepsilon_n^{-2q} n^{l/2-2q} h_n^{p(l-2q)} \leqslant C\,(q)\, l_n^{2q}\,(nh_n^p)^{-(q-l+1)} n^{1-l/2+2q\delta_1}, \\ s_n^p \varepsilon_n^{-2q} n^{-q} \leqslant C\,(q)\, l_n^{2q} n^{2q\delta_1} h_n^{p(q-1)}. \end{gathered} \tag{25}$$

Перепишем правые части неравенств (25) в тождественном виде:

$$C\,(q)\, l_n^{2q}\,(n^{1-\tilde\delta} h_n^p)^{-(q-l+1)} n^{-(1/2-\tilde\delta)(l-1)} n^{-2q(\tilde\delta/2-(4q)^{-1}-\delta_1)},\ l = 1, 2, \ldots, q, \tag{26}$$

и

$$C\,(q)\, l_n^{2q}\,(n^{\tilde\delta} h_n^p)^{q-1} n^{-2q(\tilde\delta/2-(4q)^{-1}-\delta_1)} n^{-(1/2-\tilde\delta)}, \tag{27}$$

где $0 < \tilde\delta < \delta$, δ из леммы 2. Выберем q столь большим, а δ_1 столь малым, чтобы выполнялось неравенство

$$0 < \delta_1 < \tilde\delta/2 - (4q)^{-1}. \tag{28}$$

Из (26), (27) и (28) следует, что выбор $\delta_2 = 2q\,(\tilde\delta/2 - (4q)^{-1} - \delta_1)$ влечет (23). Теорема доказана.

Пусть $y \in [-t_n,\ t_n]$, где $t_n = \ln l_n^2$. Для функции $\exp\,(-2e^{-y})$ в точках $y = \pm\, t_n$ имеем

$$\exp\,(-\,2e^{t_n}) = e^{-2l_n^2}, \quad \exp\,(-\,2e^{-t_n}) = 1 - O\,(l_n^{-2}),\ n \to \infty.$$

Внутри сегмента $[-t_n,\ t_n]$ следующий после $\exp\,(-2e^{-y})$ член разложения (19) имеет вид

$$l_n^{-2} \exp\,(-\,2e^{-y})\, e^{-y}\,(y^2 + 2y + 4/\sqrt{\pi}),$$

т. е. разность $\mathbf{P}\,(M_{2n} < l_n + y/l_n) - \exp\,(-2e^{-y})$ снова есть $O\,(l_n^{-2})$ при $n \to \infty$. Получаем

Следствие. *Существуют постоянные $C_1 > 0$ и $C_2 < \infty$ такие, что при достаточно больших n*

$$C_1 \log^{-1} s_n < \sup_{-\infty<y<\infty} \left| \mathbf{P}\left(M_{2n} < l_n + \frac{y}{l_n}\right) - \exp(-2e^{-y}) \right| < C_2 \log^{-1} s_n.$$

Аналогичное утверждение верно для M_{1n}^*.

Сделаем несколько замечаний по поводу условий теоремы 3. Если, например, $h_n = n^{-\beta}$, $0 < \beta < 1/p$, то условия теоремы выполнены с любым δ, $0 < \delta < \min(p\beta, 1 - p\beta)$. При введении оценок $f_n(x)$ всегда требуют, чтобы $nh_n^p \to \infty$ при $n \to \infty$, это условие необходимо для стремления к нулю дисперсии оценки $f_n(x)$. Следовательно, условие $n^{1-\delta}h_n^p \to \infty$ при $n \to \infty$ близко к необходимому. Если бы мы хотели получить лишь предельное распределение, условия (5) можно было бы ослабить, при этом условие $n^\delta h_n^p \to 0$ окажется излишним.

Н. В. Смирновым [5], [7] была доказана следующая теорема.

Теорема Н. В. Смирнова. *Если непрерывная плотность распределения $f(x)$ удовлетворяет на $[a, b]$ условиям*

$$\min f(x) = \mu > 0 \quad (a \leqslant x \leqslant b), \tag{A}$$

$$\int_a^b f(x)\,dx = 1 - \alpha < 1 \tag{B}$$

и при возрастании s_n

$$\overline{\lim_{n\to\infty}}\, n^{-1} s_n^3 (\ln s_n)^3 < \infty, \tag{C}$$

то для любого λ, $-\infty < \lambda < \infty$, при $n \to \infty$

$$\mathbf{P}\left(\max_{1\leqslant i\leqslant s_n} \left| \frac{f_n^*(x_i) - \bar f(x_i)}{\sqrt{\bar f(x_i)}} \right| < \left(l_n' + \frac{\lambda}{l_n'}\right)(nh_n)^{-1/2}\right) =$$
$$= \left(2\Phi\left(l_n' + \frac{\lambda}{l_n'}\right) - 1\right)^{s_n} + O\left(\frac{\sqrt{\ln s_n}}{s_n}\right) + O\left(\frac{s_n^{3/2}}{\sqrt{n}}\right) =$$
$$= \exp(-2e^{-\lambda}) + O\left(\frac{1}{\sqrt{\ln s_n}}\right). \tag{29}$$

Здесь $h_n = (b - a)/s_n$, $f_n^*(x_i) = m_i/nh_n$, m_i — число наблюдений, попавших в i-й отрезок Δ_{ni},

$$\bar f(x_i) = h_n^{-1} \int_{\Delta_{ni}} f(x)\,dx, \qquad \Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} \exp(-y^2/2)\,dy,$$

l_n' — корень уравнения $s_n^{-1} = 1 - \Phi(l)$.

Если в теореме 3 выбрать $K(x) = \chi[-1/2, 1/2]$, $p = 1$, то в качестве следствия получим усиление теоремы Н. В. Смирнова, состоящее, во-первых, в отбрасывании условия (B) и значительном ослаблении условия (C) и, во-вторых, в получении асимптотического разложения вместо оценки $O(1/\sqrt{\ln s_n})$ остаточного члена в (29). Заметим, что рассуждения из [5] на стр. 218 допускают уточнение, в результате которого остаточный член в (29) есть $O(1/\ln s_n)$, т. е. совпадает с порядком главного члена разложения (19). Отметим также, что наше определение l_n несколько отличается от определения l'_n у Н. В. Смирнова, однако легко показать, что при $n \to \infty$

$$|l_n - l'_n| = O(\ln^{-3/2} s_n).$$

Как видно из (19), если ограничиться k членами ряда (19), то остаточный член имеет порядок $O(\ln^{-k} s_n)$. Практический интерес представляет нахождение более точных аппроксимаций.

Теорема 4. *Пусть выполнены условия теоремы* 3 *и* $s_n h_n \equiv C$. *Тогда для любого* $\varepsilon > 0$ *можно указать* $\delta(\varepsilon)$ *такое, что при выборе* $s_n = n^{1/2+\delta(\varepsilon)}$, $n \to \infty$

$$\mathbf{P}\left(M_{2n} < l_n + \frac{y}{l_n}\right) = \left(2\Phi\left(l_n + \frac{y}{l_n}\right) - 1\right)^{s_n^p} + O(n^{-1/4+\varepsilon}).$$

Аналогичное утверждение верно для M_{1n}^*.

Доказательство. Согласно (20) имеем

$$\mathbf{P}\left(M_{1n}^* < l_n + \frac{y}{l_n}\right) = \exp\left(-\sum_{i=1}^{s_n^p} p_n(\mathbf{x}_i)\right) + O(s_n^{-p}). \tag{30}$$

Из определения l_n, (9), (11) и (13) следует, что

$$1 - \Phi(\lambda_{nm}^{\pm}) = 1 - \Phi(\lambda_n) + O(s_n^{-p} n^{\alpha-1/2} h_n^{p/2} \log s_n). \tag{31}$$

Из (1), (12), (15) и (31)

$$p_n(\mathbf{x}_i) = 2 - 2\Phi(\lambda_n) + O\left(\frac{\lambda_n^3}{\sqrt{n}\, h_n^{p/2} s_n^p}\right) + O(s_n^{-p} n^{\alpha-1/2} h_n^{p/2} \log s_n),$$

$$\begin{aligned}\exp\left(-\sum_{i=1}^{s_n^p} p_n(\mathbf{x}_i)\right) &= (2\Phi(\lambda_n) - 1)^{s_n^p} + O(s_n^{-p}) + \\ &+ O(n^{-1/2} h_n^{-p/2} \log^{3/2} s_n) + O(n^{\alpha-1/2} h_n^{p/2} \log s_n).\end{aligned} \tag{32}$$

Подставляя (32) в (30), легко показать, что при условии $h_n s_n = C$ остаточный член будет иметь минимальный порядок при $s_n = n^{\alpha/p}$; выбирая α близким к $1/2$, можно добиться оценки $O(n^{-1/4+\varepsilon})$, где $\varepsilon > 0$ — любое наперед заданное. Для перехода от M_{1n}^* к M_{2n} заметим, что если имеет место (23) при некоторых δ_1 и δ_2, то из (24), (30) и (32) и того, что

$$\left(2\Phi\left(l_n + \frac{y \pm n^{-\delta_1}}{l_n}\right) - 1\right)^{s_n^p} = \left(2\Phi\left(l_n + \frac{y}{l_n}\right) - 1\right)^{s_n^p} + O(n^{-\delta_1}),$$

следует

$$\begin{aligned}\mathbf{P}\left(M_{2n} < l_n + \frac{y}{l_n}\right) &= \left(2\Phi\left(l_n + \frac{y}{l_n}\right) - 1\right)^{s_n^p} + O(n^{-\delta_1} + n^{-\delta_2} + s_n^{-p}) + \\ &+ O(n^{-1/2} h_n^{-p/2} \log^{3/2} s_n) + O(n^{\alpha-1/2} h_n^{p/2} \log s_n).\end{aligned} \tag{33}$$

Как показано при доказательстве теоремы 3, можно положить

$$0 < \delta_1 < \frac{\tilde{\delta}}{2} - \frac{1}{4q}, \quad \delta_2 = 2q\left(\frac{\tilde{\delta}}{2} - \frac{1}{4q} - \delta_1\right).$$

Выравнивание влечет

$$\delta_1 = \delta_2 = \left(\frac{\tilde{\delta}}{2} - \frac{1}{4q}\right)\left(1 - \frac{1}{1+2q}\right). \tag{34}$$

Выберем $\alpha = {}^1/_2 + \varepsilon$, тогда при $s_n^p = n^{1/2+\varepsilon}$ последние два слагаемых в (33) есть $O(n^{-1/4+\varepsilon/2} \log^{3/2} n)$, а условия (5) выполнены с $\delta_0 = (1-3\varepsilon)/2$. Если $\tilde{\delta}$ близко к δ_0, q велико, то, согласно (34), δ_1 и δ_2 близки к ${}^1/_4$ и доказательство теоремы следует из (33).

В случае $p = 1$ и $K(x) = \chi[-{}^1/_2, {}^1/_2]$ результат теоремы 3 усиливает оценку Н. В. Смирнова (29). В самом деле, минимизируя по s_n погрешности в (29), получим $O(n^{-1/6} \ln^{1/2} n)$ при $s_n = n^{1/6}$. Наш выбор $s_n = n^{1/2+\delta(\varepsilon)}$ даже не допустим, ввиду излишне жесткого условия (С).

В заключение отметим, что для случая $p = 1$ оценку остаточного члена в теореме 3 можно несколько улучшить, если использовать иной подход, основанный на оценке скорости сходимости в строгом принципе инвариантности [10]. Имеет место следующее утверждение, приводимое здесь без доказательства.

Теорема 5. *Пусть* $p = 1$, $\sigma^2 = \int K^2(x)\,dx$ *и* $h_n s_n < C$. *Тогда при выборе* $h_n = n^{-1/2}$

$$\mathbf{P}\left(\max_{1 \leqslant i \leqslant s_n} \sqrt{nh_n}\left|\frac{f_n(x_i) - \mathbf{M}f_n(x_i)}{\sigma\sqrt{f(x_i)}}\right| < l_n + \frac{y}{l_n}\right) =$$
$$= \left(2\Phi\left(l_n + \frac{y}{l_n}\right) - 1\right)^{s_n} + O(n^{-1/4} \log^{3/2} n).$$

Автор благодарен Н. Н. Ченцову и Д. М. Чибисову за полезные обсуждения.

Поступила в редакцию
19.10.76

ЛИТЕРАТУРА

[1] **V. A. Statulevičius**, On large deviations, Z. Wahrscheinlichkeitstheorie verw. Geb., 6, 2 (1966), 133—144.

[2] **Н. Г. де Брейн**, Асимптотические методы в анализе, М., ИЛ, 1961.

[3] **M. Rosenblatt**, A quadratic measure of deviation of two-dimensional density estimates and a test of independence, Ann. Statist., 3, 1 (1975), 1—14.

[4] **Э. Копсон**, Асимптотические разложения, М., изд-во «Мир», 1966.

[5] **Н. В. Смирнов**, О приближении плотностей распределения случайных величин, Ученые записки МГПИ им. В. П. Потемкина, XVI, 3 (1951), 69—96 (**Н. В. Смирнов**, Теория вероятностей и математическая статистика, Избранные труды, М., изд-во «Наука», 1970, 205—223).

[6] **В. П. Леонов, А. Н. Ширяев**, К технике вычисления семиинвариантов, Теория вероят. и ее примен., IV, 3 (1959), 342—355.

[7] **Н. В. Смирнов**, О построении доверительной области для плотности распределения случайной величины, ДАН СССР, 74, 2 (1950), 189—191.

[8] **Г. Крамер**, Математические методы статистики, М., изд-во «Мир», 1975.

[9] **А. А. Боровков**, Теория вероятностей, М., изд-во «Наука», 1976.

[10] **J. Komlós, P. Major, G. Tusnády**, An approximation of partial sums of independent RV's and the sample DF.I, Z. Wahrscheinlichkeitstheorie verw. Geb., 32, 2 (1975), 111—131.

COMPLETE ASYMPTOTIC EXPANSIONS FOR THE MAXIMUM DEVIATION OF THE EMPIRICAL DENSITY FUNCTION

V. D. KONAKOV (MOSCOW)

(*Summary*)

Complete asymptotic expansion (19) is obtained for the properly normalized maximum deviation of the empirical density function. As a consequence, the exact rate of convergence to the limiting distribution is found. Some approximations, more useful in practice, and their rates of convergence are also given. In the special histogram case the results obtained are refinements and generalizations of some results of N. V. Smirnov ([5], [7]).