

**АСИМПТОТИЧЕСКИЙ АНАЛИЗ БЛИЗОСТИ
ПРОЕКЦИОННЫХ ОЦЕНОК ПЛОТНОСТИ В МЕТРИКЕ
ХЕЛЛИНГЕРА**

А. Л. Абдувахидов,
Андижанский государственный университет имени З.М.Бабура,
г.Андижан, Узбекистан.

***Аннотация.** В данной работе получено асимптотическое представление для среднего значения расстояния Хеллингера между проекционной оценкой и неизвестной плотностью распределения. Рассматривается задача восстановления вероятностной плотности на основе выборочных данных, где используется метод проекционной оценки как один из эффективных инструментов статистического анализа. В работе детально исследуются условия сходимости и асимптотические свойства полученной оценки, что позволяет дать количественную характеристику качества аппроксимации неизвестной плотности. Особое внимание уделяется поведению среднего расстояния Хеллингера при увеличении объёма выборки, что имеет важное значение для прикладных задач статистики и математической теории вероятностей. Полученные результаты могут быть использованы для разработки более точных методов статистического моделирования и анализа данных, а также для оптимизации вычислительных алгоритмов, основанных на непараметрических подходах.*

***Ключевые слова:** Оценки на основе ортогональных рядов, Расстояние Хеллингера, Асимптотический анализ, Оценка плотности, Непараметрическая статистика, Теория вероятностей.*

A. L. Abduvohidov,
Andijan state University,
170100 Andijan, Uzbekistan.

**ASYMPTOTIC ANALYSIS OF THE CLOSENESS OF PROJECTION
DENSITY ESTIMATORS IN THE HELLINGER METRIC**

Abstract. *In this work, an asymptotic representation is obtained for the expected value of the Hellinger distance between the projection estimator and the unknown probability density function. The problem of density recovery based on sample data is considered, where the projection estimator is employed as one of the effective tools of statistical analysis. The study investigates in detail the conditions of convergence and the asymptotic properties of the obtained estimator, which allows for a quantitative characterization of the approximation quality of the unknown density. Special attention is given to the behavior of the expected Hellinger distance as the sample size increases, which is of great importance for applied problems in statistics and probability theory. The results obtained can be applied to the development of more accurate methods of statistical modeling and data analysis, as well as to the optimization of computational algorithms based on nonparametric approaches.*

Keywords: *Orthogonal series estimators, Hellinger distance, Asymptotic analysis, Density estimation, Nonparametric statistics, Probability theory.*

A. L. Abduvaxidov

Andijon davlat universiteti,
Andijon shahri, O'zbekiston.

PROEKSION ZICHLIK BAHOLASHLARNING YAQINLIGINI XELLINGER METRIKASIDA ASIMPTOTIK TAHLIL QILISH

Annotatsiya. *Ushbu ishda proeksion baholash bilan noma'lum taqsimot zichligi orasidagi Xellinger masofasining o'rtacha qiymati uchun asimptotik ifoda keltirilgan. Tadqiqotda tanlama ma'lumotlarga asoslanib ehtimollik zichligini tiklash masalasi ko'rib chiqiladi va bunda statistik tahlilning samarali vositalaridan biri sifatida proeksion baholash usuli qo'llaniladi. Ishda olingan baholashning yaqinlashuv shartlari va asimptotik xossalari batafsil tahlil qilinib, noma'lum zichlikni approksimatsiya qilish sifatiga miqdoriy tavsif beriladi. Xususan, tanlanma hajmi ortishi bilan Xellinger masofasining o'rtacha qiymatidagi o'zgarishlarga alohida e'tibor qaratilgan bo'lib, bu holat statistika va ehtimollar nazariyasi bo'yicha amaliy masalalar uchun muhim ahamiyat kasb*

etadi. Olingan natijalar statistik modellash va ma'lumotlarni tahlil qilishning yanada aniq usullarini ishlab chiqishda, shuningdek, noparametrik yondashuvlarga asoslangan hisoblash algoritmlarini optimallashtirishda qo'llanishi mumkin.

Kalit so'zlar: Ortogonal qatorlarga asoslangan baholashlar, Xellingер masofasi, Asimptotik tahlil, Zichlikni baholash, Noparametrik statistika, Ehtimollar nazariyasi.

1. Введение

Расстояние Хеллингера между распределениями играет существенную роль в различных задачах математической статистики (см.[2]). В теории статистического оценивания качество оценки плотности распределения можно определить по ее отклонению от оцениваемой плотности по величине расстояния Хеллингера. В данной работе изучена задача сходимости проекционной оценки плотности распределения к истинной плотности в смысле расстояния Хеллингера.

Пусть $Z_1, Z_2, \dots, Z_n$ - статистическая выборка, состоящая из независимых и одинаково распределенных (н.о.р.) случайных величин (с.в.) с общей плотностью распределения $p(x), x \in R$. Предположим, что $p(x) \in L_p[a, b]$ и функции $\{\varphi_j(x), j=0, 1, \dots\}$ образуют полную ортонормированную систему в пространстве $L_p[a, b]$, где $-\infty \leq a < b \leq \infty$. В качестве оценки $p(x)$ используем проекционные оценки Н.Н.Ченцова [3,4]:

$$p_n(x) = \frac{1}{n} \sum_{i=1}^n K_n(x, Z_i), \quad (1)$$

где $K_n(x, y) = \sum_{j=0}^q \varphi_j(x) \varphi_j(y)$ и при $n \rightarrow \infty, q = q(n) \rightarrow \infty$ так, что $q = o(n)$.

Различные статистические свойства оценки (1) установлены в работах [3-7].

В настоящей работе исследуем сходимость величины $E\chi_n$ к нулю при $n \rightarrow \infty$,

где χ_n - расстояние Хеллингера между p_n и p :

$$\chi_n = \int_{N_p} (\sqrt{p_n(x)} - \sqrt{p(x)})^2 dx. \quad (2)$$

Здесь $N_p = \{x : p(x) > 0\}$ -носитель плотности $p(x)$ считается компактным множеством.

2. Основной результат

Рассмотрим следующие условия:

(У1) Для любого $\varepsilon > 0$ и $x \in N_p$ при $n \rightarrow \infty$:

$$\sigma_n^{-2}(x) \int_{|x-y| \geq \varepsilon} K_n^2(x, y) dy = o(1),$$

где
$$\sigma_n^2(x) = \int_{N_p} K_n^2(x, y) dy;$$

(У2) Существует абсолютная постоянная C такая, что $\sup_{j \geq 1} \sup_{a \leq x \leq b} |\varphi_j(x)| \leq C;$

(У3) Плотность $p(x)$ имеет производные $p^{(r)}(x), r \geq 1$ и при $j = 0, 1, 2, \dots$

существует число C_r такое, что

$$\left| \int_{N_p} \varphi_j(x) p(x) dx \right| \leq C_r \cdot j^{-\alpha(r)},$$

где $\alpha(r) > 0$ -некоторое число;

(У4) Для любого $x \in N_p$ где $n \rightarrow \infty$

$$\frac{\sigma_n^2(x)}{n} = o(1), \frac{\beta_n^2}{\sigma_n^2(x)n} = o(1), \quad (3)$$

где
$$\beta_n = \sup_{x \in N_p} \sigma_n^2(x).$$

Сформулируем основное утверждение.

Теорема. Пусть имеют место условия (У1)-(У4) и существует целое

число $m > 1$ такое, что при $n \rightarrow \infty, \left(\frac{q}{n}\right)^{m-1} \cdot \sqrt{q} = o(1).$ Тогда при $n \rightarrow \infty$

$$E\chi_n = \frac{1}{4n} \int_{N_p} \sigma_n^2(x) dx + O(q^{-2(\alpha(r)-1)}) + O\left(\frac{q}{n}\right). \quad (4)$$

Следствие. Если в условиях теоремы при $i, j = 0, 1, 2, \dots$

$$\int_{N_p} \varphi_i(x) \varphi_j(x) dx = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

то согласно (4)

$$E\chi_n = \frac{q}{4n} + O\left(q^{-2(\alpha(r)-1)}\right) + O\left(\frac{q}{n}\right).$$

Если же здесь выберем $q = q(n) = C^0 \cdot n^{1/(2\alpha(r)-1)}$, тогда в (4) будем иметь оценку

$$E\chi_n = C^0 n^{-\frac{2(\alpha(r)-1)}{2\alpha(r)-1}},$$

что подтверждает состоятельность проекционной оценки плотности порядка $n^{\frac{2(\alpha(r)-1)}{2\alpha(r)-1}}$. Для доказательства теоремы нам понадобятся несколько лемм.

Лемма 2.1. При условиях (У2), (У3) и $n \rightarrow \infty$

$$Ep_n(x) = p(x) + O\left(q^{-(\alpha(r)-1)}\right). \quad (5)$$

Доказательство леммы 2.1 Пусть

$$a_n(x) = \sum_{j=1+q}^{\infty} a_j \varphi_j(x),$$

где $a_j = \int \varphi_j(x) p(x) dx$ - коэффициенты Фурье в формальном разложении

$$p(x) = \sum_{j=0}^{\infty} a_j \varphi_j(x).$$

Поскольку

$$p_n(x) = \sum_{j=0}^q a_{jn} \varphi_j(x), \quad a_{jn} = \frac{1}{n} \sum_{k=1}^n \varphi_j(Z_k),$$

тогда ввиду несмещенности оценки a_{jn} для a_j , имеем

$$Ep_n(x) = p(x) - a_n(x),$$

где при условиях леммы

$$|a_n(x)| = \left| \sum_{j=1+q}^{\infty} a_j \varphi_j(x) \right| \leq C \sum_{j=1+q}^{\infty} |a_j| \leq C_1 q^{1-\alpha(r)},$$

что доказывает (5).

Лемма 2.2. Пусть выполнены условия (У1)-(У3) и первое из условий (3).

Тогда при $n \rightarrow \infty$

$$Dp_n(x) = \frac{p(x) \sigma_n^2(x)}{n} + O\left(\frac{\sigma_n^2(x)}{n}\right). \quad (6)$$

Доказательство леммы 2.2 Непосредственным вычислением имеем

$$\begin{aligned} d_n^2(x) &= Dp_n(x) = \frac{1}{n} DK_n(x, Z_1) = \frac{1}{n} \int K_n^2(x, y) p(y) dy - \frac{1}{n} \left(\int K_n(x, y) p(y) dy \right)^2 = \\ &= \frac{1}{n} \int K_n^2(x, y) p(y) dy - \frac{1}{n} (p(x) - a_n(x))^2. \end{aligned}$$

Отсюда с учетом непрерывности $p(x)$ имеем

$$\begin{aligned}
nd_n^2(x) &= p(x)\sigma_n^2(x) + \int_{|x-y|\geq \varepsilon} K_n^2(x,y)(p(y)-p(x))dy + \int_{|x-y|<\varepsilon} K_n^2(x,y)(p(y)-p(x))dy - \\
&\quad - (p(x)-a_n(x))^2 = + \int_{|x-y|<\varepsilon} K_n^2(x,y)(p(y)-p(x))dy - (p(x)-a_n(x))^2 = \\
&= p(x)\sigma_n^2(x) + O\left(\int_{|x-y|\geq \varepsilon} K_n^2(x,y)dy\right) + O(\varepsilon\sigma_n^2(x)) - (p(x)-a_n(x))^2.
\end{aligned}$$

Отсюда следует (6). Лемма 2.2 доказана.

Лемма 2.3. Пусть $\xi_1, \xi_2, \dots$ - последовательность н.о.р. с.в. с $E\xi_1 = 0$,

$D\xi_1 = 1$ и существует число λ такое, что

$$P(|\xi_k| \leq \lambda) = 1, k = 1, 2, \dots$$

Тогда существуют абсолютные постоянные K_m такие, что для любого целого $m \geq 1$

$$\mu_m = E \left| n^{-1/2} \sum_{k=1}^n \xi_k \right|^{2m} \leq \begin{cases} K_m \left(\frac{\lambda^2}{n} + \left(\frac{\lambda^2}{n} \right)^{m-1} \right), m > 1, \\ K_1 \frac{\lambda^2}{n}, m = 1. \end{cases} \quad (7)$$

Доказательство леммы 2.3 Нетрудно убедиться в том, что ввиду независимости с.в.

$$\mu_m = \frac{1}{n^m} \sum_{(r_1, \dots, r_n)} E \xi_1^{r_1} \dots E \xi_n^{r_n},$$

где суммирование ведется по всем целым $r_k, k = \overline{1, n}$ таким, что $r_1 + \dots + r_n = 2m$ ($r_k \neq 1, k = \overline{1, n}$,

т.к. при $r_k = 1, E \xi_k = 0$). Так как $E \xi_1^2 = 1$ и с вероятностью 1, $|\xi_k| \leq \lambda, k = \overline{1, n}$, то для $r_k = 2, 3, \dots$,

$$E |\xi_1|^{r_k} = E |\xi_1|^2 \cdot |\xi_1|^{r_k-2} \leq \lambda^{r_k-2}. \text{ Тогда}$$

$$\mu_m \leq \frac{1}{n^m} \sum_{(r_1, \dots, r_p)} \lambda^{r_1 + \dots + r_p - 2p},$$

где $(r_1, \dots, r_p)$ означает набор элементов совокупности $(r_1, \dots, r_n)$, отличных от 1. Число расположений элементов $r_1, \dots, r_p$ по n местам не превосходит $n^p, p \leq m-1$.

Тогда,

$$\mu_m \leq K_m \sum_{p=1}^{m-1} \left(\frac{\lambda^2}{n} \right)^{m-p} \leq \begin{cases} K_m \left(\frac{\lambda^2}{n} + \left(\frac{\lambda^2}{n} \right)^{m-1} \right), m > 1, \\ K_1 \frac{\lambda^2}{n}, m = 1. \end{cases}$$

Лемма 2.3 доказана.

$$E |p_n(x) - Ep_n(x)|^{2m} = O \left(\left(\frac{\sigma_n^2(x)}{n} \right)^m \right), m = 1, 2, 3, \dots$$

Лемма 2.3 доказана.

Лемма 2.4. Пусть выполнены условия леммы 2.2 и второе из условий (3).

Тогда для целого $m \geq 1$ при $n \rightarrow \infty$

$$E |p_n(x) - Ep_n(x)|^{2m} = O \left(\left(\frac{\sigma_n^2(x)}{n} \right)^m \right), m = 1, 2, 3, \dots$$

Доказательство леммы 2.4 По определению оценки (1)

$$p_n(x) - Ep_n(x) = \frac{1}{n} \sum_{k=1}^n \eta_k,$$

где $\eta_k = K_n(x, Z_k) - EK_n(x, Z_k)$. Обозначим $\bar{\sigma}_n^2(x) = D\eta_1$. Согласно неравенству Коши-Буняковского и условию (У2) для всех $(x, y) \in [a, b]^2$:

$$|K_n(x, y)| = \left| \sum_{j=0}^q \varphi_j(x) \varphi_j(y) \right| \leq \left(\sum_{j=0}^q \varphi_j^2(x) \sum_{j=0}^q \varphi_j^2(y) \right)^{1/2} \leq \gamma_q.$$

Заметим, что с.в $\xi_k = \bar{\sigma}_n^{-1} \eta_k, k = \overline{1, n}$ удовлетворяют условиям леммы 2.3 с $\lambda = 2\bar{\sigma}_n^{-1} \gamma_q$.

Тогда согласно (7) при $m = 1, 2, 3, \dots$

$$\begin{aligned} E |p_n(x) - Ep_n(x)|^{2m} &= \left(\frac{\bar{\sigma}_n}{\sqrt{n}} \right)^{2m} E \left| \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{\eta_k}{\bar{\sigma}_n} \right|^{2m} \leq \\ &\leq \begin{cases} K_m \cdot \left(\frac{\bar{\sigma}_n(x)}{\sqrt{n}} \right) \cdot \left(\frac{\lambda^2}{n} + \left(\frac{\lambda^2}{n} \right)^{m-1} \right), m > 1, \\ K_1 \cdot \frac{\bar{\sigma}_n^2(x)}{n} \cdot \frac{\lambda^2}{n}, m = 1, \end{cases} = O \left(\left(\frac{\sigma_n^2(x)}{n} \right)^m \right), \end{aligned}$$

так как $\bar{\sigma}_n^{-1}(x) = O(\sigma_n^{-1}(x))$ и при $n \rightarrow \infty$

$$\frac{\lambda^2}{n} = O \left(\frac{\beta_n^2}{\sigma_n^2(x)n} \right).$$

Этим доказательство леммы 2.4 завершено.

Лемма 2.5. В условиях леммы 2.4 пусть существует целое число $m > 1$

такое, что $\left(\frac{q}{n}\right)^{m-1} \cdot \sqrt{q} = o(1)$ при $n \rightarrow \infty$. Тогда для каждого $x \in N_p$

$$E\sqrt{p_n(x)} = E\sqrt{p_n(x)} - \frac{Dp_n(x)}{8(Ep_n(x))^{3/2}} + O\left(\frac{q}{n}\right).$$

Доказательство леммы 2.5 Пусть X выборочное пространство с заданным на нем распределением Q выборки. Это пространство разобьем на две части $X = X_1 \cup X_2$, где для достаточно малого $\varepsilon > 0$

$$X_1 = \left\{ |p_n(x) - Ep_n(x)| < \varepsilon \right\},$$

$$X_2 = \left\{ |p_n(x) - Ep_n(x)| \geq \varepsilon \right\}.$$

Тогда

$$E\sqrt{p_n(x)} = \int_{X_1} \sqrt{p_n(x)} dQ + \int_{X_2} \sqrt{p_n(x)} dQ. \quad (8)$$

Далее используя ход доказательства теоремы в [1], оценим каждое слагаемое в (8) в отдельности. Согласно неравенству Маркова и леммы 2.4

$$\int_{X_2} \sqrt{p_n(x)} dQ \leq C^* \sqrt{q} Q(X_2) \leq C_1^* \sqrt{q} \varepsilon^{-2m} \left(\frac{\sigma_n^2(x)}{n} \right)^m. \quad (9)$$

Из (9) при условии $\left(\frac{q}{n}\right)^{m-1} \cdot \sqrt{q} = 0$ (1) при $n \rightarrow \infty$ имеем

$$\int_{X_2} \sqrt{p_n(x)} dQ = O\left(\frac{q}{n}\right). \quad (10)$$

Согласно разложению Тейлора с остаточным членом в форме Лагранжа на множестве X_1 имеем

$$\begin{aligned} \sqrt{p_n(x)} &= \sqrt{Ep_n(x)} \cdot \sqrt{1 + \left[(p_n(x) - Ep_n(x)) / Ep_n(x) \right]} = \sqrt{Ep_n(x)} + \frac{p_n(x) - Ep_n(x)}{2\sqrt{Ep_n(x)}} - \\ &- \frac{[p_n(x) - Ep_n(x)]^2}{8(Ep_n(x))^{3/2}} + \frac{[p_n(x) - Ep_n(x)]^3}{16(Ep_n(x))^{5/2} \left\{ 1 + \theta \cdot \frac{[p_n(x) - Ep_n(x)]}{(Ep_n(x))^{5/2}} \right\}}, \end{aligned} \quad (11)$$

где $0 < \theta < 1$. Подставим (11) в (8) и тогда имеем

$$\int_{X_1} \sqrt{p_n(x)} dQ = \sum_{k=1}^4 R_k, \quad (12)$$

где

$$R_1 = \sqrt{Ep_n(x)}Q(X_1), \quad R_2 = \left[2\sqrt{Ep_n(x)}\right]^{-1} \int_{X_1} [p_n(x) - Ep_n(x)] dQ,$$

$$R_3 = - \left[8(Ep_n(x))^{3/2}\right]^{-1} \int_{X_1} [p_n(x) - Ep_n(x)]^2 dQ,$$

$$R_4 = \left[16(Ep_n(x))^{5/2}\right]^{-1} \int_{X_1} \frac{[p_n(x) - Ep_n(x)]^3 dQ}{1 + \theta \cdot \frac{[p_n(x) - Ep_n(x)]}{(Ep_n(x))^{5/2}}}.$$

Исследуем каждое слагаемое в (12). Согласно неравенству Маркова и леммы 2.4 при $n \rightarrow \infty$ имеем

$$R_1 = \sqrt{Ep_n(x)}(1 - Q(X_2)) = \sqrt{Ep_n(x)} - \sqrt{Ep_n(x)}Q(X_2) = \sqrt{Ep_n(x)} + O\left(\left(\frac{q}{n}\right)^m\right). \quad (13)$$

В силу лемм 2.2 и 2.4 и неравенства Коши-Буняковского для достаточно больших n имеем

$$|R_2| \leq \left[2\sqrt{Ep_n(x)}\right]^{-1} \cdot \sqrt{Q(X_1) \int_{X_1} (p_n(x) - Ep_n(x))^2 dQ} \leq C_2^* \cdot \left[\left(\frac{q}{n}\right)^{m+1}\right]^{-1/2}. \quad (14)$$

Нетрудно проверить, что

$$R_3 = \frac{Dp_n(x)}{8(Ep_n(x))^{3/2}} - \left(8(Ep_n(x))^{3/2}\right)^{-1} \cdot \int_{X_1} (p_n(x) - Ep_n(x))^2 dQ. \quad (15)$$

Используя лемму 2.4 и неравенство Коши-Буняковского для достаточно больших n получаем

$$\int_{X_1} (p_n(x) - Ep_n(x))^2 dQ \leq \sqrt{Q(X_1) \int_{X_1} (p_n(x) - Ep_n(x))^4 dQ} \leq C_3^* \left(\frac{q}{n}\right)^{-\frac{m+2}{2}}. \quad (16)$$

Используя (16) в (15), имеем

$$R_3 = \frac{Dp_n(x)}{8(Ep_n(x))^{3/2}} + O\left(\left(\frac{q}{n}\right)^{-\frac{m+2}{2}}\right). \quad (17)$$

При $\min_{x \in N_p} Ep_n(x) > \varepsilon$ для всех $x \in N_p$:

$$0 < 1 - \frac{\varepsilon}{Ep_n(x)} < 1 - \theta \cdot \frac{\varepsilon}{Ep_n(x)} < 1 + \theta \cdot \frac{[p_n(x) - Ep_n(x)]}{Ep_n(x)}.$$

Следовательно,

$$|R_4| \leq \frac{E|p_n(x) - Ep_n(x)|}{16(Ep_n(x) - \varepsilon)^{5/2}} \leq C_4^* (Ep_n(x) - \varepsilon)^{-5/2} \left(\frac{q}{n}\right)^{3/2}. \quad (18)$$

Теперь из соотношений (8), (10), (12)-(14), (17) и (18) получаем утверждение леммы 2.5. Лемма 2.5 доказана.

Доказательство теоремы. Легко убедиться в том, что

$$E\chi_n = \int_{N_p} E(\sqrt{p_n(x)} - E\sqrt{p_n(x)})^2 dx + \int_{N_p} (E\sqrt{p_n(x)} - \sqrt{p(x)})^2 dx = A_n + B_n. \quad (19)$$

Для A_n используя лемму 2.5, имеем

$$\begin{aligned} A_n &= \int_{N_p} E(\sqrt{p_n(x)} - E\sqrt{p_n(x)})^2 dx = \int_{N_p} Ep_n(x) dx - \int_{N_p} (E\sqrt{p_n(x)})^2 dx = \int_{N_p} Ep_n(x) dx - \\ &\quad - \int_{N_p} (E\sqrt{p_n(x)})^2 dx = \frac{1}{4} \int_{N_p} \frac{Dp_n(x)dx}{Ep_n(x)} + o\left(\frac{q}{n}\right). \end{aligned}$$

Отсюда согласно лемме 2.2

$$A_n = \frac{1}{4n} \int_{N_p} \frac{p(x)\sigma_n^2(x)}{Ep_n(x)} dx + o\left(\frac{q}{n}\right). \quad (20)$$

Используя лемму 2.1 для $x \in N_p$, имеем

$$(Ep_n(x))^{-1} = (p(x))^{-1} (1 + O(q^{-2(\alpha(r)-1)})). \quad (21)$$

Следовательно, из (20), (21) следует, что

$$A_n = \frac{1}{4n} \int_{N_p} \sigma_n^2(x) dx + o\left(\frac{q}{n}\right). \quad (22)$$

Теперь рассмотрим слагаемое B_n в (19). При $x \in N_p$ используем разложение Тейлора

$$\sqrt{Ep_n(x)} = p(x) \left\{ 1 + \frac{[Ep_n(x) - p(x)]}{p(x)} \right\}^{1/2} = p(x) + \frac{[Ep_n(x) - p(x)]}{2\sqrt{p(x)}} + O\left(\frac{[Ep_n(x) - p(x)]^2}{(p(x))^{3/2}}\right). \quad (23)$$

В (23) применим леммы 2.1, 2.2 и 2.5 и тогда будем иметь

$$\begin{aligned} B_n &= \int_{N_p} (E\sqrt{p_n(x)} - \sqrt{p(x)})^2 dx = \int_{N_p} \left\{ E\sqrt{p_n(x)} - \sqrt{p(x)} - \frac{Dp_n(x)}{8(Ep_n(x))^{3/2}} + o\left(\frac{q}{n}\right) \right\}^2 dx = \\ &= \int_{N_p} \left\{ \frac{[Ep_n(x) - p(x)]}{2\sqrt{p(x)}} + O(q^{-2(\alpha(r)-1)}) + o\left(\frac{q}{n}\right) \right\}^2 dx = O(q^{-2(\alpha(r)-1)}). \end{aligned} \quad (24)$$

Суммируя соотношения (19), (22) и (24), получаем утверждение теоремы. Теорема доказана полностью.

Заключение. В данной работе исследована задача восстановления плотности распределения с помощью проекционной оценки в терминах расстояния Хеллингера. Основным результатом является доказательство того, что при выполнении условий (У1–У4) проекционная оценка асимптотически сходится к истинной плотности. Сформулированная теорема

и соответствующие леммы показывают, что расстояние Хеллингера стремится к нулю, что подтверждает состоятельность и надежность метода. Кроме того, установлено, что с увеличением объема выборки качество оценки существенно улучшается. Полученные результаты имеют важное значение как для теоретических исследований в области теории вероятностей и непараметрической статистики, так и для прикладных задач моделирования и анализа данных.

Список литературы

1. Абдушукуров А.А., Сагидуллаев К.С. Асимптотические разложения для числовых характеристик оценки функции среднего остаточного времени безотказной работы. // Сб. Статистические методы оценивания и проверки гипотез. Пермь. Пермский гос.университет. 2007. вып.20. С.72-81.
2. Боровков А.А. Математическая статистика. М.: Наука. 1984. 472 С.
3. Ченцов Н.Н. Оценка неизвестной плотности распределения по наблюдениям. // ДАН СССР. 1962. т. 147. N.1. С. 45-48.
4. Ченцов Н.Н. Статистические решающие правила и оптимальные выводы. М.: Наука. 1972. 520 С.
5. Hall P. Comparison of two orthogonal series methods of estimating a density and its derivatives on an interval. // J. Multivar. Anal.1982. v.12. P.432-449.
6. Kronmal R., Tarter M. The estimation of probability densities and cumulatives by Fourier series methods. // J.A.S.A. 1990. V.63. P.925-952.
7. Viollaz A. Nonparametric estimation of a probability density functions based on or-thogonal expansions. // Revista. Math. 1989. V.2. P. 41-82.