УДК 519.21

О. Меляева

Предельное распределение интегральных функционалов от ядерных оценок плотности вероятности

Рассмотрим независимую случайную выборку $X_1, X_2, \ldots, X_n$ объема n из совокупности с неизвестной плотностью распределения $f(x)$.

В качестве оценки для $f(x)$ выберем ядерную оценку вида

$$f_n(x) = (na_n)^{-1} \sum_{i=1}^{n} W\left(\frac{x - X_i}{a_n}\right),$$

где положительные числа $a_n \to 0$ при $n \to \infty$, $W(x)$ — известная симметрическая относительно нуля плотность распределения на R.

Следуя [1, с. 34], при любом целом $p \geqslant 1$ определим величину

$$B_n^{2p} = \int (f_n(x) - Ef_n(x))^{2p} h(x)\, dx. \tag{1}$$

Здесь $h(x)$ — весовая функция, интегрирование ведется по всему множеству R, функции W, f и h предполагаются такими, что интеграл в (1) существует.

Исследованию распределения случайной величины B_n^{2p} при $p = 1$ посвящено много работ (см., например, [1, 2]). Цель данной работы состоит в том, чтобы изучить асимптотику распределения B_n^{2p} при любом фиксированном $p \geqslant 2$. При этом будем применять в основном метод работы [1].

Сформулируем основной результат.

Теорема. *Предположим, что функции f, W и h являются непрерывными и ограниченными на R, а последовательность $\{a_n\}$ такова, что при $n \to \infty$ $a_n \to 0$ и существует число $\alpha > 0$ такое, что $n^\alpha a_n \to \infty$.*

Тогда имеет место предельное соотношение

$$\sigma^{-1}(B_n^{2p})(B_n^{2p} - EB_n^{2p}) \xrightarrow{d} \tau, \tag{2}$$

где $\xrightarrow{d}$ обозначает сходимость по распределению, τ — стандартная нормальная случайная величина.

Доказательство. Сначала запишем B_n^{2p} в виде функционала Мизеса

$$B_n^{2p} = \sum_{i_1=1}^{n} \ldots \sum_{i_{2p}=1}^{n} \psi(X_{i_1}, \ldots, X_{i_{2p}}) \tag{3}$$

с ядром

$$\psi(x_1, \dots, x_{2p}) = \int \prod_{s=1}^{2p} (H_n(x_s, x) - EH_n(x, X_1))\, h(x)\, dx,$$

где $H_n(x, y) = (na_n)^{-1} W\left(\dfrac{x-y}{a_n}\right)$.

Из (3) [1, с. 35] имеем

$$B_n^{2p} = \sum_{i=1}^{2p} \binom{n}{m} \mathcal{U}_{nm}, \tag{4}$$

где $\mathcal{U}$-статистика $\mathcal{U}_{nm}$ равна

$$\mathcal{U}_{nm} = \binom{n}{m}^{-1} \sum_{1 \leqslant i_1 < \dots < i_m \leqslant n} \psi_m(X_{i_1}, \dots, X_{i_m})$$

с ядром

$$\psi_m(X_{i_1}, \dots, X_{i_m}) = \sum_{\substack{\gamma_1 + \dots + \gamma_m = 2p \\ \gamma_j \geqslant 1}} \frac{(2p)!}{\gamma_1! \dots \gamma_m!} \int \prod_{s=1}^{m} (H_n(x_s, x) - EH_n(x, X_1))^{\gamma_s} \times$$
$$\times h(x)(dx).$$

Из (4) имеем (см. формулу (5.9) [1, с. 36])

$$B_n^{2p} - EB_n^{2p} = \sum_{m=1}^{2p} \sum_{1 \leqslant i_1 < \dots < i_m \leqslant n} \Phi_m(X_{i_1}, \dots, X_{i_m}), \tag{5}$$

где $\Phi_m(x_1, \dots, x_m) = \psi_m(x_1, \dots, x_m) - E\psi_m(X_1, \dots, X_m)$.

Очевидно, что функция $\Phi_m(x_1, \dots x_m)$ в (5) является симметрической относительно любой перестановки аргументов $x_1, \dots, x_m$. Для применения теоремы 4.1 из [1, с. 33] необходимо по ядру $\Phi_m(x_1, \dots, x_m)$ вычислить функции $g_{mc}(x_1, \dots, x_c)$, $c = 1, \dots, m$, (см. формулы (2.2) в [1, с. 19]).

При замене подынтегральной переменной получаем

$$EH_n(x, X_1) = \frac{1}{n} \int W(y) f(x + a_n y)\, dy. \tag{6}$$

Для функции

$$\Phi_{mc}(x_1, \dots, x_c) = E(\Phi_m(X_1, \dots, X_m) \mid X_1 = x_1, \dots, X_c = x_c)$$

имеем явный вид

$$\Phi_{mc}(x_1, \dots, x_c) = \sum_{\gamma_1 + \gamma_2 + \dots + \gamma_m = 2p} \frac{(2p)!}{\gamma_1! \dots \gamma_m!} \left\{ \int \prod_{s=1}^{c} (H_n(x_s, x) - EH_n(x, X_1))^{\gamma_s} \times \right.$$
$$\times \prod_{s=c+1}^{m} E(H_n(X_s, x) - EH_n(x, X_1))^{\gamma_s} h(x)\, dx -$$
$$\left. - \int \prod_{s=1}^{m} E(H_n(X_s, x) - EH_n(x, X_1))^{\gamma_s} h(x)\, dx \right\}. \tag{7}$$

При выводе (7) учтено, что

$$E\psi_m(X_1, \dots, X_m) =$$

$$= \sum_{\substack{\gamma_1+\dots+\gamma_m=2p \\ \gamma_j \geqslant 1}} \frac{(2p)!}{\gamma_1! \dots \gamma_m!} \int \prod_{s=1}^{m} E\,(H_n(X_s, x) - EH_n(x, X_1))^{\gamma_s} h(x)\, dx. \tag{8}$$

В (8) с учетом (6) $E\,(H_n(X_s, x) - EH_n(x, X_1))^{\nu_s} = a_n \int \left(\frac{1}{na_n} W(z) - \right.$ $\left. - \frac{1}{n} \int W(y) f(x + a_n y)\, dy \right)^{\nu_s} f(x + a_n z)\, dz$. Отсюда при $n \to \infty$ получаем для $\nu_s \geqslant 2$

$$n^{\nu_s} a_n^{\nu_s - 1} E\,(H_n(X_s, x) - EH_n(x, X_1))^{\nu_s} \to f(x) \int W^{\nu_s}(z)\, dz.$$

В случае $m < 2p$ с учетом (8) при $n \to \infty$ имеем

$$n^{2m} a_n^{2m-1} E\Psi_m(X_1, \dots, X_m) \to m^{2p} \int h(x) f(x)\, dx \int W^{2p}(z)\, dz. \tag{9}$$

Пусть $m = 2p$. Тогда $\Psi_{2p}(x_1, \dots, x_{2p}) = (2p)! \int \prod_{s=1}^{2p} (H_n(x_s, x) -$

$$- EH_n(x, X_1))\, h(x)\, dx. \tag{10}$$

Формула (10) показывает, что $E\Psi_{2p}(X_1, \dots, X_{2p}) = 0$. Поэтому в (5) $\Phi_{2p}(x_1, \dots, x_{2p}) = \Psi_{2p}(x_1, \dots, x_{2p})$ и, кроме того, ядро $\Phi_{2p}(x_1, \dots, x_{2p})$ в силу (7) и (10) обладает свойством полной вырождаемости, т. е. для всех $c = 1, \dots, 2p - 1$ справедливо

$$\Phi_{2p,c}(x_1, \dots, x_c) \equiv 0. \tag{11}$$

Далее, представим (5) в виде

$$B_n^{2p} - EB_n^{2p} = \sum_{1 \leqslant i_1 < \dots < i_{2p} \leqslant n} \Phi_{2p}(X_{i_1}, \dots, X_{i_m}) + R_{np}, \tag{12}$$

где $R_{np} = \sum_{m=1}^{2p-1} \sum_{1 \leqslant i_1 < \dots < i_m \leqslant n} \Phi_m(X_{i_1}, \dots, X_{i_m})$. Обозначим для $m = 1, 2, \dots, 2p$ $S_{n,m} = \sum_{1 \leqslant i_1 < \dots < i_m \leqslant n} \Phi_m(X_{i_1}, \dots, X_{i_m})$. По формуле Гефдинга для дисперсии $\mathcal{U}$-статистики с учетом свойства (11) имеем

$$E\,(S_{n,2p})^2 = \binom{n}{2p}^{-1} E\Phi_{2p}^2(X_1, \dots, X_{2p}). \tag{13}$$

Оценим теперь математическое ожидание $E\Phi_{2p}^2(X_1, \dots, X_{2p})$, применяя (10). По определению имеем

$$E\Phi_{2p}^2(X_1, \dots, X_{2p}) = \int \dots \int \Phi_{2p}^2(x_1, \dots, x_{2p}) f(x_1) \dots f(x_{2p})\, dx_1 \dots dx_{2p} =$$

$$= ((2p)!)^2 \int \Big\{ \int \prod_{s=1}^{2p} (H_n(x_s, y_1) - EH_n(y_1, X_1))\, (H_n(x_s, y_2) - \tag{14}$$

$$- EH_n(y_2, X_1))\, f(x_1) \dots f(x_{2p})\, dx_1 \dots dx_{2p} \Big\}\, h(y_1)\, h(y_2)\, dy_1 dy_2.$$

Здесь справа сделаем замены подынтегральных переменных по формулам

$$x_s = a_n \cdot z_s + y_1, \quad s = 1, \dots, 2p,$$
$$y_2 = y_1 + a_n u \tag{15}$$

и затем положим $n \to \infty$. В результате некоторых несложных преобразований получим предельное соотношение

$$n^{4p} a_n^{2p-1} E\Phi_{2p}^2 \to (2p!)^2 C_p, \tag{16}$$

$$\text{где } C_p = \int h^2(x) f^{2p}(x)\,dx \int \left(\int W(z)\,W(z+u)\,dz\right)^{2p} du.$$

Далее оценим ER_{np}^2. По определению и неравенству $|x_1 + \ldots + x_m|^2 \leqslant \leqslant m^2 (x_1^2 + \ldots + x_m^2)$ имеем

$$ER_{np}^2 \leqslant (2p)^{2p} \sum_{m=1}^{2p-1} ES_{n,m}^2. \tag{17}$$

В (17) математическое ожидание $ES_{n,m}^2$ при $m = 1, \ldots, 2p-1$ можно вычислить с помощью формулы для дисперсии $\mathcal{U}$-статистики, где появляющиеся интегралы оцениваются так же, как (14), с применением замены переменных типа (15).

В результате аналогичных оценок находим, что при $n \to \infty$ $n^{4p} a_n^{2p-1} ES_{n,m}^2 \to 0$ для всех $m = 1, \ldots, 2p-1$. Отсюда и из (17) при $n \to \infty$ получаем

$$n^{4p} a_n^{2p-1} ER_{np}^2 \to 0. \tag{18}$$

Далее, случайную величину $\sigma^{-1}(B_n^{2p})(B_n^{2p} - EB_n^{2p})$ в соответствии с (12) представим как

$$\sigma^{-1}(B_n^{2p})(B_n^{2p} - EB_n^{2p}) = \sigma^{-1}(B_n^{2p})\,S_{n,2p} + Q_{n,p}, \tag{19}$$

где $Q_{n,p} = \sigma^{-1}(B_n^{2p})\,R_{np}$.

Из (18) и неравенства Чебышева при $n \to \infty$ вытекает соотношение

$$Q_{n,p} \xrightarrow{p} 0. \tag{20}$$

Поэтому осталось показать, что при $n \to \infty$

$$\sigma^{-1}(B_n^{2p})\,S_{n,2p} \xrightarrow{d} \tau. \tag{21}$$

Сначала заметим, что ввиду (13), (16) и (18) при $n \to \infty$

$$\sigma(S_{n,2p})/\sigma(B_n^{2p}) \to 1.$$

Следовательно, (21) эквивалентно соотношению

$$\sigma^{-1}(S_{n,2p})\,S_{n,2p} \xrightarrow{d} \tau. \tag{22}$$

Для доказательства (22) достаточно проверить условия (4.2) и (4.3) из [1, с. 25]. В соответствии с неравенством (4.3) из [1, с. 26] при $\delta = 2$, $m = r = 2p$, $g_{2p} = \Phi_{2p}$ имеем

$$l_1(\Phi) \leqslant n^{-1} (E\Phi_{2p}^2)^{-2} E|\Phi_{2p}|^4. \tag{23}$$

Оценка для $E\Phi_{2p}^2$ следует из (16). Поэтому достаточно рассмотреть $E|\Phi_{2p}|^4$. По определению

$$E|\Phi_{2p}|^4 = (2p!)^4 \int \ldots \int |\Phi_{2p}(x_1, \ldots, x_{2p})|^4 f(x_1) \ldots f(x_{2p})\,dx_1 \ldots dx_{2p} =$$

$$= (2p!)^4 \int \left| \int \prod_{s=1}^{2p} (H_n(x_s, x) - EH_n(x, X_1))\,h(x)\,dx \right|^4 \times$$

$$\times f(x_1) \ldots f(x_{2p})\,dx_1 \ldots dx_{2p} = (2p!)^4 \int \Big\{ \prod_{s=1}^{2p} (H_n(x_s, y_1) - EH_n(y_1, X_1)) \times$$

$$\times (H_n(X_s, y_2) - EH_n(y_2, X_1))(H_n(x_s, y_3) - EH_n(y_3, X_1)) \times$$

$$\times (H_n(X_s, y_4) - EH_n(y_4, X_1))\,f(x_1) \ldots f(x_{2p})\,dx_1 \ldots dx_{2p} \Big\} \prod_{s=1}^{4} h(y_s)\,dy_s. \tag{24}$$

В (24) сделаем замену подынтегральных переменных по формулам вида (15). В результате некоторых преобразований получим асимптотическое при $n \to \infty$ соотношение

$$n^{8p} a_n^{6p-3} E \mid \Phi_{2p} \mid^4 \to \gamma_p \tag{25}$$

с некоторой постоянной $\gamma_p > 0$. Сравнивая (16), (23) и (25), видим, что при $n \to \infty$

$$l_1 (\Phi) = O \left(\frac{1}{n a_n^{2p-1}} \right).$$

Так как по условию $n^{\frac{1}{2p-1}} a_n \to \infty$, то $l_1 (\Phi) \to 0$ при $n \to \infty$. Таким образом, выполняется условие (4.2) из [1, с. 26]. Аналогично проверяется справедливость условия (4.3) из [1, с. 26]. Поэтому имеет место (22), а вместе с ним и (2). Теорема доказана.

1. *Королюк В. С., Боровских Ю. В.* Мартингальная теория $\mathscr{U}$-статистик.— Киев, 1984.— 56 с.— (Препринт / АН УССР. Институт математики; 84.63).
2. *Королюк В. С., Боровских Ю. В.* Асимптотический анализ распределений статистик.— Киев : Наук. думка, 1984.— 304 с.

Туркм. ун-т, Ашхабад Получено 28.08.86