П. Б. Бобров, Е. И. Островский

ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ ПРИ АДАПТИВНОМ ОЦЕНИВАНИИ

1. Введение

В последнее время адаптивные оценки вызывают особый интерес. Основное преимущество таких оценок заключается в их независимости от априорной информации об оцениваемой функции на весьма широких классах функций. В данной статье рассматриваются три классические задачи непараметрической статистики.

1) Оценивание функции регрессии. Необходимо восстановить функцию $f(x)$ по наблюдениям: $y_i : y_1 = f(x_i) + \xi_i$ в точках $x_i = i/n$, $i = 1, 2, \ldots, n$. $\{\xi_i\}$ – независимые одинаково распределенные центрированные случайные ошибки.

2) Оценивание плотности. Мы будем оценивать плотность распределения $f(x)$ по выборке $\xi_1, \ldots, \xi_n$ независимых одинаково распределенных величин.

3) Оценивание спектральной плотности. Рассматривается гауссовская центрированная стационарная последовательность $\{\xi_i\}$ с неизвестной спектральной плотностью $f(x)$. Мы оцениваем $f(x)$ на основе наблюдений этой последовательности.

Первые адаптивные оценки для задачи регрессии были предложены в [6, 5] и др., а для оценивания плотности – в [8]. В этих статьях рассматривались оценки адаптивные на классах эллипсоидов:

$$\mathcal{F}_\alpha = \left\{ f : \sum_{j=N}^{\infty} j^{2\alpha}\, c_j^2(f) = Q_N < \infty \right\}.$$

Здесь $c_j(f)$ – коэффициенты Фурье функции f по тригонометрической системе. В [1] класс оцениваемых функций был расширен до $\mathcal{L}_2$ в задаче оценивания регрессии. Адаптивные на классе $\mathcal{F}_\alpha$ оценки спектральной плотности были предложены в [9].

В работах [3, 4] вводятся адаптивные оценки, построенные при помощи эмпирических *wavelet*-коэффициентов в задачах оценивания регрессии и плотности соответственно. В этих статьях были также изучены минимаксные риски в широкой шкале пространств Бесова $B_{s,p,q}$.

Однако вопрос о построении доверительных границ для оценивания функции исследован достаточно слабо. В [2] был построен доверительный интервал для функции регрессии:

$$\mathbf{P}\left((A_n^*)^{-1}\|f_{nN}(x)-f(x)\|_{\mathcal{L}_2}>u\right)\leqslant Cu^{-2/3}(1+\ln^2 u),$$

где A_n^* – величина, которая будет описана ниже. Здесь и далее C – некоторые константы, при этом необязательно одинаковые. Этот интервал был построен при условии $\mathbf{E}\,\xi_i^4=\mu_4<\infty$. В [7] показатель степени при u был улучшен до -1 для этого же условия и до $-k/2$ при условии $\mathbf{E}\,\xi_i^{2k}=\mu_{2k}<\infty$, $k>2$.

Отметим, что эти доверительные интервалы неадаптивны, так как A_n^* зависит от класса гладкости оцениваемой функции.

В данной статье предлагаются универсальные адаптивные оценки для всех трех выше перечисленных задач. Мы строим как неадаптивные, так и асимптотически адаптивные доверительные интервалы в интегральной матрике.

2. Основные результаты

Мы будем предполагать, что неизвестная функция $f(x)$ принадлежит пространству $\mathcal{L}_2[0,1]$. Следовательно, она разлагается в ряд Фурье

$$f(x)=\sum_{j=1}^{\infty}c_j\,\varphi_j\,(x),$$

где $\{\varphi_j(x)\}$ – ортонормированная тригонометрическая система на $[0,1]$:

$$\varphi_{2l}(x)=\sqrt{2}\,\cos 2\pi lx,\quad \varphi_{2l+1}(x)=\sqrt{2}\,\sin 2\pi lx,\quad \varphi_1(x)=1,\quad l\geqslant 1.$$

Условие $x\in[0,1]$ не ограничивает общности. Функция $f(x)$ может быть определена на любой ограниченной области $D\subset\mathbf{R}^1$. Обозначим

$$\rho(N)=\sum_{j=N}^{\infty}c_j^2,\quad A_{nN}=\rho(N)+\frac{\delta N}{2n},\quad A_n^*=\min_{1\leqslant N\leqslant n}A_{nN}.$$

Здесь $\delta=\sigma^2+\frac{1}{n}\sum\limits_{i=1}^{n}f^2(x_i)$, $\sigma^2=\mathbf{D}\,\xi_i$ в задаче регрессии, $\delta=1$ в задаче оценивания плотности и $\delta=\|f\|^2_{\mathcal{L}_2}$ в задаче оценивания спектра.

Оценка функции $f(x)$ строится по формуле

$$f_{nN}(x) = \sum_{j=1}^{N} \hat{c}_j \, \varphi_j(x), \tag{2.1}$$

где $\hat{c}_j$ – состоятельные оценки коэффициентов Фурье. Эти оценки имеют следующий вид:

$$\hat{c}_j = \frac{1}{n}\sum_{i=1}^{n} y_i \, \varphi_j(x_i), \quad \hat{c}_j = \frac{1}{n}\sum_{i=1}^{n} \varphi_j(\xi_i), \quad \hat{c}_j = \frac{1}{n}\sum_{i=1}^{n-j} \xi_i \, \xi_{i+j}$$

для задач регрессии, плотности и спектра соответственно. Оптимальный порядок N_n задается выражением:

$$N_n = \operatorname*{argmin}_{N \in [1,n]} A_{nN}$$

и при этом A_n^* – скорость сходимости оценки (2.1). Однако величина ρ_N неизвестна, и мы должны предполагать некоторый класс гладкости оцениваемой функции. В данной статье мы предлагаем ряд критериев для выбора порядка оценки (2.1), не требующих таких предположений:

$$N_n = \operatorname*{argmin}_{1 \leqslant N \leqslant n} \left(\frac{\delta_n \, N}{n} - \sum_{j=1}^{N} \hat{c}_j^2 \right) = \operatorname*{argmin}_{1 \leqslant N \leqslant n} \tau_N ,$$

где δ_n – состоятельная оценка параметра δ и

$$N_n = \operatorname*{argmin}_{1 \leqslant N \leqslant n} \sum_{j=N}^{2N} \hat{c}_j^2 = \operatorname*{argmin}_{1 \leqslant N \leqslant n} \tau_N .$$

Для второго критерия необходимо, чтобы $f(x)$ удовлетворяла дополнительному условию

$$\exists \gamma, \quad 0 < \gamma < 1; \quad \gamma \rho_N > \rho_{2N}.$$

Первый критерий не требует данного условия, но включает в себя оценку параметра δ. Далее мы будем рассматривать только второй критерий. Идея этих критериев состоит в следующем:

$$\frac{\tau_N}{\mathbf{E}\, \tau_N} \xrightarrow{p} 1, \quad \mathbf{E}\, \tau_N \sim \rho_N + \frac{\delta N}{2n}.$$

Обозначим

$$P_f(u) = \mathbf{P}\left((A_n^*)^{-1}\int\limits_0^1 \big(f_{nN_n}(x) - f(x)\big)^2\,dx > u\right)$$

и сформулируем несколько основных теорем.

Теорема 2.1. *Если* $\mathbf{E}\xi_i^{2k} = \mu_{2k} < \infty$, $k > 2$ *и функция регрессии* $f(x)$ *интегрируема по Риману, то*

$$P_f(u) \leqslant 2^{2k}\,k^k\,\mu_{2k}\,u^{-k/2}, \quad u > 0.$$

Теорема 2.2. *Если погрешности* ξ_i *удовлетворяют дополнительному условию:*

$$\exists\, q > 0, \quad Q > 0; \quad \mathbf{P}\,\big(\,|\,\xi_i\,| > t\big) \leqslant \exp\big[-Q\,t^q\big], \quad t > 0, \qquad (2.2)$$

то

$$P_f(u) \leqslant C\exp\big[-u^{q/2(q+2)}\big], \quad u > 0,$$

где

$$C = \exp\left[-\frac{2q}{q+2} + \exp[-1] + 1 - 2\ln 2 - 2/q\right].$$

Теорема 2.3. *Если в теореме* (2.1) $k = 2$, *то*

$$P_f(u) \leqslant \big(1 + \ln^2 u\big)\,u^{-1}, \quad u \geqslant 1.$$

Теорема 2.4. *Пусть плотность распределения* $f(x)$ *ограничена. Тогда справедливо неравенство:*

$$P_f(u) \leqslant \exp\big[1 - 2\ln 2\big]\exp\big[-u^{1/2}\big].$$

Теорема 2.5. *Пусть спектральная плотность* $f(x)$ *интегрируема по Риману и ограничена. Тогда имеет место оценка:*

$$P_f(u) \leqslant \exp\big[-C(m)\,u^{1/4}\big].$$

Здесь $C(m) = (4e)^{-1}(2\pi m)^{1/2m}$.

Для построения асимптотически адаптивных доверительных интервалов сформулируем следующую теорему.

Теорема 2.6. *При $n \to \infty$ имеет место сходимость по вероятности:*

$$\frac{\min\limits_{1 \leqslant N \leqslant n} \tau_N}{A_n^*} \xrightarrow{p} 1.$$

Эта теорема справедлива для всех рассматриваемых задач. С ее помощью можно получить асимптотически адаптивные доверительные интервалы. В выражении для доверительного интервала величина A_n^* может быть заменена на $\min\limits_{1 \leqslant N \leqslant n} \tau_N$. При этом

$$\|f_{nN}(x) - f(x)\|_{\mathcal{L}_2} < \frac{u}{\min\limits_{1 \leqslant N \leqslant n} \tau_n}$$

с вероятностью не меньшей чем $1 - C\exp[-\gamma]$, где

$$\gamma = \frac{q}{2(2+q)}, \quad \gamma = 1/2, \quad \gamma = 1/4$$

для задач оценивания регресси, плотности и спектра соответственно.

3. Вспомогательные результаты

Перед тем как перейти к доказательствам, мы приведем здесь основную информацию о пространствах $\mathbf{G}_p$ и $\mathbf{B}_\varphi$. Теория этих пространств легла в основу применяемых нами методов. Для более детального ознакомления с этими пространствами мы рекомендуем читателям работы [10, 16] соответственно.

$\mathbf{G}_p$-пространства. Случайная величина η принадлежит пространству $\mathbf{G}_p$, если

$$\|\eta\|_p = \sup_{k \geqslant 2} k^{-k/q} \|\eta\|_{\mathcal{L}_k(\Omega)} < \infty,$$

$$\text{где} \quad p^{-1} + q^{-1} = 1 \quad \text{и} \quad \|\eta\|_{\mathcal{L}_k(\Omega)} = \big(\mathbf{E} \mid \eta \mid^k\big)^{1/k}.$$

Пусть $\eta(t)$ – случайное поле, где $t \in T$ и $\sup\limits_{t \in T} \|\eta(t)\|_p < \infty$. Положим $d_p(t,s) = \|\eta(t) - \eta(s)\|_p$. Пусть $\mathcal{N}(d_p, \varepsilon)$ – наименьшее число шаров d_p-радиуса ε, покрывающих множество T. Если сходится энтропийный интеграл

$$\int\limits_0^1 \big(\ln \mathcal{N}(d_p, \varepsilon)\big)^{1/q}\, d\varepsilon,$$

то

$$\mathbf{P}\left(\sup_{t\in T} \mid \eta(t) \mid > u\right) \leqslant C\exp[-C\,u^q].$$

Для пространств $\mathcal{L}_k(\Omega)$ аналогичный результат был получен в [17]. Если $r_k(t,s) = \|\eta(t)-\eta(s)\|_{\mathcal{L}_k(\Omega)}$ и

$$\int_0^1 \left(\mathcal{N}(r_k,\varepsilon)\right)^{1/k} d\varepsilon < \infty,$$

то

$$\mathbf{E}\sup_{t\in T} \mid \eta(t) \mid^k = C < \infty.$$

$\mathbf{B}_\varphi$-пространства. Случайная величина η принадлежит пространству $\mathbf{B}_\varphi$, если $\mathbf{E}\,\eta = 0$ и $\|\eta\|_\varphi < \infty$, где

$$\|\eta\|_\varphi = \sup_{\lambda\neq 0} \varphi^{-1}\big(\ln \mathbf{E}\,\exp[\,\pm\lambda\eta\,]\big) \mid \lambda \mid^{-1}.$$

Здесь φ – четная выпуклая функция, $\varphi(0) = 0$, $\varphi'(0) = 0$. Мы будем использовать результат, полученный в [16]:

$$\eta \in \mathbf{B}_\varphi \Longleftrightarrow \exists\, K > 0; \quad \mathbf{P}\,(\mid \eta \mid) > u) \leqslant 2\exp[-\varphi^*(Ku)],$$

где φ^* – преобразование Юнга–Фенхеля функции φ:

$$\varphi^*(x) = \sup_{y\geqslant 0}\big(xy - \varphi(y)\big).$$

4. Доказательства

Доказательства теорем 2.1–2.5 похожи. Мы будем отмечать некоторые встречающиеся различия. По аналогии с работами [1, 2] введем вероятности:

$$\pi_{1n} = \mathbf{P}\,(\exists\, N_n : N_n \leqslant N_{1n} - 1), \quad \pi_{2n} = \mathbf{P}\,(\exists\, N_n : N_n \leqslant N_{2n} + 1),$$

где

$$N_{1n} = \max(N : 1 \leqslant N < N_n^-,\ v\,A_n^* \leqslant A_{nN}) + 1,$$

$$N_{2n} = \min\,(N : N_n^+ < N \leqslant n,\ v\,A_n^* \leqslant A_{nN}) - 1$$

и

$$N_n^- = \min(N : A_{nN} = A_n^*), \quad N_n^+ = \max(N : A_{nN} = A_n^*).$$

$v \geqslant 3$ – некоторое фиксированное число. Обозначим:

$$P_0(u) = \mathbf{P}\big(\max_{N_{1n} \leqslant N \leqslant N_{2n}} (A_n^*)^{-1} \| f_{nN} - f \|^2_{\mathcal{L}_2[0,1]} > u\big).$$

Заметим, что

$$P_f(u) \leqslant \pi_{1n} + \pi_{2n} + P_0(u).$$

Вероятности π_{1n} и π_{2n} оцениваются следующим образом:

$$\pi_{1n} \leqslant \mathbf{P}\big(\min_{N_{1n} < N < N_{2n}} \tau_N > \min_{N \leqslant N_{1n}} \tau_N\big)$$

и

$$\pi_{2n} \leqslant \mathbf{P}\big(\min_{N_{1n} < N < N_{2n}} \tau_N > \min_{N \geqslant N_{2n}} \tau_N\big).$$

Функционал τ_N представляется в виде:

$$\tau_N = \rho(N) + \frac{\delta N}{2n} - \Psi_{nN} + o(1) \sim A_{nN} + L_{nN}.$$

Это разложение получается простым вычислением. В задаче регрессии функция Ψ_{nN} имеет вид: $\Psi_{nN} = 2\Psi_{1n}(1, N) + \Psi_{2n}(N)$, где так же, как и в [1, 2],

$$\Psi_{1n}(N', N'') = \sum_{j=N'}^{N''} f_{jn}\, q_{jn}, \qquad N'' \geqslant N',$$

$$\Psi_{2n}(N) = \sum_{j=1}^{N} (q_{jn}^2 - \mathbf{E}\, q_{jn}^2)$$

и

$$f_{jn} = \frac{1}{n} \sum_{i=1}^{n} f(x_i)\, \varphi_j(x_i), \qquad q_{jn} = \frac{1}{n} \sum_{i=1}^{n} \xi_i\, \varphi_j(x_i).$$

Функция Ψ_{nN} в задаче оценивания плотности представляется в виде:

$$\Psi_{nN} = 2\Psi_{1n}(1, N) + \Psi_{2n}(N), \quad \text{где} \quad \Psi_{1n} = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{N} c_j\, \varphi_j(\xi_i) - c_j^2$$

и Ψ_{2n} – вырожденная U-статистика с ядром:

$$\Phi(\xi, \eta) = \sum_{j=1}^{N} \big(\varphi_j(\xi) - c_j\big)\big(\varphi_j(\eta) - c_j\big).$$

Для задачи оценивания спектра функция Ψ_{nN} имеет другой вид. Здесь $\Psi_{nN} = J_4 + J_3 + J_2 + J_1$, где J_m – m-кратный стохастический интеграл Ито–Винера.

Для того, чтобы оценить вероятности π_{1n} и π_{2n} мы докажем следующие леммы.

Лемма 4.1. *Пусть выполнены условия теоремы* 2.1. *Тогда имеет место оценка:*

$$\mathbf{E} \sup_{[N',N'']} \mid \Psi_{1n}(M', M'') \mid^k \leqslant 9(2\pi)^k\, 2^{1/k}\, k^k\, \mu_k (N'')^{k/2-1} \left(\frac{\rho\,(N')}{n} \right)^{k/2}$$

и

$$\mathbf{E} \sup_{M<N} \mid \Psi_{2n}(M) \mid^k \leqslant 2^{2k}\, 9k^k\, \mu_{2k}\, N^{k/2}\, n^{-k}.$$

Лемма 4.2. *Пусть выполнены условия теоремы* 2.4. *Тогда*

$$\mathbf{E} \sup_{[N',N'']} \mid \Psi_{1n}(M', M'') \mid^k \leqslant 9(2\pi)^k\, 2^{1/k}\, k^k\, \mu_k (N'')^{k/2-1} \left(\frac{\rho\,(N')}{n} \right)^{k/2}$$

и

$$\mathbf{E} \sup_{M<N} \mid \Psi_{2n}(M) \mid^k \leqslant 9k^k\, N^{k/2}\, n^{-k}.$$

Доказательство леммы 4.1. Доказательство этой леммы аналогично доказательствам лемм 1 и 2, описанному в [7]. Здесь мы приведем это доказательство кратко. Из [1] следует, что $\mathbf{D}\, \Psi_{1n} \leqslant C\, \rho\,(N)/n$. Оценим $\mathbf{E} \mid \Psi_{1n} \mid^k$. Имеем

$$\Psi_{1n}(N', N'') \sim \sum_{l=N'}^{N''} c_l(f) \frac{1}{n} \sum_{i=1}^{n} \xi_i\, \varphi_l(x_i),$$

где $c_l(f) = \int f(x)\, \varphi_l(x)\, dx$.

Используя неравенство Дхармадхикари–Йогдео [11], получим

$$\mathbf{E} \mid \Psi_{1n} \mid^k \leqslant C(k) \mu_k\, n^{-k/2} \| \, \Phi_f(N'') - \Phi_f(N') \, \|^k_{\mathcal{L}_k[0,1]}.$$

Здесь

$$C(k) = \frac{1}{2} k(k-1) \max\left(1, 2^{k-3}\right) \left(1 + \frac{2}{k}\, K_{2l}^{(k-2)/2l} \right),$$

где l – целое число, которое удовлетворяет условию $2l \leqslant k < 2l+2$ и

$$K_{2l} = \sum_{r=1}^{l} \frac{r^{2l-1}}{(r-1)!}$$

и $\Phi_f(N)$ – частичная сумма Фурье функции $f(x)$:

$$\Phi_f(N) = \sum_{j=1}^{N} c_j\, \varphi_j(x).$$

Применяя неравенство Никольского [12], мы получим оценку:

$$\mathbf{E} \mid \Psi_{1n}(N', N'') \mid^k \leqslant 2\pi C(k)\mu_k(N'')^{k/2-1}\, n^{-k/2} \mid \rho(N'') - \rho(N') \mid^{k/2}.$$

Пусть $M \in [N', N'']$. Тогда

$$\| \Psi_{1n}(N', M) \|_{\mathcal{L}_k(\Omega)} \leqslant (2\pi C(k)\mu_k)^{1/k}(N'')^{1/2-1/k}\, n^{-1/2}\big(\rho(N')\big)^{1/2}.$$

Обозначая

$$\tau(M) = \frac{\Psi_{1n}(N', M)}{(2\pi C(k)\,\mu_k)^{1/k}(N'')^{1/2-1/k}\, n^{-1/2}\big(\rho(N')\big)^{1/2}},$$

мы имеем:

a) $\|\tau(M)\|_{\mathcal{L}_k(\Omega)} \leqslant 1$

b) для M', $M'' \in [N', N'']$ имеет место $d_k = \|\tau(M'') - \tau(M')\|_{\mathcal{L}_k(\Omega)} \leqslant C\sqrt{|t-s|}$, где $t = \rho(M'')/\rho(N')$, $s = \rho(M')/\rho(N')$ и $t,\ s \in [0,1]$.

Энтропия множества $[N/, N'']$ в d_k-метрике не превосходит энтропию отрезка $[0,1]$ в метрике $\sqrt{|t-s|}$, т.е. величину $2\varepsilon^{-2}$ при $\varepsilon \in [0,1]$. Энтропийный интеграл

$$\int_0^1 \big(\mathcal{N}(d_k, \varepsilon)\big)^{1/k}\, d\varepsilon$$

оценивается величиной $C\int_0^1 \varepsilon^{-2/k}\, d\varepsilon$, т.е. абсолютной константой при $k \geqslant 3$. Из результатов пункта 3 настоящей статьи следует

$$\mathbf{E} \sup_{M \in [N', N'']} \mid \tau(M) \mid^k \leqslant 9\frac{k+1}{k}\, 2^{1/k}.$$

Переходя к Ψ_{1n}, мы получаем первое утверждение леммы 4.1.

Функция Ψ_{2n} – квадратичная форма от случайных величин. Ее дисперсия равна $\mathbf{D}\,\Psi_{2n} \leqslant CN/n^2$. Для того, чтобы оценить k-ый момент этой квадратичной формы, мы оценим ее k-ый семиинвариант. В соответствии с [13, стр. 101]

$$\Gamma_k(\Psi_{2n}(N))/B_n^k \leqslant C^k\, k^k\, \mu_{2k}(W_n/B_n)^{k-2},$$

где $B_n^2 = \mathbf{D}\,\Psi_{2n}$ и

$$W_n = \max_i \sum_{j=1}^{n} |\, a_{ij}^{(n)} \,|, \quad a_{ij}^{(n)} = n^{-2} \sum_{l=1}^{N} \varphi_l(x_i)\, \varphi_l\,(x_j).$$

Величина W_n может быть оценена неравенством:

$$W_n \leqslant n^{-n} \max_y \sum_{j=1}^{n} \left| \sum_{l=1}^{N} \varphi_l(y)\, \varphi_l\,(x_j) \right|.$$

Величины вида:

$$L(N) = \max_y \int \left| \sum_{l=1}^{N} \varphi_l(y)\, \varphi_l\,(x) \right| dx$$

называются константами Лебега и $L(N) \leqslant C \ln(N+1)$. Мы можем по аналогии получить оценку $W_n \leqslant 4n^{-1} \ln\, N$ и, следовательно:

$$\Gamma_k\big(\Psi_{2n}(N)\big)/B_n^k \leqslant 2^{2k}\, k^k\, \mu_{2k} \frac{\ln N}{N} \quad \text{для} \quad k \geqslant 3.$$

Переходя от семиинвариантов к моментам по формулам из [18, с. 311], получаем оценку:

$$\|\Psi_{2n}(N)\|_{\mathcal{L}_k(\Omega)} \leqslant 4k\, \mu_{2k}^{1/k}\, N^{1/2}\, n^{-1}.$$

Пусть

$$\tau(M) = \frac{\Psi_{2n}(M)}{4k\, \mu_{2k}^{1/k}\, N^{1/2}\, n^{-1}}.$$

Тогда

a) $\|\tau(M)\|_{\mathcal{L}_k(\Omega)} \leqslant 1$ и

b) для M', $M'' \leqslant N$ мы имеем:

$$d_k = \|\tau(M'') - \tau(M')\|_{\mathcal{L}_k(\Omega)} \leqslant C \sqrt{\left| \frac{M'}{N} - \frac{M''}{N} \right|} = C\sqrt{|t-s|},$$

где $t = M'/N$, $s = M''/N$. Как и выше, энтропийный интеграл оценивается константой $C < \infty$. Следовательно,

$$\mathbf{E} \sup_M \mid \tau(M) \mid^k \leqslant 9 \frac{k+1}{k} 2^{1/k}.$$

Переходя к Ψ_{2n}, мы получаем второе утверждение леммы 4.1.

Доказательство леммы 4.2. Оценим k-тый момент функции Ψ_{1n}:

$$\mathbf{E} \left| n^{-1} \sum_{i=1}^{n} \sum_{j=N'}^{N''} c_j \, \varphi_j(\xi_i) - c_j^2 \right|^k \leqslant$$

$$\leqslant C(k) \, n^{-k/2} \int \left| \sum_{j=N'}^{N''} c_j \, \varphi_j(x) \right|^k f(x) \, dx \leqslant$$

$$\leqslant C(k) \, n^{-k/2} \left\| \sum_{j=N'}^{N''} c_j \, \varphi_j(x) \right\|^k_{\mathcal{L}_k[0,1]}.$$

Если мы применим неравенство Никольского, то получим следующую оценку:

$$\mathbf{E} \mid \Psi_{1n} \mid^k \leqslant 2\pi C(k) \, N^{k/2-1} \, \rho^{k/2}(N) \, n^{-k/2}.$$

k-тый момент Ψ_{1n} имеет такую же оценку как и в лемме 4.1. Далее первое утверждение леммы 4.2 доказывается аналогично лемме 4.1.

Из [13, стр. 108] следует, что дисперсия Ψ_{2n} имеет оценку:

$$\mathbf{D} \, \Psi_{2n} \leqslant \frac{4N}{n^2}.$$

Оценим теперь k-тый момент. Асимптотика таких моментов описана в [15]. Однако, нам требуется оценка сверху. Мы несколько модифицируем предложенный метод.

Имеет место следующее неравенство:

$$\mathbf{E} \mid \Phi(\xi, \eta) \mid^m \leqslant N^{m-1}, \quad m = 3, 4, \ldots, k.$$

Действительно,

$$\mathbf{E} \mid \Phi(\xi, \eta) \mid^m \sim \int \int \mid \mathcal{D}_N(x, y) \mid^m \, dx \, dy \leqslant N^{m-1}.$$

Здесь $\mathcal{D}_N(x,y)$ – ядро Дирихле:

$$\mathcal{D}_N(x,y) \sim \frac{\sin N(x-y)}{x-y}.$$

Так как U-статистика сингулярна, то мы имеем:

$$\mathbf{E} \mid \Psi_{2n} \mid^k \leqslant \frac{1}{n^{2k}} \sum_{i_1=1}^{n} \sum_{j_1=1}^{n} \ldots \sum_{i_k=1}^{n} \sum_{j_k=1}^{n} \mathbf{E}\, \Phi(\xi_{i_1}, \xi_{j_1}) \ldots \Phi(\xi_{i_k}, \xi_{j_k}) \leqslant$$

$$k^k \frac{1}{n^{2k}} \sum_{i_1=1}^{n} \sum_{j_1=1}^{n} \ldots \sum_{i_{k/2}=1}^{n} \sum_{j_{k/2}=1}^{n} \mathbf{E}\left(\Phi(\xi_{i_1}, \xi_{j_1})\right)^2 \ldots \mathbf{E}\left(\Phi(\xi_{i_{k/2}}, \xi_{j_{k/2}})\right)^2$$

$$= k^k \|\Psi_{2n}\|^k_{\mathcal{L}_2(\Omega)} \leqslant k^k \frac{N^{k/2}}{n^k}.$$

Используя путь, описанный в лемме 4.1, мы получаем второе утверждение леммы 4.2.

В теореме 2.5 мы имеем $\Psi_{nN} = J_4 + J_3 + J_2 + J_1$, где J_m – m-кратный стохастический интеграл Ито–Винера. Тогда

$$B_n^2 = \mathbf{D}\, \Psi_{nN} = \mathbf{D}\, J_4 + \mathbf{D}\, J_3 + \mathbf{D}\, J_2 + \mathbf{D}\, J_1.$$

Нетрудно показать, что $B_n \leqslant C(\rho(N)/n + N/n^2)$ и $\mathbf{D}\, J_m \leqslant B_n^2$. В соответствии с [14], мы получаем неравенство:

$$\mathbf{P}\left(\mid J_m/B_n \mid > u\right) \leqslant \exp\left[-C(m)\, u^{2/m}\right],$$

где $C(m) = (4e)^{-1}(2\pi m)^{1/2m}$. Для дисперсии справедливы оценки:

$$\mathbf{D}\, \Psi_{nN} \leqslant \begin{cases} 2\rho(N)/n & \text{при} \quad N \leqslant N_{1n}, \\ 2N/n^2 & \text{при} \quad N \geqslant N_{2n}. \end{cases}$$

Вернемся к оцениванию вероятностей π_{1n} и π_{2n}. Пусть τ' и τ'' – случайные величины такие, что

$$\sup \mid \Psi_{1n} \mid = \tau'\, K'(N_n^-)^{1/2-1/k} (\rho_N/n)^{1/2},$$

$$\sup \mid \Psi_{2n} \mid = \tau''\, K''(N_{n1})^{1/2}/n,$$

где K' и K'' – некоторые константы. Заметим, что

$$\max(\tau', \tau'') = \tau, \quad \|\tau\|_{\mathcal{L}_k(\Omega)} \leqslant 2,$$

$$\max(K', K'') \leqslant K\, \mu_{2k}^{1/k}\, k.$$

Используя неравенства (3.9), (3.13), (3.24) из [2]:

$$\begin{aligned} &2vn\,A_n^* > \delta\,N_{2n}, && \delta(N_{2n}+1) > 2(v-1)n\,A_n^*, \\ &\rho\,(N_{1n}-1) \geqslant (v-1)\,A_n^*, && \delta\,N_{1n} \leqslant n\,A_n^*, \end{aligned} \tag{4.1}$$

мы видим, что

$$\max\bigl((N_n^-)^{1/2-1/k}\bigl(\rho\,(N)/n\bigr)^{1/2},\ (N_{n1})^{1/2}/n\bigr)$$

$$= (N_n^-)^{1/2-1/k}\bigl(\rho\,(N)/n\bigr)^{1/2}$$

при $N \leqslant N_{1n}$ и

$$\max\bigl(N^{1/2-1/k}\bigl(\rho\,(N_{n2}+1)/n\bigr)^{1/2},\ N^{1/2}/n\bigr) = N^{1/2}/n$$

при $N \geqslant N_{2n}$.

Для того, чтобы оценить π_{1n} и π_{2n}, мы представим их в виде

$$\pi_{1n} \leqslant \mathbf{P}\,\bigl(K\tau/(A_n^*\,n)^{1/k}(v-1)^{1/2} > 1-1/v\bigr),$$

$$\pi_{2n} \leqslant \mathbf{P}\,\bigl(K\tau/(A_n^*\,n)^{1/2}(v-1)^{1/2} > 1-1/v\bigr).$$

Используя неравенство Чебышева, лемму 4.1 и (4.1), мы получим:

$$\pi_{1n} + \pi_{2n} \leqslant 9(2\pi)^k\,2^{1/k}\,2^{2k}\,k^k\,\mu_{2k}\,v^{-k/2}.$$

Для плотности и спектра:

$$\pi_{1n} + \pi_{2n} \leqslant 9(2\pi)^k\,2^{1/k}\,k^k\,v^{-k/2}$$

и

$$\pi_{1n} + \pi_{2n} \leqslant \exp\bigl[-C(m)\,v^{1/4}\bigr]$$

соответственно.

Доказательство теоремы 2.1. По аналогии с формулой (3.27) из [2]:

$$P_0(u) \leqslant \frac{1}{(A_n^*(u-v))^k}\mathbf{E}\sup_{[N_{1n},N_{2n}]} |\,\Psi_{2n}\,|^k$$

и используя второе неравенство леммы 4.1, мы получим:

$$P_0(u) \leqslant 9\cdot 2^{1/k}\,2^{2k}\,k^k\,\mu_{2k}\,v^{k/2}(u-v)^{-k}, \quad v < u.$$

Складывая оценки π_{1n}, π_{2n} и $P_0(u)$, мы выберем $v = u/2$ и получим:

$$P_f(u) \leqslant 2^{2k}\, k^k\, \mu_{2k}\, u^{-k/2}.$$

Это неравенство доказывает теорему 2.1.

Доказательство теоремы 2.2. Из условия (2.2) следует, что $\mu_k = C(q)\, k^{k/q}$. Мы имеем

$$P_f(u) \leqslant 2^{2k}\, k^{k+2k/q}\, u^{-k/2}.$$

Минимизируя по k выражение для $P_f(u)$, мы доказываем теорему 2.2.

Доказательство теоремы 2.3. Эта теорема доказывается по аналогии с теоремой 2.1. При $k = 2$ энтропия $\mathcal{N}(d,\varepsilon)$ допускает кроме оценки $\mathcal{N}(d,\varepsilon) \leqslant \varepsilon^{-2}$ еще и тривиальную оценку $\mathcal{N}(d,\varepsilon) \leqslant N'' - N'$. Следовательно,

$$\mathcal{N}(d,\varepsilon) \leqslant \min(\varepsilon^{-2}, N'' - N').$$

Здесь энтропийный интеграл неограничен, но допускает оценку:

$$\int\limits_0^1 \big(\mathcal{N}(d,\varepsilon)\big)^{1/2}\, d\varepsilon \leqslant \ln(N'' - N' + 1).$$

Следовательно,

$$P_f(u) \leqslant \frac{v \ln^2 v}{(u/2 - v)^2} + \frac{\ln^2 v}{v}.$$

Выбирая $v = u/3$, мы получим утверждение теоремы 2.3.

Доказательство теоремы 2.4. Заметим, что

$$f_{nN}(x) - f(x) = \sum_{j=1}^{N} (\hat{c}_j - c_j)\, \varphi_j(x) + \sum_{j=N+1}^{\infty} c_j\, \varphi_j(x).$$

Тогда

$$\|f_{nN} - f\|^2_{\mathcal{L}_2[0,1]} = \sum_{j=1}^{N} (\hat{c}_j - c_j)^2 + \rho(N).$$

В соответствии с пунктом 3, мы получаем, что если

$$\|n^{1/2}(\hat{c}_j - c_j)\|_{\mathbf{B}_\varphi} \leqslant 1 \quad (\text{здесь } \ \mathbf{B}_\varphi = B_2, \quad \varphi(t) = t^2/2),$$

то
$$\mathbf{P}\left(n \mid \hat{c}_j - c_j \mid^2 > u\right) \leqslant \frac{1}{2}\exp[-u/2].$$

Нетрудно убедиться, что
$$\sum_{j=1}^{N} \| \hat{c}_j - c_j \|_{\mathbf{B}_2} \leqslant \frac{N}{n^{1/2}}.$$

Следовательно,
$$P_0(u) \leqslant \mathbf{P}\left(\frac{n}{N_{n2}-1}\sum_{j=1}^{N_{n2-1}} (\hat{c}_j - c_j)^2 > (u-v)\right) \leqslant \frac{1}{2}\exp\left[-(u-v)/2\right].$$

Сложим π_{1n}, π_{2n} и $P_0(u)$ и выберем $v = u/2$. Тогда мы получим:
$$P_f(u) \leqslant 2^{2k}\, k^k\, u^{-k/2}.$$

Минимизируя по k, мы завершаем доказательство теоремы 2.4.

Доказательство теоремы 2.5. Вероятность $P_0(u)$ оценивается почти так же как и в теореме 2.4. Здесь величины $n^{1/2}(\hat{c}_j - c_j)$ принадлежат пространству $\mathbf{B}_1$, (т.е. $\varphi_0(t) = \mathbf{E}\exp[t(\zeta^2 - 1)]$, где $\mathcal{L}(\zeta) = \mathcal{N}(0,1)$ и $\varphi(t) = \varphi_0(t)\,\varphi_0(-t)$ при $|t| < 1/2$). Следовательно, в соответствии с пунктом 3, справедливо неравенство:
$$P_0(u) \leqslant \frac{1}{\sqrt{2}}\exp\left[-C(\varphi)(u-v)^{1/2}\right].$$

Складывая вероятности π_{1n}, π_{2n} и $P_0(u)$, при $v = u/2$, мы получаем утверждение теоремы 2.5.

Доказательство теоремы 2.6. Оценим вероятность:
$$P(\varepsilon) = \mathbf{P}\left(\left|\frac{\min\limits_{1\leqslant N\leqslant n} \tau_N}{A_n^*} - 1\right| > \varepsilon\right).$$

Эта вероятность оценивается следующим образом:
$$P(\varepsilon) \leqslant \pi_{1n} + \pi_{2n} + P_0(\varepsilon),$$

где
$$P_0(\varepsilon) = \mathbf{P}\left(\left|\frac{\min\limits_{N_{1n}\leqslant N\leqslant N_{2n}} \tau_N}{A_n^*} - 1\right| > \varepsilon\right).$$

Вероятности π_{1n} и π_{2n} (локализация) определяются так же, как и при доказательстве предыдущих теорем. В этом случае величина v в определении чисел N_{1n} и N_{2n} иная, чем в теоремах 2.1–2.5. Мы имеем:

$$P_0(\varepsilon) \leqslant \mathbf{P}\big(\max_{N_{1n} \leqslant N \leqslant N_{2n}} L_{nN} > \varepsilon A_n^*\big).$$

Используя леммы 4.1 и 4.2, мы получим, что при $\forall\, \varepsilon > 0$

$$P_0(\varepsilon) \leqslant C\, v^{k-1}(n\, A_n^*)^{-1}\, \varepsilon^{-k}.$$

Таким образом, мы имеем:

$$P(\varepsilon) \leqslant C\, k^k\, \mu_{2k}\, v^{-k/2}(n\, A_n^*)^{-1} + C\, v^{k-1}(n\, A_n^*)^{-1}\, \varepsilon^{-k}$$

для задачи регрессии,

$$P(\varepsilon) \leqslant C\, k^k\, v^{-k/2}(n\, A_n^*)^{-1} + C\, v^{k-1}(n\, A_n^*)^{-1}\, \varepsilon^{-k}$$

для задачи оценивания плотности и

$$P(\varepsilon) \leqslant C \exp\big[-C(vn\, A_n^*)^{1/4}\big] + C \exp\big[-C(v)(n\, A_n^*\, \varepsilon)^{1/4}\big]$$

для задачи оценивания спектра. Во всех этих задачах мы получаем: $P(\varepsilon) \to 0$ при $n \to \infty$, так как $n\, A_n^* \to \infty$. Теорема 2.6 доказана.

Замечание 4.1. Пусть для спектральной плотности $f(x)$ сходится ряд

$$\sum_n \exp\big[-C(n\, A_n^*\,)^{1/4}\big].$$

Тогда

$$\frac{\min\limits_{1 \leqslant N \leqslant n} \tau_N}{A_n^*} \xrightarrow{\text{п.н.}} 1.$$

Доказательство этого замечания следует из доказательства теоремы 2.6. Действительно, если ряд $\sum_n \exp\big[-C(n\, A_n^*\,)^{1/4}\big]$ сходится, то

$$\sum_n P_n(\varepsilon) = \sum_n \mathbf{P}\left(\left|\frac{\min\limits_{1 \leqslant N \leqslant n} \tau_N}{A_n^*} - 1\right| > \varepsilon\right) \leqslant$$

$$\leqslant \sum_n C \exp\big[-C(\varepsilon n A_n^*\,)^{1/4}\big] < \infty.$$

Таким образом,

$$\frac{\min\limits_{1\leqslant N\leqslant n} \tau_N}{A_n^*}$$

сходится к 1 почти наверное.

Литература

1. Б. Поляк, А. Цыбаков, *C_P-критерий при проекционном оценивании регрессии.* — Теор. вероятн. и ее прим. **35(2)** (1990), 305–317.
2. Б. Поляк, А. Цыбаков, *Семейство асимптотически оптимальных методов выбора порядка проекционной оценки регрессии.* — Теор. вероятн. и ее прим. **37(3)** (1992), 502–512.
3. D. Donoho, I. Johnstone, *Adapting to unknown smoothness via wavelet shrinkage.* Technical report **425**, Dept. of Stat., Stanford University (1993).
4. D. Donoho, I. Johnstone, G. Kerkyacharian, D. Picard, *Density estimation by wavelet thresholding.* Technical report **426**, Dept. of Stat., Stanford University (1993).
5. S. Efroimovich, *The adaptic algorithm of nonparametrical regression.* Second IFAC Symp. Stoch. Control,. Vilnius, USSR, Preprint 2 (1986), pp. 112–114.
6. M. Nussbaum, *Spline smoothing in regression models and asymptotic efficiency in $\mathcal{L}_2$.* — Ann. Stat. **13(3)** (1985), 984–997.
7. П. Бобров, Е. Островский, *Доверительные интервалы при адаптивном оценивании регрессии.* — Пробл. передачи информ. **32(3)** (1996), 51–59.
8. С. Ефроймович, *Непараметрическое оценивание плотности неизвестной гладкости.* — Теор. вероятн. и ее прим. **30(3)** (1985), 524–534.
9. Г. Голубев, *Непараметрическое оценивание гладких спектральных плотностей гауссовских стационарных последовательностей.* — Теор. вероятн. и ее прим. **38(3)** (1993), 775–786.
10. С. Ермаков, Е. Островский, *Условия непрерывности, экспоненциальные оценки в ЦПТ для случайных полей.* Деп. ВИНИТИ, 3752-B86 (1986).
11. S. Dharmadhikari, K. Jogdeo, *Bounds on moments of certain random wariables.* — Ann. Stat. **40(4)** (1969), 1506–1508.
12. С. Никольский, *Неравенства для целых функций конечной степени и их применения в теории дифференцируемых функций многих переменных.* — Труды Матем. Ин-та им. В. В. Стеклова АН СССР **38** (1951), 244–278.
13. L. Saulis, V. Statulevichius, *Limit theorems for large deviations.* Vilnius: Mokslas Publishers (1989).
14. A. Plikusas, *Some properties of multiple integrals of Ito.* — Liet. Mat. Rink **21** (1981), 163–173.
15. А. Ронжин, *Асимптотические формулы для моментов U-статистик с вырожденными ядрами.* — Теор. вероятн. и ее прим. **27(1)** (1982), 47–55.
16. Ю. Козаченко, Е. Островский, *Банаховы пространства случайных величин типа субгауссовских.* — Теор. вероятн. и матем. стат.. Киев, КГУ, **32**, 1985, 42–53.

17. G. Pizier, *Conditions d'entropie assupant la continuite de certains processus et applications a l'analyse harmonique.* — Sem. d'anal Fonct. **23–24** (1979–1980), 1–43.
18. А. Ширяев, *Вероятность.* Наука, М. (1989).

Bobrov P. B., Ostrovskii E. I. Confidence intervals in adaptive estimation.

In the modern nonparametric statisties an adaptic methods of estimation became in the last time field of investigation. An estimates independent of a priori information about the function from $\mathcal{L}_2$ under estimation (adaptive estimates) are suggested. These estimates are used in the dufferent problems concerning the estimation of the regression, the density and the spectral function.

Институт атомной энергетики

Поступило 18 июня 1997 г.