I. A. Ibragimov, R. Z. Khas'minskii, Estimation of distribution density, *Zap. Nauchn. Sem. LOMI*, 1980, Volume 98, 61–85

И.А.Ибрагимов, Р.З.Хасьминский

ОБ ОЦЕНКЕ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ

1. Введение. Пусть X – случайная величина со значениями в евклидовом пространстве R^k. Пусть X имеет плотность распределения $f(x)$, о которой известно лишь, что она принадлежит некоторому заданному классу функций Σ в R^k. Для оценки неизвестной плотности f производится n независимых наблюдений величины X:

$$X_1, X_2, \ldots, X_n. \tag{1}$$

Погрешность оценки f_n^* неизвестной плотности f по наблюдениям (1) характеризуется математическим ожиданием ее L_p-нормы, $1 \le p \le \infty$, или, более общо, величиной

$$E_f \ell(\| f_n^* - f \|_p),$$

где ℓ – некоторая вещественнозначная функция на прямой.

Ниже мы интересуемся поведением величины

$$\inf_{f_n^*} \sup_{f \in \Sigma} E_f \ell(\| f_n^* - f \|_p), \tag{2}$$

когда $n \to \infty$. Одна из первых работ в этом направлении принадлежит Н.Н.Ченцову [8], см.также [9], § 25; одна из самых недавних выполнена J.Bretagnollé и C.Huber [2]. Наша работа дополняет [2] в следующих отношениях: 1. рассматривается более широкий класс функций потерь; 2. наряду с L_p – нормами, $1 \le p < \infty$ рассматривается также L_∞ -норма; 3. исследуются также и такие классы функций Σ, которые не рассматривались в [2], напр., классы аналитических функций.

2. Верхние границы для функций риска. Напомним, что функция $g(z)$, комплексного переменного $z = (z_1, \ldots, z_k)$ называется целой функцией экспоненциального типа $\nu = (\nu_1, \ldots, \nu_k)$, если $g(z)$ аналитична при всех z и для любого $\varepsilon > 0$ существует положительное число A_ε такое, что для всех z

$$|g(z)| \le A_\varepsilon \exp\left\{ \sum_1^k (\nu_j + \varepsilon)|z_j| \right\}.$$

Обозначим через $\mathcal{M}_{\bar\nu p}(R^k) = \mathcal{M}_{\bar\nu p}$ совокупность всех целых функций экспоненциального типа $\bar\nu$, которые как функции от действительного $x \in R^k$ принадлежат $L_p(R^k)$. Через $\mathcal{M}_{\bar\nu p}$

обозначается ниже также и сужение функций из $\mathfrak{M}_{\bar{\nu}p}$ на R^k.

В дальнейшем существенную роль будет играть следующее неравенство С.М.Никольского, позволяющее оценить $L_{p'}$ норму функции $g \in \mathfrak{M}_{\bar{\nu}p}$ через ее L_p норму, $p < p'$.

ЛЕММА I. (см. [5], стр.150). Если $1 \le p \le p' \le \infty$, то для функции $g \in \mathfrak{M}_{\bar{\nu}p}$ имеет место неравенство

$$\|g\|_{p'} \le 2^k \left(\prod_1^k \nu_j\right)^{1/p - 1/p'} \|g\|_p, \quad \bar{\nu} = (\nu_1, \dots, \nu_k).$$

Для заданной функции $\varphi \in L_p(R^k)$ через $\mathcal{E}_{\bar{\nu}}^{(p)}(\varphi)$ обозначим величину наилучшего приближения в норме L_p функции φ целыми функциями экспоненциального типа $\bar{\nu}$, т.е.

$$\mathcal{E}_{\bar{\nu}}^{(p)}(\varphi) = \inf_{g \in \mathfrak{M}_{\bar{\nu}p}} \|\varphi - g\|_p .$$

Положим далее

$$\lambda(\bar{\nu}, \Sigma, p) = \sup_{\varphi \in \Sigma} \mathcal{E}_{\bar{\nu}}^{(p)}(\varphi),$$

и, когда это не может вызвать недоразумения, вместо $\lambda(\bar{\nu}, \Sigma, p)$ будем писать просто $\lambda(\bar{\nu})$.

Наконец, если $\bar{\nu} = (\nu, \dots, \nu)$, будем вместо $\bar{\nu}$ писать всюду просто ν, т.е. $\mathcal{E}_{\nu}^{(p)}$ вместо $\mathcal{E}_{\bar{\nu}}^{(p)}$ и т.д.

Рассмотрим далее ядро (аналог ядра Валле-Пуссена, см. [5], стр.358).

$$V(x) = \prod_1^k \frac{\cos x_j - \cos 2x_j}{\pi x_j^2}, \quad x = (x_1, \dots, x_k);$$

$$V_\nu(x) = \prod_1^k \frac{\cos \nu x_j - \cos 2\nu x_j}{\pi \nu x_j^2} = \nu^k V(\nu x).$$

Очевидно $V_\nu \in \mathfrak{M}_{\nu p}$ для всех p с $1 \le p \le \infty$.

Нам понадобятся ниже следующие свойства ядер V_ν.

ЛЕММА 2. Пусть функция $\varphi \in L_p(R^k)$. Тогда функция $\Psi = \int_{R^k} V_\nu(x-y)\varphi(y)\,dy$ принадлежит $\mathfrak{M}_{2\nu p}$ и

$$\left\| \int_{R^k} V_\nu(x-y)\varphi(y)\,dy \right\|_p \le \|V\|_1 \|\varphi\|_p, \quad 1 \le p \le \infty \tag{3}$$

Доказательство (3) см. в [5], стр.32; включение $\Psi \in \mathfrak{M}_{2\nu p}$ следует из [5], стр.162.

ЛЕММА 3 (см. [5], стр.360). Если функция $\varphi \in L_p(R^k)$, то

$$\left\| \varphi(x) - \int_{R^k} V_\nu(x-y)\varphi(y)\,dy \right\|_p \leq (1+\|V\|_1)\mathcal{E}_\nu^{(p)}(\varphi) \leq$$

$$\leq \left(1+\left(\frac{4}{\pi^2}\ln 3 + 2\right)^k\right)\mathcal{E}_\nu^{(p)}(\varphi) \leq (1+3^k)\,\mathcal{E}_\nu^{(p)}(\varphi). \qquad (4)$$

Относительно оценки $\|V\|_1$ см. [1], стр.254.

Определим теперь следующую оценку σ_ν^* для неизвестной плотности f:

$$\sigma_\nu^*(x) = \frac{1}{n}\sum_1^n V_\nu(x-X_j). \qquad (5)$$

Если ввести эмпирическое распределение

$$\mathcal{P}_n(A) = \frac{1}{n}\sum_1^n \mathbb{1}_A(X_j),$$

то оценку (5) можно записать в виде

$$\sigma_\nu^*(x) = \int_{R^k} V_\nu(x-y)\,\mathcal{P}_n(dy). \qquad (6)$$

ТЕОРЕМА 1. Пусть плотность $f \in \Sigma$, $\Sigma \subset L_p(R^k)$, $2 \leq p < \infty$. Рассмотрим оценку $\sigma_\nu^*(X)$, где ν выбрано так, чтобы $\frac{\nu^k p\, r}{n} \leq 1/2$ или $\frac{\nu^k p'}{n} \leq 1$. Тогда

$$E_f\|f-\sigma_\nu^*\|_p^r \leq A^{kr}(r+1)^{\frac{r+1}{2}}\left(1+\|f\|_p\right)^{\frac{r+1}{2}}\left(\left(\mathcal{E}_\nu^{(p)}(f)\right)^r + \left(\frac{\nu^k}{n}\right)^{r/2}\right), \text{ если } \frac{\nu^k p(r+1)}{2n} \leq 1,$$

$$(7)$$

$$E_f\|f-\sigma_\nu^*\|_p^r \leq A^{kr}\left(\frac{r+1}{|\ln(r+1)|}\right)^{r+1}\left(1+\|f\|_p\right)^{\frac{r+1}{2}}\left(\left(\mathcal{E}_\nu^{(p)}(f)\right)^r + \left(\frac{\nu^k}{n}\right)^{r/2}\right), \text{ если } \frac{\nu^k p}{n} \leq 1.$$

Здесь и ниже через A обозначаются абсолютные константы, не обязательно одни и те же.

ТЕОРЕМА 2. Пусть плотность $f \in \Sigma \subset L_\infty(R^k)$. Рассмотрим оценку $\sigma_\nu^*(X)$, где ν выбрано так, чтобы $\frac{\nu^k k \ln \nu r}{n} \leq 1/2$.

Тогда

$$E_f\|f-\sigma_\nu^*\|_\infty^r \le A^{kr}\Big(\big(\varepsilon_\nu^{(\infty)}(f)\big)^r+\Big(\frac{\nu^k k\ln\nu}{n}\Big)^r\Big)(r+1)^{\frac{r+1}{2}}\big(1+\|f\|_\infty\big)^{\frac{r+1}{2}}. \tag{8}$$

Если же $\frac{\nu^k k\ln\nu}{n}\le 1$, то

$$E_f\|f-\sigma_\nu^*\|_\infty^r \le A^{kr}\Big(\frac{r+1}{|\ln(r+1)|}\Big)^{r+1}\big(1+\|f\|_\infty\big)^{\frac{r+1}{2}}\Big(\big(\varepsilon_\nu^{(\infty)}(f)\big)^r+\Big(\frac{\nu^k k\ln\nu}{n}\Big)^{r/2}\Big).$$

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 1. Условимся ниже для краткости вместо $E_f(\cdot)$ писать просто $E(\cdot)$. Очевидно,

$$E\|f-\sigma_\nu^*\|_p^r \le 2^{r-1}\Big(\|f-E\sigma_\nu^*\|_p^r+E\|\sigma_\nu^*-E\sigma_\nu^*\|_p^r\Big). \tag{9}$$

Первое слагаемое в (9), смещение, оценивается совсем просто. Так как

$$E\sigma_\nu^*(x)-E\{V_\nu(x-X_1)\}=\int_{R^k}V_\nu(x-y)f(y)\,dy,$$

то по лемме 2

$$\|f-E\sigma_\nu^*\|_p^r \le \big(1+\|V\|_1\big)^r\big(\varepsilon_\nu^{(p)}(f)\big)^r. \tag{10}$$

Оценка второго слагаемого в (9) составляет содержание следующей леммы:

ЛЕММА 4. Справедливы следующие неравенства: если $2\le p<\infty$ то

$$E\|\sigma_\nu^*-E\sigma_\nu^*\|_p^r \le \begin{cases} A^{rk}(r+1)^{\frac{r+1}{2}}\big(1+\|f\|_p\big)^{r/2}\Big(\frac{\nu^k}{n}\Big)^{r/2}, & \frac{\nu^k(r+1)p}{2n}\le 1 \\ A^{rk}\Big(\frac{r+1}{\ln(r+1)}\Big)^{r+1}p^{r/2}\big(1+\|f\|_p\big)^{r/2}\Big(\frac{\nu^k}{n}\Big)^{r/2}, & \frac{\nu^k p}{n}\le 1 \end{cases} \tag{11}$$

если же $p=\infty$, то

$$E\|\sigma_\nu^*-E\sigma_\nu^*\|_\infty^r \leqslant \begin{cases} A^{rk}k^{r/2}(r+1)^{\frac{(r+1)}{2}}(1+\|f\|_\infty)^{r/2}\left(\frac{\nu^k \ln\nu}{n}\right)^{r/2}, & \frac{\nu^k k(r+1)\ln\nu}{n}<1, \\ A^{rk}k^{r/2}\left(\frac{r+1}{\ln(r+1)}\right)^{\frac{r+1}{2}}\left(\frac{\nu^k\ln\nu}{n}\right)^{r/2}(1+\|f\|_\infty)^r, & \frac{k\nu^k\ln\nu}{n}<1, \end{cases} \tag{12}$$

ДОКАЖЕМ ЛЕММУ 1. Предположим сперва, что ρ – целое четное число, r – целое. Тогда

$$E\|\sigma_\nu^*-E\sigma_\nu^*\|_\rho^r \leqslant E^{1/\rho}\|\sigma_\nu^*-E\sigma_\nu^*\|_\rho^{r\rho} =$$
$$= n^{-r}\Big(\int_{R^k}\ldots\int_{R^k} dx_1\ldots dx_r\, E\prod_{j=1}^{r}\Big(\sum_{i=1}^{n}(V_\nu(x_j-X_i)- \tag{13}$$
$$-EV_\nu(x_j-X_i))\Big)^\rho\Big)^{1/\rho}.$$

Далее

$$E\Big\{\prod_{j=1}^{r}\Big(\sum_{i=1}^{n}(V_\nu(x_j-X_i)-EV_\nu(x_j-X_i))\Big)^\rho\Big\} =$$

$$=E\Big\{\prod_{j=1}^{r}\sum_{k_{j1}+\ldots+k_{jn}=\rho}\frac{\rho!}{k_{j1}\ldots k_{jn}!}\prod_{i=1}^{n}\big(V_\nu(x_j-X_i)-EV_\nu(x_j-X_i)\big)^{k_{ji}}\Big\}= \tag{14}$$

$$=\frac{(\rho!)^r}{(\rho r)!}\sum_{\substack{k_1+\ldots+k_n=\rho r\\ k_{i1}+\ldots+k_{ir}=k_i}}\prod_{i=1}^{n}\frac{1}{k_i!}E\Big\{k_i!\prod_{j=1}^{r}\frac{1}{k_{ji}!}\big(V_\nu(x_j-X_i)-EV_\nu(x_j-X_i)\big)^{k_{ji}}\Big\}.$$

Всякий раз, когда одно из чисел k_{ji}, $j=1,\ldots,r$ равно единице, а остальные нулю, т.е., когда $k_i=\sum_{j=1}^{r}k_{ji}=1$, необходимо

$$E\prod_{j=1}^{r}\big(V_\nu(x_j-X_i)-EV_\nu(x_j-X_i)\big)=0.$$

Кроме того, из неравенства Гельдера и элементарного неравенства

$$(|a|+|b|)^k \leqslant 2^{k-1}(|a|^k+|b|^k).$$

следует, что

$$\Big|E\prod_{j=1}^{r}\big(V_\nu(x_j-X_i)-EV_\nu(x_j-X_i)\big)^{k_{ji}}\Big|\leqslant$$

$$\leqslant 2^{\sum_{j=1}^{r}k_{ji}}E\prod_{1}^{r}\big|V_\nu^{k_{ji}}(x_j-X_i)\big|.$$

Подставляя последнее неравенство в (14), и заменяя факториалы по формуле Стирлинга, найдем, что

$$E\Big\{\prod_{j=1}^{r}\Big(\sum_{i=1}^{n}V_\nu(x_j-X_i)-EV_\nu(x_j-X_i)\Big)^{\rho}\Big\}\leqslant$$

$$\leqslant(2e\pi\rho)^{r/2}\Big(\frac{2}{r}\Big)^{\rho r}\sum_{\substack{k_1+\ldots+k_n=\rho r\\ k_i\neq 1}}\frac{(\rho r)!}{k_1!\ldots k_n!}E\Big\{\sum_{k_{i1}+\ldots+k_{ir}=k_i}\frac{k_i!}{k_{i1}!\ldots k_{ir}!}\cdot$$

$$\cdot\prod_{j=1}^{r}\big|V_\nu^{k_{ij}}(x_j-X_i)\big|\Big\}\leqslant$$

$$\leqslant(2e\pi\rho)^{r/2}\Big(\frac{2}{r}\Big)^{\rho r}\sum_{l=1}^{\min(n,\frac{\rho r}{2})}C_n^l\sum_{\substack{k_1+\ldots+k_l=\rho r\\ k_i\geqslant 2}}\frac{(\rho r)!}{k_1!\ldots k_l!}E\Big\{\prod_{i=1}^{l}\Big(\sum_{i=1}^{r}|V_\nu(x_j-X_i)|\Big)^{k_i}\Big\}.$$

Следовательно,

$$E^{\rho}\|\sigma_\nu^*-E\sigma_\nu^*\|_\rho^r\leqslant(2e\pi\rho)^{r/2}\Big(\frac{2}{nr}\Big)^{\rho r}\sum_{l=1}^{\min(n,\frac{\rho r}{2})}C_n^l\times$$

$$\times\sum_{\substack{k_1+\ldots+k_l=\rho r\\ k_i\geqslant 2}}\frac{(\rho r)!}{k_1!\ldots k_l!}\sum_{\Sigma k_{ji}=k_i}\frac{k_1!\ldots k_l!}{k_{11}!\ldots k_{rn}!}E\Big\{\int_{R^k}\ldots\int_{R^k}dx_1\ldots dx_r\prod_{i=1}^{l}\prod_{j=1}^{r}|V_\nu(x_j-X_i)|^{k_{ji}}\Big\}. \qquad (15)$$

Далее, если $r\geqslant l$, то, применяя сперва неравенство Гельдера, а затем неравенство Минковского для интегралов, найдем, что

$$E\Big\{\int_{R^k}\ldots\int_{R^k}dx_1\ldots dx_r\prod_{i=1}^{l}\prod_{j=1}^{r}|V_\nu(x_j-X_i)|^{k_{ji}}\Big\}=$$

$$=\int_{R^k}\ldots\int_{R^k}dx_i\ldots dx_r\prod_{i=1}^{l}E\Big\{\prod_{j=1}^{r}|V_\nu(x_j-X_i)|^{k_{ji}}\Big\}\leqslant$$

$$\leq \prod_{i=1}^{\ell}\left\{\int_{R^k}\ldots\int_{R^k} dx_1\ldots dx_r E^{\ell}\left\{\prod_{j=1}^{r}|V_\nu(x_j-X_i)|^{k_{ji}}\right\}\right\}^{1/\ell} \leq$$

$$\leq \prod_{i=1}^{\ell} E\left\{\left(\int_{R^k}\ldots\int_{R^k} dx_1\ldots dx_r \prod_{j=1}^{r}|V_\nu(x_j-X_i)|^{k_{ji}\ell}\right)^{1/\ell}\right\} = \tag{16}$$

$$= \prod_{i=1}^{\ell}\prod_{j=1}^{r}\left(\int_{R^k}|\nu V(\nu x)|^{k_{ji}\ell}dx\right)^{1/\ell} \leq \prod_{i=1}^{\ell}\prod_{j=1}^{r}\left\{\nu^{(k_{ji}-1/\ell)k}\|V\|_\infty^{k_{ji}-1/\ell}\|V\|_1^{1/\ell}\right\} \leq$$

$$\leq \|V\|_\infty^{\rho r-r}\|V\|_1^{r}\nu^{(\rho r-r)k} \leq \|V\|_\infty^{\rho r-r}\|V\|_1^{r}\nu^{(\rho r-\ell)k} \leq 3^{\rho r k}\nu^{(\rho r-\ell)k}$$

Если же $\ell>r$, то, снова используя неравенства Гельдера и Минковского, получим, что

$$E\left\{\int_{R^k}\ldots\int_{R^k} dx_1\ldots dx_r \prod_{i=1}^{\ell}\prod_{j=1}^{r}|V_\nu(x_j-X_i)|^{k_{ji}}\right\} \leq$$

$$\leq \prod_{j=1}^{r}\int_{R^k} dx \prod_{i=1}^{\ell}\left(\int_{R^k}|V_\nu(x-y)|^{rk_{ji}}f(y)\,dy\right)^{1/r} \leq \tag{17}$$

$$\leq \prod_{j=1}^{r}\prod_{i=1}^{\ell}\left(\int_{R^k} dx\left(\int_{R^k}|V_\nu(y)|^{rk_{ji}}f(x-y)dy\right)^{\ell/r}\right)^{1/\ell} \leq$$

$$\leq \prod_{j=1}^{r}\prod_{i=1}^{\ell}\left(\int_{R^k}(f(x))^{\ell/r}dx\right)^{1/\ell}\left(\int_{R^k}(\nu|V(\nu y)|)^{rk_{ji}}dy\right)^{1/r} \leq$$

$$\leq \|f\|_{\ell/r}^{\ell}\prod_{1}^{r}\prod_{1}^{\ell}\nu^{k_{ji}-1/r}\|V\|_\infty^{k_{ji}-1/r}\|V_1\|^{1/r} =$$

$$= \nu^{(\rho r-\ell)k}\|f\|_{\ell/r}^{\ell}\|V\|_\infty^{\rho r-\ell}\|V\|_1^{\ell} \leq 3^{\rho r k}\nu^{(\rho r-\ell)k}\left(1+\|f\|_{\rho/2}\right)^{\frac{\rho r}{2}}$$

Подставляя (16) и (17) в (15), получим, что

$$E^{\rho}\|\sigma_\nu^*-E\sigma_\nu^*\|_\rho^{r} \leq (2e\pi\rho)^{r/2}\left(\frac{2\cdot 3^k}{nr}\right)^{\rho r}\left(1+\|f\|_{\rho/2}\right)^{\frac{\rho r}{2}}.$$

$$\sum_{l=1}^{\min(n,\frac{pr}{2})} \frac{n^l}{l!}\nu^{(pr-l)k} l^{pr} r^{pr} = (2\pi p)^{r/2}\left(2e3^k\right)^{pr}\left(1+\|f\|_{p/2}\right)^{\frac{pr}{2}} \cdot \sum_{l=1}^{\min(n,\frac{pr}{2})}\left(\frac{\nu^k l}{n}\right)^{pr-l} \tag{18}$$

Следовательно, если $\frac{\nu^k pr}{2n} \leq 1$, то

$$E\|\sigma_\nu^* - E\sigma_\nu^*\|_p^r \leq \left(2e^3 3^k\right)^r\left(\frac{\nu^k}{n}\right)^{r/2}(pr)^{r/2}\left(1+\|f\|_{p/2}\right)^{r/2} = A^{rk}\left(1+\|f\|_{p/2}\right)^{r/2}(pr)^{r/2}\left(\frac{\nu^k}{n}\right)^{r/2}, \tag{19}$$

а также

$$E\|\sigma_\nu^* - E\sigma_\nu\|_p^r \leq A^{rk}\left(1+\|f\|_p\right)^{r/2}(pr)^{r/2}\left(\frac{\nu^k}{n}\right)^{r/2}. \tag{20}$$

Далее, легко проверяется, что если $1 \leq l \leq \frac{pr}{2}$, то

$$\left(\frac{l}{p}\right)^{pr-l} \leq \left(\frac{er}{\ln r}\right)^{pr}$$

Отсюда и из (18) следует, что для $\frac{\nu^k p}{n} < 1$

$$E\|\sigma_\nu^* - E\sigma_\nu^*\|_p^r \leq A^{kr}\left(\frac{r}{\ln r}\right)^2\left(\frac{\nu^k p}{n}\right)^{r/2}\left(1+\|f\|_{p/2}\right)^{r/2} \leq A^{kr}\left(\frac{r}{\ln r}\right)^r\left(\frac{\nu^k p}{n}\right)^{r/2}\left(1+\|f\|_p\right)^{r/2}. \tag{21}$$

Таким образом, для четных p и целых r лемма доказана.

2. Если p по-прежнему целое четное число, а r произвольное положительное число, то полагая $s = [r]+1 > r$, найдем

$$E\|\sigma_\nu^* - E\sigma_\nu^*\|_p^r \leq E^{r/s}\|\sigma_\nu^* - E\sigma_\nu^*\|_p^s .$$

Следовательно, неравенства (19),(20),(21) справедливы и для этого случая.

3. Пусть теперь $2 \leq p < \infty$, $r > 0$. Выберем целое

число m так, что $2m \le p < 2m+2$. По неравенству Гельдера

$$\|\sigma_\gamma^* - E\sigma_\gamma^*\|_p \le \|\sigma_\gamma^* - E\sigma_\gamma^*\|_{2m}^{\frac{2m+2-p}{2}} \|\sigma_\gamma^* - E\sigma_\gamma^*\|_{2m+2}^{\frac{p-2m}{2}} . \qquad (22)$$

Из последнего неравенства и результата предыдущего пункта (неравенства (19),(21) уже следует утверждение леммы для всех $\tau > 0$ и всех $p < \infty$.

4. Рассмотрим случай $p = \infty$. Так как функции $\sigma_\gamma^*(x)$ и $E\sigma_\gamma^*(x)$ принадлежат $\mathfrak{M}_{2\gamma p}$, $1 \le p \le \infty$ (см. лемму 2), то по неравенству С.М.Никольского (лемма 1)

$$\|\sigma_\gamma^* - E\sigma_\gamma^*\|_\infty^\tau \le 2^{k\tau} \gamma^{\frac{k\tau}{p}} \|\sigma_\gamma^* - E\sigma_\gamma^*\|_p^\tau . \qquad (23)$$

Выберем здесь $p = k\ln\gamma$. Если $n \ge \frac{\tau k \ln\gamma\, \gamma^k}{2}$, то из (20) и (23) следует неравенство

$$E\|\sigma_\gamma^* - E\sigma_\gamma^*\|_\infty^\tau \le A^{\tau k} k^{\tau/2} (\tau)^{\tau/2} \left(\frac{\gamma^k \ln\gamma}{n}\right)^{\tau/2} . \qquad (24)$$

Аналогичным образом неравенство (21) приводит к неравенству: если $\frac{k\gamma^k \ln\gamma}{n} \le 1$, то

$$E\|\sigma_\gamma^* - E\sigma_\gamma^*\|_\infty^\tau \le A^{k\tau} k^{\tau/2} \left(\frac{\gamma^k \ln\gamma}{n}\right)^{\tau/2} \left(\frac{\tau}{\ln\tau}\right)^\tau .$$

Лемма 4 доказана.

Теоремы 1 и 2 следуют немедленно из неравенства (9) и оценок (10),(11) (теорема 1) или (10),(12) (теорема 2).

4. Примеры. Поскольку многие функциональные классы $\sum$ гладких функций полностью характеризуются поведением их наилучших приближений, теоремы 1 и 2 позволяют указать для таких классов порядок приближения неизвестной плотности f оценками σ_γ^*. Приведем несколько примеров.

1. Класс $H_p^\beta L$ (см. [5], стр.189, для целых β наше определение слегка упрощено). Пусть $\beta = \tau + \alpha$, где $\tau \ge 0$ – целое и $0 < \alpha \le 1$, $L > 0$. Класс $H_p^\beta L$ состоит из функций $g \in L_p$, имеющих частные производные до порядка τ включительно, и для любой такой производной $g^{(\tau)}$

$$\|g^{(\tau)}(x+h) - g^{(\tau)}(x)\|_p \le L|h|^\alpha$$

ЛЕММА 5 ([5] , стр.244). Пусть $g \in H_p^\beta L$, тогда

$$\mathcal{E}_\gamma^\rho(g) \leq \frac{cL}{\gamma^\beta}, \tag{25}$$

где константа C зависит только от k и β .

ТЕОРЕМА 3. Пусть неизвестная плотность распределения $f \in \Sigma \subset H_p^\beta L$, $2 \leq \rho < \infty$ и $\sup\limits_{\Sigma} \|f\|_\rho = M < \infty$ Тогда найдется такая последовательность оценок σ_γ^*, $\gamma = \gamma(n)$ плотности f , что для $n > \left(\frac{\rho z}{2}\right)^{\frac{2\beta+k}{2\beta}} (CL)^{k/\beta}$

$$\sup_{f \in \Sigma} E_f \|f - \sigma_\gamma^*\|_\rho^z \leq A^{kz} (cL)^{\frac{kz}{k+2\beta}} (1+M)^{z/2} (\rho(z+1))^{\frac{z+1}{2}} n^{-\frac{z\beta}{2\beta+k}},$$

где постоянные A и C те же, что и в неравенствах (7) и (25).

ДОКАЗАТЕЛЬСТВО. Обращаясь к неравенству (7) теоремы 1, оценим в нем $\mathcal{E}_\gamma^{(\rho)}(f)$ с помощью леммы 5 . Выберем теперь γ ближайшим к корню уравнения

$$\frac{c^z L^z}{\gamma^{\beta z}} = \left(\frac{\gamma^k}{n}\right)^{z/2}.$$

При таком выборе γ

$$\left(\mathcal{E}_\gamma^\rho(f)\right)^z + \left(\frac{\gamma^k}{n}\right)^{z/2} \leq A n^{-\frac{\beta z}{k+2\beta}} (cL)^{\frac{kz}{k+2\beta}}.$$

Теорема доказана. Аналогично доказывается и

ТЕОРЕМА 4. Пусть неизвестная плотность распределения $f \in \Sigma \subset H_\infty^\beta L$ и $\sup\limits_{\Sigma} \|f\|_\infty = M < \infty$. Тогда найдется такая последовательность оценок σ_γ^* , $\gamma = \gamma(n)$, плотности f , что для больших n

$$\sup_{f \in \Sigma} E_f \|f - \sigma_\gamma^*\|_\rho^z \leq A^{kz} (cL)^{zk/(k+2\beta)} z^{\frac{z+1}{2}} C_1 \left(\frac{\ln n}{n}\right)^{\frac{\beta z}{k+2\beta}} (1+M)^{\frac{z+1}{2}},$$

где постоянные A и C те же, что в (7) и (25), а постоянная C_1 зависит только от k и β .

Таким образом, мы видим, что для $f \in H_\rho^\beta L$ сближение оценки σ_γ^* с f происходит со скоростью $n^{-\frac{\beta}{k+2\beta}}$, если $2 \leq \rho < \infty$, и со скоростью $\left(\frac{n}{\ln n}\right)^{\frac{-\beta}{k+2\beta}}$, если $\rho = \infty$; во всяком случае если мерить уклонение σ_γ^* от f

с помощью $\|\cdot\|_p^r$. Покажем, что этот результат сохраняется и для более общих функций потерь.

ТЕОРЕМА 5. Пусть $f\in\Sigma\subset H_p^\beta L$, $\sup_{f\in\Sigma}\|f\|_p=M<\infty$. Пусть $\ell(y)$, $y\geq 0$ неотрицательная функция, удовлетворяющая неравенству

$$\ell(y)\leq he^{Hy}, \tag{26}$$

где h и H некоторые неотрицательные величины. Для любой такой функции ℓ можно указать последовательность оценок $\sigma_\nu^*(x), \nu=\nu(n)$, что

$$\begin{gathered}\overline{\lim_n}\ \sup_f E_f\,\ell\left(n^{\frac{\beta}{k+2\beta}}\|f-\sigma_\nu^*\|_p\right)<\infty,\quad 2\leq p<\infty\ ;\\ \overline{\lim_n}\ \sup_f E_f\,\ell\left(\left(\frac{n}{\ln n}\right)^{\frac{\beta}{k+2\beta}}\|f-\sigma_\nu^*\|_\infty\right)<\infty .\end{gathered}\tag{27}$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим для определенности случай $p<\infty$. Достаточно проверить, что можно подобрать ν так, чтобы

$$\sup_n\sup_f E_f\exp\left\{Hn^{\frac{\beta}{k+2\beta}}\|f-\sigma_\nu^*\|_p\right\}<\infty .$$

Выберем $\nu=\gamma n^{\frac{1}{k+2\beta}}$, где γ достаточно малое число, выбранное так, чтобы $\frac{\nu^k p}{n}\leq 1$.

По лемме 5

$$\mathcal{E}_\nu^{(p)}(f)\leq\frac{cL}{\gamma^\beta}n^{-\frac{\beta}{k+2\beta}}$$

в силу теоремы 1 для $r>1$

$$E\|f-\sigma_\nu^*\|_p^r\leq C^r n^{\frac{-\beta r}{k+2\beta}}\left(\frac{r}{\ln(r+1)}\right)^r, \tag{28}$$

где постоянная C не зависит ни от r, ни от n. Следовательно,

$$E_f\exp\left\{Hn^{\frac{\beta}{k+2\beta}}\|f-\sigma_\nu^*\|_p\right\}\leq$$
$$\leq\sum\frac{H^rC^r}{r!}\left(\frac{r}{\ln(r+1)}\right)^r\leq\sum_r\left(\frac{e^2HC}{\ln(r+1)}\right)^r<\infty .$$

Теорема доказана.

2. Классы аналитических функций. Ограничимся случаем $k=1$. Скажем, что функция $g\in L_p(R^1)$ принадлежит классу $A_p^\lambda L$, если она допускает аналитическое продолжение в полосу $|\Im m z|<L$ комплексного переменного $z=x+iy$ и

$$\|g(\cdot\pm i\lambda)\|_p\le L .$$

Если $g\in A_p^\lambda L$, то (см. [I], стр.267)

$$\mathcal{E}_\nu^p(g)\le CLe^{-\lambda\nu} . \tag{29}$$

Оценивая в неравенствах (7) $\mathcal{E}_\nu^{(p)}(f)$ с помощью (29), аналогично тому, как это было сделано выше, придем к следующим результатам.

ТЕОРЕМА 6. Пусть плотность $f\in\Sigma\subset A_p^\lambda L$. Тогда найдется такая последовательность оценок σ_ν^*, $\nu=\nu(n)$, что

$$\sup_f E_f\|f-\sigma_\nu^*\|_p^r\le\left(\frac{\ln n}{n}\right)^{r/2}, \quad 2\le p<\infty ;$$

$$\sup_f E_f\|f-\sigma_\nu^*\|_\infty^r\le C\left(\frac{\ln n\ \ln\ln n}{n}\right)^{r/2}, \tag{30}$$

где постоянная C зависит только от r, L и λ. Наконец, подобно теореме 5 доказывается следующая

ТЕОРЕМА 7. Пусть плотность распределения $f\in\Sigma\subset A_p^\lambda L$, а функция потерь l удовлетворяет неравенству (26). Тогда найдется такая последовательность оценок σ_ν^*, что

$$\sup_n\sup_f E_f\, l\left(\sqrt{\frac{n}{\ln n}}\,\|f-\sigma_\nu^*\|_p\right)<\infty, \quad 2\le p<\infty ;$$

$$\sup_n\sup_f E_f\, l\left(\sqrt{\frac{n}{\ln n\ \ln\ln n}}\,\|f-\sigma_\nu^*\|_\infty\right)<\infty .$$

3. Класс целых функций экспоненциального типа $\tau=(\tau_1,\dots,\tau_k)$. Допустим, что плотность $f\in\Sigma\subset\mathcal{M}_{\bar\tau p}$, $\bar\tau=(\tau_1,\dots,\tau_k)$. Тогда для всех $\nu>\max_j\tau_j$

$$\int_{R^k}V_\nu(x-y)f(y)\,dy=f(x) ,$$

т.е.

$$E_f \sigma_\nu^*(x) \equiv f(x),$$

и, следовательно, $n > \rho\nu^k$

$$\sup_f E\|f-\sigma_\nu^*\|_\rho = E\|\sigma_\nu^* - E\sigma_\nu^*\|_\rho \leq A^k\left(\frac{\nu^k}{n}\right)^{1/2}. \qquad (31)$$

Хотя множество Σ, вообще говоря, нельзя вложить ни в какое конечномерное пространство, неожиданным образом оказывается, что порядок приближения здесь такой же, как и для конечномерных параметрических семейств. Заметим, что слегка меняя оценку σ_ν^*, можно в правой части неравенства (31) получить величину вида $A^k\left(\frac{\tau_1\ldots\tau_k}{n}\right)^{1/2}$, т.е. произведение $\tau_1\ldots\tau_k$ играет роль размерности.

ЗАМЕЧАНИЕ. Во всех доказанных теоремах предполагалось, что $\rho \geq 2$. В случае $1 \leq \rho < 2$ получается та же величина скорости сходимости, что и в разобранных случаях ($n^{-\frac{\beta}{2k+\beta}}$, если $f \in H_\rho^\beta L$ и т.д.), если дополнительно предположить, что $\|f\|_{\rho/2}$ равномерно ограничены в Σ. Мы не будем останавливаться здесь на этом вопросе, отметив лишь, что ограничения на норму $\|f\|_{\rho/2}$ диктуются существом дела. Это будет доказано ниже, см. теорему II.

4. Нижние границы для функций риска. Мы покажем, что установленные выше границы сверху для функций риска не могут быть существенно усилены, во всяком случае в обстановке примеров п.3. К этим примерам мы вернемся в следующем п.5, а здесь докажем один сравнительно общий результат. Идеи лежащие в основе доказательства сходны с использованными в [4], [7].

Именно, обозначим Λ класс функций потерь $\ell: R^1 \to R^1$ четных, монотонно возрастающих на $[0,\infty)$ и таких, что $\ell(0)=0$, $\ell \not\equiv 0$. Допустим далее, что в множестве Σ удается выделить $\rho(\delta)$ плотностей $f_{i\delta}$ так, что $\|f_{i\delta} - f_{j\delta}\| \geq \delta$ и $\ell(\frac{\delta}{2}) > 0$. Тогда, разумеется, для любой оценки $\tilde{f}_n$ плотности f по наблюдениям $X_1,\ldots,X_n$ справедливо неравенство

$$\sup_{f\in\Sigma} E_f \ell(\|f-\tilde{f}_n\|) \geq \sup_{f_{i\delta}} E_f \ell(\|f_{i\delta}-\tilde{f}_n\|) \geq \qquad (32)$$

$$\geqslant \frac{\ell\left(\frac{\delta}{2}\right)}{\rho(\delta)} \sum P_{f_{i\delta}} \left\{ \| f_{i\delta} - \tilde{f}_n \| \geqslant \frac{\delta}{2} \right\}.$$

Таким образом, нужно суметь оценить снизу величину

$$P_e = \frac{1}{\rho(\delta)} \sum P_{f_{i\delta}} \left\{ \| f_{i\delta} - \tilde{f}_n \| \geqslant \frac{\delta}{2} \right\},$$

причем в силу неравенства $\| f_{i\delta} - f_{j\delta} \| \geqslant \delta$ можно ограничиться оценками $\tilde{f}_n$, принимающими лишь значения $f_{i\delta}$. Т.е. P_e есть минимальная вероятность ошибки при проверке сложной гипотезы $\{\eta = f_{i\delta}\}$, где случайная величина η принимает значения $f_{i\delta}$ с вероятностями $\frac{1}{\rho(\delta)}$, по наблюдениям $X_1, \ldots, X_n$. Вероятность ошибки можно оценить с помощью леммы Фэно (см. [4], стр.428):

$$P_e \geqslant 1 - \frac{I(\eta ; (X_1, \ldots, X_n))}{\ln \rho(\delta)}, \tag{33}$$

где вообще $I(\xi_1 ; \xi_2)$ количество информации по Шеннону между случайными элементами $\xi_1, \xi_2, \ldots$

Нетрудно показать (см. [7]), что

$$I(\eta ; (X_1, \ldots, X_n)) \leqslant n I(\eta ; X_1)$$

так что

$$P_e \geqslant 1 - n \frac{I(\eta ; X_1)}{\ln \rho(\delta)}. \tag{34}$$

Это последнее неравенство и будет положено в основу нашего доказательства. Нужно лишь оценить I и ρ в терминах, допускающих проверку для семейств гладких функций.

ТЕОРЕМА 8. Пусть множество $\sum$ плотностей из R^k таково, что при любом $\delta > 0$ существуют плотности $f_{0\delta}$ и $f_{i\delta} = f_i$, $i = 1, 2, \ldots, \rho(\delta)$ из $\sum$ такие, что

1) $\| f_{i\delta} - f_{j\delta} \|_\rho \geqslant \delta, \qquad i, j = 1, 2, \ldots, \rho(\delta),$

2) $\left\| \frac{f_{i\delta}}{f_{0\delta}} - 1 \right\|_\infty \leqslant C_1 < 1,$

где положительная постоянная C_1 не зависит от δ.

Пусть

$$\delta(n,\Sigma)=\sup\left\{\delta:\sup_{i=\overline{1,\rho(\delta)}}\frac{\left\|\frac{f_{i\delta}-f_{o\delta}}{\sqrt{f_{o\delta}}}\right\|_2^2}{\ln\rho(\delta)}\leqslant\frac{c_2}{n}\right\}.$$

Тогда для любой $l\in\Lambda$

$$\inf_{\tilde f_n}\sup_{f\in\Sigma}E_f\,l\left(\frac{\|f-\tilde f_n\|}{\delta(n,\Sigma)}\right)\geqslant l\left(\frac{1}{2}\right)\left(1-\frac{c_2}{2(1-c_1)}\right).$$

ДОКАЗАТЕЛЬСТВО. В силу (32) и (34)

$$\sup_f E_f\,l\left(\frac{\|f-\tilde f_n\|_\rho}{\delta}\right)\geqslant l\left(\frac{c_0}{2}\right)P_e\geqslant$$

$$\geqslant l\left(\frac{c_0}{2}\right)\left(1-n\,\frac{I(X_1,\eta)}{\ln\rho(\delta)}\right). \qquad (35)$$

По определению информации Шеннона

$$I(X_1;\eta)=E\left\{\ln\frac{dP_{X_1|\eta}}{dP_{X_1}}(X_1,\eta)\right\}=$$

$$=E\left\{\ln\frac{dP_{X_1|\eta}}{dx}\right\}-E\left\{\ln\frac{dP_{X_1}}{dx}\right\},$$

где $\frac{dP}{dx}$ обозначает производную меры P по мере Лебега в R^k.

Обозначим P_0 меру с плотностью f_0 в R^k. По неравенству Йенсена

$$E\left\{\ln\frac{dP_{X_1}}{dx}(X_1)\right\}=\int\ln\frac{dP_{X_1}}{dx}(x)\frac{dP_{X_1}}{dx}(x)\,dx\geqslant$$

$$\geqslant\int\ln\frac{dP_0}{dx}(x)\frac{dP_{X_1}}{dx}(x)\,dx.$$

Таким образом

$$I(X_1;\eta)\leqslant E\ln\frac{dP_{X_1|\eta}}{dP_0}(X_1,\eta)=\frac{1}{\rho(\delta)}\sum_{i=1}^{\rho(\delta)}\int_{R^k}\ln\frac{f_{i\delta}(x)}{f_{o\delta}(x)}f_{o\delta}(x)\,dx$$

В области $|x| \le c_1 < 1$ справедливо неравенство

$$\ln(1+x) \le x + \frac{x^2}{2(1-c_1)},$$

и потому

$$I(X_1;\eta) = \frac{1}{\rho(\delta)} \sum_1^{\rho(\delta)} \int_{R^k} \ln\left(1 + \frac{f_{i\delta} - f_{o\delta}}{f_{o\delta}}\right) f_{o\delta}(x)\,dx \le$$

$$\le \frac{1}{\rho(\delta)} \frac{1}{2(1-c_1)} \sum_1^{\rho(\delta)} \left\| \frac{f_{i\delta} - f_{o\delta}}{\sqrt{f_{o\delta}}} \right\|_2^2 \le \frac{c_2}{2n(1-c_1)} \ln\rho(\delta).$$

Возвращаясь к (35), получим

$$\inf_{\tilde f_n} \sup_f E_f\, l\left(\frac{\|f - \tilde f_n\|_p}{\delta(n,\Sigma)}\right) \ge l\left(\frac{1}{2}\right)\left(1 - \frac{c_2}{2(1-c_1)}\right).$$

Теорема доказана.

Разумеется $\delta(n,\Sigma)$ нужно выбирать так, чтобы

$$\frac{c_2}{2(1-c_1)} < 1.$$

Например, выбрать $c_2 = 1 - c_1$.

ЗАМЕЧАНИЕ. Пусть $l \in \Lambda$, так что в условиях теоремы

$$\inf_{\tilde f_n} \sup_f E_f\, l\left(\frac{\|f - \tilde f_n\|}{\delta(n,\Sigma)}\right) \ge l\left(\frac{1}{2}\right)\left(1 - \frac{c_2}{2(1-c_1)}\right).$$

Может статься, что $l\left(\frac{1}{2}\right) = 0$. Пусть $l(\varkappa) > 0$. Т.к. $l_\varkappa(x) = l(2\varkappa x)$ тоже принадлежит Λ, то

$$\inf_{\tilde f_n} \sup_f E_f\, l\left(2\varkappa \frac{\|f - \tilde f_n\|}{\delta(n,\Sigma)}\right) \ge l_\varkappa\left(\frac{1}{2}\right)\left(1 - \frac{c_2}{2(1-c_1)}\right) =$$

$$= l(\varkappa)\left(1 - \frac{c_2}{2(1-c_1)}\right) > 0.$$

5. <u>Примеры</u>. 1. Допустим, что Σ состоит из всех плотностей распределения в R^k принадлежащих $H_p^\beta L$ и ограниченных в L_p сверху константой M. Покажем, что в этом случае для любой функции $l \in \Lambda$ можно найти константу $\gamma > 0$ такую, что

$$\lim_n \sup_{f \in \Sigma} \inf_{\tilde f_n} E_f\, l\left(\gamma n^{\frac{\beta}{k+2\beta}} \|f - \tilde f_n\|_p\right) > 0. \qquad (36)$$

если $1 \leq p < \infty$, и

$$\varliminf_{n} \sup_{f\in\Sigma} \inf_{\tilde f_n} E_f\, l\left(\gamma\left(\frac{n}{\ln n}\right)^{\frac{\beta}{k+2\delta}} \|f-\tilde f_n\|_\infty\right) > 0. \qquad (37)$$

Неравенства (36) и (37) вместе с теоремами 3 и 4 позволяют заключить, что верна следующая

ТЕОРЕМА 9. Пусть Σ состоит из всех плотностей распределения f в R^k , принадлежащих классу $H_p^\beta L$ и ограниченных в L_p константой M . Тогда при $n \to \infty$

$$\sup_{f\in\Sigma} \inf_{\tilde f_n} E_f\left\{\|f-\tilde f_n\|_p^2\right\} \asymp n^{-\frac{2\beta}{2\beta+k}} ,$$

если $2 \leq p < \infty$, и

$$\sup_{f\in\Sigma} \inf_{\tilde f_n} E_f\left\{\|f-\tilde f_n\|_\infty^2\right\} \asymp \left(\frac{n}{\ln n}\right)^{-\frac{2\beta}{2\beta+k}}$$

Докажем неравенства (36) и (37). Они оба выводятся из теоремы 8. Надлежит лишь выбрать по возможности больше функций $f_{i\delta} \in \Sigma$, отстоящих друг от друга не менее чем на δ , и проследить за соблюдением условия 2 теоремы 8 и неравенств: $C_2 < 2(1-C_1)$.

Выберем в Σ какую-нибудь функцию f_0 , которая отлична от нуля на целом интервале. Можно не умаляя общности предположить, что

$$\inf_{|x_i|\leq 1} f_0(x) = q > 0 .$$

Потребуем также, чтобы $f_0 \in H_p^\beta \frac{1}{2} L$ и $\|f_0\|_p \leq \frac{1}{2} M$. Функции $f_{i\delta}$ будут теперь строиться с помощью малых возмущений функции f_{0x} сдвигами и "растяжениями" некоторой фиксированной гладкой функции $\varphi(x)$

Более точно, пусть $\Psi(x)$ – какая-нибудь финитная бесконечно дифференцируемая функция, обращающаяся в нуль вне отрезка $[-\frac{1}{2}, \frac{1}{2}]$ и такая, что

$$\int_{-1/2}^{1/2} \Psi(x)\,dx = 0$$

Например, можно положить

$$\Psi(x) = -\Psi(x)$$
$$\Psi(x) = \begin{cases} 0, & x > \frac{1}{2},\ x = 0, \frac{1}{2} \\ exp\left\{-\frac{1}{x} - \frac{1}{1-2x}\right\}, & 0 < x < \frac{1}{2}. \end{cases}$$

Определим теперь $\varphi(x)$ **следующим образом**

$$\varphi(x_1, \ldots, x_k) = \frac{1}{M_1} \prod_1^k \Psi(x_j),$$

где постоянная M_1 выбрана так, чтобы

$$\|\varphi\|_\infty \le \frac{M}{2}, \quad \varphi \in H_p^\beta \frac{1}{2} L.$$

Задавшись целым числом m, положим

$$\varphi_{mj}(x) = m^{-\beta} \varphi(mx - \bar{j}),$$

где векторы $\bar{j} = (j_1, \ldots, j_k)$ таковы, что их координаты j_s могут принимать значения 1,2, ..., $m-1$. Рассмотрим, наконец, "возмущенные" функции

$$f_a(x) = f_0(x) + \sum_{\bar{j}} a(\bar{j}) \varphi_{m\bar{j}}(x), \tag{37'}$$

где суммирование ведется по всем допустимым векторам $\bar{j} = (j_1, \ldots, j_k)$, $j_s = 1, 2, \ldots, m-1$, а определенная на векторах $\bar{j}$ функция $a(\bar{j})$ принимает значения 0,1. Заметим, что имеется всего $2^{(m-1)^k}$ функций $a(\bar{j})$ указанного вида, и, таким образом, имеется $2^{(m-1)^k}$ различных функций вида (37').

Понятно, что все функции $\varphi_{m\bar{j}}(x)$ отличны от нуля лишь на непересекающихся интервалах вида $\left(\ldots \frac{j_s}{m} \pm \frac{1}{2m} \ldots\right)$ и потому

$$\left\| \sum_{\bar{j}} a(\bar{j}) \varphi_{m\bar{j}}^{(s)} \right\|_p \le m^{-\beta+s} \left(\frac{m-1}{m}\right)^{k/p} \|\varphi\|_p ; \tag{38}$$

$$\left\| \sum_{\bar{j}} a(\bar{j}) \left[\varphi_{m\bar{j}}^{(r)}(x+h) - \varphi_{m\bar{j}}^{(r)}(x)\right] \right\|_p \le \left(\frac{m-1}{m}\right)^{k/p} \sup_{|z| \le h} \left\|\varphi^{(r)}(x+z) - \varphi^{(r)}(x)\right\|_p :$$

Следовательно, для всех достаточно больших m все функции f_a суть плотности распределения из множества $\sum$.

Функции f_a, точнее некоторые из них и будут функциями $f_{i\delta}$ теоремы 8. За функцию $f_{0\delta}$ мы примем функцию f_0. В силу (38)

$$\left\|\frac{f_a-f_0}{f_0}\right\|_\infty \le \frac{1}{q}\left\|\sum_{\bar{j}} a(\bar{j})\varphi_{m\bar{j}}\right\|_\infty \le m^{-\beta} q^{-1}\|\varphi\|_\infty ,$$

так что условие 2) теоремы 8 выполняется с $C_1 = m^{-\beta} q^{-1}\|\varphi\|_\infty$.

Выясним теперь сколько имеется функций f_a отстоящих друг от друга достаточно далеко. Имеем

$$\|f_a - f_{a'}\|_p = \left\|\sum_{\bar{j}} (a(\bar{j}) - a'(\bar{j}))\varphi_{mj}\right\|_p =$$

$$= m^{-\beta-\frac{k}{p}}\left(\sum_{\bar{j}} |(a(\bar{j}) - a'(\bar{j})|\right)^{1/p}\|\varphi\|_p . \tag{39}$$

Для $p=\infty$ это равенство выглядит особенно просто:

$$\|f_a - f_{a'}\|_\infty = m^{-\beta}\|\varphi\|_\infty , \tag{40}$$

и мы сперва рассмотрим этот технически более простой случай, т.е. докажем неравенство (37).

Обозначим D_0 множество тех функций $a(\bar{j})$, которые принимают значение 0 во всех точках $\bar{j}$, кроме одной, а в этой последней принимается значение 1. Плотности f_a, $a\in D_0$ занумеровав их каким-либо образом, мы и примем за функции $f_{i\delta}$ $i\ne 0$, теоремы 8. В силу (40) $\delta = m^{-\beta}\|\varphi\|_\infty$, а $\rho(\delta) = (m-1)^k$. Далее

$$\left\|\frac{f_{i\delta}-f_{0\delta}}{\sqrt{f_{0\delta}}}\right\|_2^2 \le \frac{1}{q}\|f_{i\delta} - f_{0\delta}\|_2^2 = \frac{m^{-k-2\beta}}{q}\|\varphi\|_2^2 . \tag{41}$$

Выберем теперь наименьшее m, для которого

$$\frac{\left\|\frac{f_{i\delta}-f_{0\delta}}{\sqrt{f_{0\delta}}}\right\|_2^2}{\ln\rho(\delta)} \le \frac{m^{-k-2\beta}\|\varphi\|_2^2}{qk\ln(m-1)} \le \frac{1}{2n} ,$$

$$C_1 = m^{-\beta} q^{-1}\|\varphi\|_\infty \le \frac{1}{2} .$$

Ясно, что $m \sim \left(\frac{2k+\beta}{qk} 2\|\varphi\|_2^2\right)^{\frac{1}{2\beta+k}}\left(\frac{n}{\ln n}\right)^{1/(2\beta+k)}$ при $n\to\infty$.

По теореме 8 для любой функции риска $\ell\in\Lambda$

$$\inf_{\tilde{f}_n}\sup_{f\in\Sigma} E_f\,\ell\left(\frac{\|f-\tilde{f}_n\|_\infty}{\delta}\right) = \inf_{\tilde{f}_n}\sup_{f\in\Sigma} E_f\,\ell\left(\frac{\|f-\tilde{f}_n\|_\infty}{m^{-\beta}\|\varphi\|_\infty}\right) \ge \frac{1}{2}\ell\left(\frac{1}{2}\right).$$

Из последнего соотношения легко следует неравенство (37). По существу мы повторили рассуждения из [7], где (37) доказано для случая $k=1$.

Обратимся к доказательству неравенства (36).

ЛЕММА 6. Обозначим через $\mathcal{D}_1$ множество всех функций $a(\bar{j})$ таких, что, если $a, a' \in \mathcal{D}_1$

$$\sum |a(\bar{j}) - a'(\bar{j})| > \frac{(m-1)^k}{4}. \tag{42}$$

Тогда $\operatorname{card} \mathcal{D}_1 \geq \exp\left\{\frac{1}{8}(m-1)^k\right\}$.

Отложив пока доказательство леммы, покажем как с ее помощью получить неравенство (36).

Примем теперь за функции $f_{i\delta}$ теоремы 8 все функции f_{a_i}, $a_i \in \mathcal{D}_1$. В силу (39) и (42)

$$\|f_{i\delta} - f_{j\delta}\|_p \geq m^{-\beta}\left(\frac{m-1}{m}\right)^{k/p} \frac{\|\varphi\|_p}{4^{1/p}},$$

правую часть этого неравенства мы и примем за δ. По лемме 6 $\rho(\delta) \geq \exp\left\{\frac{1}{8}(m-1)^k\right\}$. Наконец,

$$\left\|\frac{f_{i\delta} - f_{o\delta}}{\sqrt{f_{o\delta}}}\right\|_2^2 \leq \frac{1}{q}\|f_{i\delta} - f_{o\delta}\|_2^2 \leq \frac{1}{q}\sum_{\bar{j}} \|\varphi_{m\bar{j}}\|_2^2 \leq$$
$$\leq q^{-1} m^{-2\beta}\left(\frac{m-1}{m}\right)^k \|\varphi\|_2^2.$$

Выберем теперь в качестве m наименьшее целое число, для которого

$$\frac{1}{\ln \rho(\delta)}\left\|\frac{f_{i\delta} - f_{o\delta}}{\sqrt{f_{o\delta}}}\right\| \leq \frac{8}{q}\|\varphi\|_2^2 m^{-2\beta-k} \leq \frac{1}{2n};$$

$$c_1 = m^{-\beta} q^{-1} \|\varphi\|_\infty \leq \frac{1}{2}.$$

Тогда $m \asymp n^{\frac{1}{2\beta+k}}$ и в силу теоремы 8 для любой функции $\ell \in \Lambda$

$$\inf_{\tilde{f}_n} \sup_f E_f \ell\left(\frac{\|f - \tilde{f}_n\|_p}{\delta}\right) = \inf_{\tilde{f}_n} \sup_f E_f \ell\left(\frac{\|f - \tilde{f}_n\|_p}{m^{-\beta}}\left(\frac{m}{m-1}\right)^{k/p} \frac{4^{1/p}}{\|\varphi\|_p}\right) \geq$$
$$\geq \frac{1}{2}\ell(1/2),$$

что эквивалентно (36).

Осталось доказать лемму 6. Ее можно вывести из следующего результата, так называемой границы Варшамова – Гилберта из теории кодирования.

ЛЕММА 7. (граница Варшамова – Гилберта). Рассмотрим функции $a(w)$ определенные на конечном множестве $\Omega=\{w\}$, $card\,\Omega=N$ и принимающие значения ± 1. Обозначим $\mathcal{D}$ множество тех функций $a(w)$, что для $a, a'\in\mathcal{D}$

$$\sum|a(w)-a(w')|\geq d.$$

Тогда

$$card(\mathcal{D})\geq\left[2^{-N}\sum_{i=0}^{d-1}C_N^i\right]^{-1}. \tag{43}$$

Доказательство этой леммы можно найти в [3], стр.546 и 620. В силу экспоненциальных оценок для вероятностей больших уклонений (см., напр., неравенство Хёфдинга в [6], стр.76)

$$2^{-N}\sum_{i=0}^{[N/4]}C_N^i\leq exp\left\{-\frac{N}{8}\right\}. \tag{44}$$

Из неравенств (43) и (44) с очевидностью следует утверждение леммы 6. Итак, оба неравенства (36) и (37), а с ними и теорема 9 доказаны.

ЗАМЕЧАНИЕ. На самом деле мы доказали несколько больше, чем (36) и (37). Именно, пусть множество Σ такое же, что и в теореме 9, пусть плотность f_0 отлична от нуля в некотором k -мерном интервале и является "внутренней" точкой Σ, т.е.
$f_0\in H_p^\beta(L-\lambda)$, $\|f_0\|_p\leq M-\lambda$ для какой-нибудь $\lambda>0$
Пусть U_ε есть пересечение L_p – шара радиуса C с центром f_0 и множества Σ. Для любой функции потерь $l\in\Lambda$ найдется число $\gamma>0$ такое, что, если, например, $1\leq p<\infty$

$$\lim_{\delta\to 0}\lim_{n\to\infty}\inf_{\tilde f_n}\sup_{f\in U_\delta}E_f l\left(\gamma_n^{\frac{\beta}{2\beta+k}}\|f-\tilde f_n\|_p\right)>0.$$

2. ТЕОРЕМА 10. Пусть Σ есть множество всех плотностей распределения в R^1 принадлежащих классу $A_p^\lambda L$ (см. Пример 2 п.3). Тогда при $n\to\infty$

$$\inf_{\tilde f_n}\sup_{f\in\Sigma}E_f\|f-\tilde f_n\|_p^r\asymp\left(\frac{\ln n}{n}\right)^{r/2},\quad 2\leq p\leq\infty;$$

$$\inf_{\tilde f_n} \sup_{f\in\Sigma} E_f \|f-\tilde f_n\|_\infty^2 \asymp \left(\frac{\ln n \ln\ln n}{n}\right)^{2/2}. \tag{45}$$

В силу теоремы 6 достаточно оценить левые части (45) снизу. Подобно предыдущему примеру эта оценка выводится из теоремы 8. Рассмотрим только более сложный случай $p<\infty$ и, поскольку ход рассуждений тот же, что и в предыдущем примере, мы лишь наметим доказательство. При больших τ плотность $f_0(x)=\frac{c_\tau}{\tau^2+x^2}$ заведомо принадлежит $A_p^\lambda L$. Положим

$$\varphi(x)=\frac{\sin^5 x}{x^4}, \quad \varphi_{mj}(x)=\varphi(mx-j)e^{-m\lambda_0},$$

где $\lambda_0\geqslant 5\lambda$ выбрано так, чтобы $\varphi_{mj}\in A_p^\lambda \frac{1}{2}L$. Семейство $\{f_{i\delta}\}$ состоит из функций

$$f_a=f_0+\sum_{j=1}^{m-1} a(j)\varphi_{mj}(x),$$

где функция $a(j)$ по-прежнему принимает значения 0,1. В качестве $f_{0\delta}$ мы принимаем функцию f_0. С помощью достаточно простых выкладок проверяется, что

$$\|f_a-f_{a'}\|_p=\Big\|\sum_{j=1}^{m-1}(a(j)-a'(j))\varphi_{mj}\Big\|_p\geqslant$$

$$\geqslant e^{-m\lambda_0}\varkappa_0\Big(\sum|a(j)-a'(j)|\Big)^{1/p} m^{-1/p}, \quad \varkappa_0>0.$$

По лемме 6 можно выделить множество $\mathcal{D}_1$ плотностей f_a, мы их запишем в виде $f_{i\delta}$, такое, что $\rho(\delta)=\operatorname{card}\mathcal{D}_1> e^{\frac{m-1}{8}}$ и

$$\|f_{i\delta}-f_{j\delta}\|\geqslant\varkappa_1 e^{-m\lambda_0}=\delta, \quad \varkappa_1>0. \tag{46}$$

Далее, при больших m

$$c_1=\Big\|\frac{f_{i\delta}}{g}-1\Big\|_\infty<\frac{1}{2}. \tag{47}$$

Наконец,

$$\Big\|\frac{f_{i\delta}-f_{0\delta}}{\sqrt{f_{0\delta}}}\Big\|_2^2\leqslant\varkappa_2 e^{-2\lambda_0 m}, \quad \varkappa_2>0 \tag{48}$$

Выберем минимальное m для которого выполняется (47) и

$$\frac{\varkappa_2 e^{-2\lambda_0 m}}{\ln \rho(\delta)} = \frac{\varkappa_2 e^{-2\lambda_0 m}}{m-1} \leq \frac{1}{2n}$$

При $n \to \infty$

$$m \sim \frac{\ln n}{2\lambda_0} - \ln\ln n,$$

$$\delta = \varkappa_1 e^{-m\lambda_0} \sim \sqrt{\frac{\ln n}{n}}$$

По теореме 8 для любой функции риска $l \in \Lambda$ найдется $\gamma > 0$, что

$$\varliminf_{n} \inf_{\tilde{f}_n} \sup_{f \in \Sigma} E_f l\left(\gamma \sqrt{\frac{n}{\ln n}} \| f - \tilde{f}_n \|_p\right) > 0$$

3. Выше, исследуя скорость сходимости δ_γ^* к f в метрике L_p, мы предполагали, что $p \geq 2$. Покажем, что это ограничение диктуется существом дела. Для простоты мы рассмотрим только тот случай, когда $\Sigma \subset H_1^\beta L$. Мы докажем, что каковы-бы ни были оценки $\tilde{f}_n$ плотности f, верхняя грань $\sup_f E_f \| f - \tilde{f}_n \|_1$ даже не стремится к нулю при $n \to \infty$. Доказательство по-прежнему основано на теореме 8.

Определим финитные бесконечно дифференцируемые функции $g_n(x)$ следующим образом

$$g_n(x) = \begin{cases} 1, & |x| \leq 1 \\ \frac{1}{|x|}, & x \in [-2^n, -2] \cup [2, 2^n] \\ 0, & |x| > 2^n + 1 . \end{cases}$$

В интервалах $[-2^n-1]$, $[2^n, 2^n+1]$ функции g_n удовлетворяют неравенствам $0 < g_n' < 1/|x|$. Определим плотности распределения

$$f_{n_0}(x) = \frac{C_n}{n} g_n(x)',$$

где C_n - нормирующие постоянные. Очевидно, $C_n \asymp 1$, когда $n \to \infty$.

Пусть далее $\varphi(x)$ **бесконечно** дифференцируемая функция такая, что

$$0 \leqslant \varphi \leqslant 1, \quad \varphi(x)=1, \ x \in \left[\frac{4}{3}, \frac{5}{3}\right], \ \varphi(x)=0, \quad x \bar{\in} [1,2].$$

Положим

$$\varphi_{mn}(x) = \begin{cases} \varphi(x 2^{-m}) f_{no}(x), & x>0 \\ -\varphi_{mn}(-x), & x<0 \end{cases}.$$

Рассмотрим теперь семейство плотностей

$$f_{an}(x) = f_{n_0} + \gamma \sum_{m=1}^{n-1} a(m) \varphi_{mn}(x),$$

где как и раньше функции $a(m)$ принимают значения 0,1, а $\gamma < 1$ – малое положительное число, выбор которого мы уточним позднее.

Имеем,

$$\left\| \frac{f_{n_0} - f_{an}}{f_{n_0}} \right\|_\infty \leqslant \gamma < 1 .$$

Как уже отмечалось, можно выбрать множество $\mathcal{D}_1$ функций a так, что $\operatorname{card} \mathcal{D}_1 > e^{(n-1)/8}$ и для $a, a' \in \mathcal{D}_1$

$$\|f_{an} - f_{a'n}\| = \gamma \sum_{m=1}^{n-1} |a(m) - a'(m)| \, \|\varphi_{mn}\|_1 \geqslant$$

$$\geqslant \frac{\gamma}{n} \ln \frac{5}{4} \sum_{m=1}^{n-1} |a(m) - a'(m)| > \frac{\gamma}{32} .$$

Плотности f_{an} с $a \in \mathcal{D}_1$ мы примем за семейство $f_{i\delta}$ теоремы 8 с $\delta = \frac{\gamma}{32}$, $\rho(\delta) > \exp\left\{\frac{n-1}{8}\right\}$ (и $f_{o\delta} = f_{no}$).

Для этого семейства

$$\left\| \frac{f_{o\delta} - f_{i\delta}}{\sqrt{f_{i\delta}}} \right\|_2 \leqslant 2\gamma^2 \sum_{m=1}^{n-1} a^2(m) \int_{2^m}^{2^{m+1}} f_{no}(x)\, dx \leqslant \gamma^2 .$$

Положим $\gamma = \frac{1}{4}$. Для $n \geqslant 2$

$$\frac{1}{\rho(\delta)} \left\| \frac{f_{o\delta} - f_{i\delta}}{\sqrt{f_{i\delta}}} \right\|_2^2 \leqslant \frac{\gamma^2}{2(n-1)} \leqslant \frac{3}{4n} .$$

По теореме 8

$$\inf_{f_n} \sup_{f_{i\delta}} E_f \|f_{i\delta} - \tilde{f}_n\|_1 \geqslant 2^{-9} .$$

Очевидно, наконец, что плотности f_{n0} можно определить так, что для всех n и всех β плотности $f_{n\alpha} \in H_1^\beta L$, где L зависит разве лишь от β, но не от n. Следовательно, мы получили такой результат:

ТЕОРЕМА 11. Пусть Σ состоит из всех плотностей распределения в R^1, принадлежащих $H_1^\beta L$. Тогда

$$\inf_{n\geq 1} \inf_{\tilde{f}_n} \sup_{f\in\Sigma} E_f \|f-\tilde{f}_n\|_1 > 0.$$

По-видимому, подобный результат верен и для семейств аналитических плотностей.

Литература

1. А х и е з е р Н.И. Лекции по теории аппроксимации, Наука, М., 1965.
2. B r e t a g n o l l e J., H u b e r C. Estimation de densites: risque minimax, - Z. Wahrscheinlichkeitstheorie verw. Gebiete, v.47, 1979, p.119-137.
3. Г а л л а г е р Р. Теория информации и надежная связь, Сов. Радио, М., 1974.
4. И б р а г и м о в И.А., Х а с ь м и н с к и й Р.З. Асимптотическая теория оценивания, Наука, М., 1979.
5. Н и к о л ь с к и й С.М. Приближение функций многих переменных и теоремы вложения, Наука, М., 1969.
6. П е т р о в В.В. Суммы независимых случайных величин, Наука, М., 1972.
7. Х а с ь м и н с к и й Р.З. О границе снизу рисков непараметрических оценок плотности в равномерной метрике, - Теория вероятн. и ее примен., 1978, т.23, № 4, с.824-828.
8. Ч е н ц о в Н.Н. Оценка неизвестной плотности по наблюдениям, Докл. АН СССР, 1962, т.147, № 1, с.45-48.
9. Ч е н ц о в Н.Н. Статистические решающие правила и оптимальные выводы, Наука, М., 1972.