УДК 519.24

Ю.А.Кошевник
(Московский университет)

ОБ АСИМПТОТИЧЕСКОМ РАСПРЕДЕЛЕНИИ НЕПАРАМЕТРИЧЕСКИХ ОЦЕНОК ФУНКЦИИ РАСПРЕДЕЛЕНИЯ ПРИ УСЛОВИИ СИММЕТРИИ

Рассмотрим последовательность независимых одинаково распределенных случайных величин $X_1, \ldots, X_n$, каждая из которых принимает значения в k -мерном евклидовом пространстве $\mathfrak{R}^k$ (k -мерная выборка). Распределение F величины X_i предполагается неизвестным, но удовлетворяющим условию симметрии и некоторым условиям гладкости (существование плотности, в некоторых случаях непрерывной дифференцируемости плотности и т.п.). В случае одномерной выборки условие симметрии имеет вид $F(A) = F(-A)$ или $F(c+A) = F(c-A)$ для всех измеримых множеств A и для некоторого известного или неизвестного значения параметра c. Для k -мерной выборки ($k > 1$) условия симметрии могут варьироваться в более широких пределах (все отражения или некоторые отражения относительно координатных плоскостей или плоскостей, параллельных координатным и т.п.). При выполнении этих условий эмпирическая функция распределения может быть улучшена путем симметризации. Представляет интерес асимптотическое распределение нормированного уклонения $\sqrt{n}\,(\widetilde{F}_n - F)$ улучшенной оценки от теоретического распределения и $\sqrt{n}\,(F_n - \widetilde{F}_n)$ – от эмпирического распределения. При этом уклонения и нормированные уклонения, а также их пределы по распределению рассматриваются как элементы некоторого пространства Д функций на $\mathfrak{R}^k$, несепарабельного, но банахова. Сходимость мер на таких пространствах была определена и изучена Дадли [1; 2]. Условия симметрии являются как бы предельным случаем оценивания функции распределения при наличии ограничений, когда число ограничений стремится к бесконечности. Тем не менее, в известном смысле, положение аналогично случаю конечного числа ограничений [3; 4] . В статье излагается бескоординатное описание условий регулярности семейства вероятностных мер, перекликающееся с [5; 6] и охватывающее как параметрическое оценивание, так и непараметрическое оценивание.

§ I. Регулярное семейство вероятностных мер

Пусть $\mathcal{F}$ -семейство вероятностных мер на k -мерном евклидовом пространстве $\mathfrak{R}^k$, абсолютно непрерывных относительно меры Лебега λ ($\mathcal{F}$ может содержать и не все абсолютно непрерывные меры). Определим некоторые геометрические понятия, позволяющие обобщить понятие оценки максимального правдоподобия на непараметрический случай.

Для вероятностной меры $F \in \mathcal{F}$ ее плотность $\frac{dF}{d\lambda} = f$, $f \in L^1(\mathfrak{R}^k, \lambda) = L^1(\lambda)$; вместо f рассмотрим $\sqrt{f} \in L^2(\lambda)$.

Определение I.I. Отображение γ из интервала $(-\delta, \delta)$ в $\mathcal{F}$ называется дифференцируемым в точке $h = 0$, если путь $\tilde{\gamma}$ в $L^2(\lambda)$, задаваемый формулой $\tilde{\gamma}(h)(\cdot)\sqrt{\frac{d\gamma(h)}{d\lambda}}(\cdot)$, дифференцируем (в $L^2(\lambda)$) в точке $h = 0$; дифференцируемость рассматривается по Фреме.

Если $\xi^* \in L^2(\lambda)$ является производной пути $\tilde{\gamma}$, то мы будем называть ξ также производной пути γ ; заметим, что пространство $L^2(F)$ ($F \in \mathcal{F}$) удобно считать вложенным в $L^2(\lambda)$, отождествляя $\varphi \in L^2(\mathcal{F})$ с $\varphi\sqrt{f} \in L^2(\lambda)$. Так как отображение $j_F : \varphi \mapsto \varphi\sqrt{f}$ является изометрическим изоморфизмом, то производные некоторых дифференцируемых путей могут рассматриваться и как элементы $L^2(F)$.

Определение I.2. Отображение $\gamma : (-\delta, \delta) \to \mathcal{F}, \gamma(0) = F$, дифференцируемое в точке $h = 0$, называется абсолютно непрерывно дифференцируемым (а.н.д.) путем, если производная γ в точке 0 имеет вид $\xi(\cdot) = \xi_1(\cdot)\sqrt{f(\cdot)}$, $\xi_1 \in L^2(F)$, $f = \frac{dF}{d\lambda}$.

Определение I.3. Для $F \in \mathcal{F}$ рассмотрим подпространство в $L^2(F)$, порожденное производными в точке $h = 0$ а.н.д. путей γ, таких, что $\gamma(0) = F$. Это подпространство мы обозначим $T_F\mathcal{F}$ и назовем касательным пространством к $\mathcal{F}$ в точке F.

Лемма I.I. Функция I (тождественно равная I) ортогональна $T_F\mathcal{F}$ в $L^2(F)$.

Доказательство. Пусть γ - а.н.д. путь в $\mathcal{F}$, $\gamma(0) = F$; $\tilde{\gamma}$ - путь в $L^2(\lambda)$, $\tilde{\gamma}(0) = \sqrt{f}$; $\int[\tilde{\gamma}(h)(x)]^2\, \lambda(dx) = 1$. Дифференцируя по h и полагая $h = 0$, получим $\int \tilde{\gamma}(0)(x)\tilde{\gamma}'(0)(x)\lambda(dx) = \int \sqrt{f}\,\xi_1\sqrt{f}\,d\lambda = 0$. Поэтому $\int \xi_1(x)\sqrt{f(x)}\cdot 1\cdot\sqrt{f(x)}\,\lambda(dx) = 0$, т.е.

$$\int \xi_1(x) f(x)\lambda(dx) = \int \xi_1(x) F(dx) = 0.$$

Обозначим $\hat{L}^2(F)$ подпространство в $L^2(F)$, ортогональное I, т.е. $\varphi \in \hat{L}^2(F)$ тогда и только тогда, когда $\int \varphi dF = 0$.

Пример I.I. Пусть $\mathcal{F}$ – семейство всех абсолютно непрерывных по λ вероятностных мер. Тогда $T_F\mathcal{F} = \hat{L}^2(F)$.

Пример I.2. Пусть $\mathcal{F}$ –конечномерное семейство мер, $F_\theta = F(\cdot;\theta)$, с плотностями $f(\cdot;\theta)$, где θ пробегает область $\Theta \subset \mathcal{R}^p$. Пусть $O \in \Theta$. Если функция f достаточно гладкая по θ, то **пути, направленные из нуля по координатным осям, будут а.н.д. путями**, а $T_F\mathcal{F}$ является p -мерным пространством, натянутым на функции

$$\xi_i(x) = \left[\left(\frac{\partial}{\partial\theta_i}\sqrt{f(x;\theta)}\right)/\sqrt{f(x;\theta)}\right]\Big|_{\theta=0} = \frac{1}{2}\,\frac{\partial f(x;\theta)}{\partial\theta_i}\cdot\frac{1}{f(x;\theta)}\Big|_{\theta=0}.$$

Рассмотрим вместе с семейством $\mathcal{F}$ семейство касательных пространств $\{T_F\mathcal{F}\ ;\ F\in\mathcal{F}\}$. Пусть $\mathcal{T}_F$ –ортогональный проект из $L^2(\lambda)$ на $T_F\mathcal{F}$, причем $\mathcal{T}_F$ – оператор, действующий в $L^2(\lambda)$. Можно рассматривать различные топологии в пространстве линейных операторов $\mathcal{B}(L^2(\lambda))$: $\mathcal{N}$ – топология, порождаемая нормой оператора, для оператора ρ $\|\rho\| = \sup\limits_{g\in L^2(\lambda)\setminus\{0\}}(\|\rho(g)\|/\|g\|)$; и S – топология сходимости поточечной, порождаемой семейством полунорм

$$\{p_g(\rho) = \|\rho(g)\|,\ g\in L^2(\lambda)\}.$$

О п р е д е л е н и е I.4. Семейство $\mathcal{F}$ называется S - (соответственно, $\mathcal{N}$) регулярным, если операторнозначная функция $\mathcal{T}: \mathcal{F}\to\mathcal{B}(L^2(\lambda)), \mathcal{T}(F) = \mathcal{T}_F$, является непрерывной (в $\mathcal{F}$ – топология сходимости по вариации) в S – (соответственно, в $\mathcal{N}$) топологии.

Пример I.3. Для конечномерного семейства $\mathcal{F}$ из примера I.2 рассмотрим ортопроекцию из $L^2(\lambda)$ на $T_F\mathcal{F}$. Если $\varphi\in L^2(\lambda)$, а $\mathcal{T}_F\varphi = (\alpha_1\xi_1 + \ldots + \alpha_p\xi_p)\sqrt{f}$, то коэффициенты $\alpha_1,\ldots,\alpha_p$ определяются как решение системы линейных уравнений

$$\sum_{i=1}^{p}\alpha_i\int\xi_i(x)\xi_j(x)F(dx) = \int\varphi(x)\xi_j(x)\sqrt{f(x)}\,\lambda(dx);\quad j=1,\ldots,p.$$

Если потребовать, чтобы матрица $I(F) = (\beta_{ij}), \beta_{ij} = 4\int\xi_i\xi_j\,dF$ была невырожденной, а ее элементы были бы непрерывными функциями θ, то и $\alpha_1,\ldots,\alpha_p$ будут непрерывно зависеть от θ, т.е. от распределения $F\in\mathcal{F}$.

Опишем теперь проекцию из $L^2(\lambda)$ на $L^2(F)$. Эта проекция ставит в соответствие функции $\varphi \in L^2(\lambda)$ функцию $\xi\sqrt{f} \in L^2(\lambda)$, причем $\varphi - \xi\sqrt{f}$ ортогонально всем $\eta\sqrt{f}$ в $L^2(\lambda)$, $\eta \in L^2(F)$.

Л е м м а I.2. Если $\varphi \in L^2(\lambda)$, $\xi \in L^2(F)$, $\zeta \in L^2(\lambda)$,

$$\xi \perp L^2(F), \quad \varphi = \xi\sqrt{f} + \zeta,$$

то для меры ν_{φ^2}, такой, что $d\nu_{\varphi^2}/d\lambda = \varphi^2$ разложение на абсолютно непрерывную относительно F и сингулярную относительно F компоненты имеет вид

$\nu_{\varphi^2} = \nu_{\xi^2\varphi} + \nu_{\zeta^2}$, а для плотностей соответственно –

$\varphi^2 = \xi^2 f + \zeta^2$.

Доказательство следует из соотношения

$$\int_A \varphi^2 d\lambda = \int_A \xi^2 f d\lambda + \int_A \zeta^2 d\lambda + 2\int \xi \mathbb{1}_A \sqrt{f}\, \zeta\, d\lambda,$$

где $\mathbb{1}_A(x)$ – характеристическая функция множеств. Первый интеграл определяет меру, абсолютно непрерывную относительно F : $d\nu_{\xi^2 f}/dF = \xi^2 f/f = \xi^2$. Третий интеграл равен нулю, так как $\xi \in L^2(F)$, $\eta = \xi \mathbb{1}_A \in L^2(F)$, а ζ ортогональна всем функциям вида $\eta\sqrt{f}$; второй интеграл определяет некоторую меру. Чтобы показать ее сингулярность с F, заметим, что если на множестве A f всюду отлична от 0, то функция $\zeta \mathbb{1}_A$ может быть представлена в виде

$$\zeta \cdot \mathbb{1}_A = \frac{\zeta \cdot \mathbb{1}_A}{\sqrt{f}} \cdot \sqrt{f} = \eta\sqrt{f}, \; \eta \in L^2(F) \quad ; \text{ поэтому } \int_A \zeta^2 d\lambda = \int \eta\sqrt{f}\,\zeta\, d\lambda = 0.$$

Пусть теперь $F_n \in \mathcal{F}$, f_n – плотность F_n; $F_n \to F$, f – плотность F. Заметим, что в силу положительности функций f_n и f сходимость $f_n \to f$ в $L^1(\lambda)$ эквивалентна сходимости $\sqrt{f_n} \to \sqrt{f}$ в $L^2(\lambda)$.

Нетрудно видеть, что семейство $\mathcal{F}_0$ всех вероятностных мер, взаимно абсолютно непрерывных относительно меры Лебега в $\mathcal{R}^k$, является $\mathcal{N}$ – регулярным. В дальнейшем, говоря о регулярных семействах вероятностных мер, будем иметь всегда в виду подсемейства $\mathcal{F}_0$.

§ 2. Оценивание функции распределения. Обобщение метода максимума правдоподобия

Пусть производятся независимые наблюдения $X_1, \dots, X_n$ случайной величины X, принимающей значения в $\mathcal{R}^k$ и подчиняющейся распределению $F \in \mathcal{F}$. Оценкой F мы назовем измеримую функцию наблюдений $X_1,\dots,X_n$ со значениями в пространстве всех линейных функционалов на $L^2(\lambda) = H$, т.е. оценка F — это функция $F_n[\varphi; X_1,\dots,X_n]$, измеримая по $X_1,\dots,X_n$ и линейная по первому аргументу $(\varphi \in L^2(\lambda))$.

Определение 2.1. Стохастической мерой на $(\mathcal{R}^k, \mathcal{R}^k)$ называется отображение $W: \mathcal{R}^k \to L^2(\Omega, \mathcal{B}, P)$ σ - алгебры борелевских подмножеств $\mathcal{R}^k$ в пространство квадратично интегрируемых случайных величин на вероятностном пространстве $(\Omega, \mathcal{B}, P)$, являющемся мерой на $\mathcal{R}^k$ со значениями в $L^2(\Omega, \mathcal{B}, P)$. Стохастическая мера W называется ортогональной, если случайные величины $W(A)$ и $W(B)$ ортогональны всякий раз, когда $A \cap B = \varnothing$; W называется m - ортогональной, если существует конечная положительная мера m на $(\mathcal{R}^k, \mathcal{R}^k)$, такая, что

$$E_p[W(A)W(B)] = m(A \cap B)$$

(E_p - математическое ожидание на Ω по мере p).

Стохастическая мера W называется среднеквадратично непрерывной относительно конечной положительной меры m, если линейное отображение $\mathcal{J}: L^2(m) \to L^2(\Omega, \mathcal{B}, P)$, определяемое для конечных линейных комбинаций $\varphi = \sum_{i=1}^{r} c_i \mathbb{1}_{A_i}$ по формуле $\mathcal{J}(\varphi) = \sum_{i=1}^{r} c_i W(A_i)$ (нетрудно видеть, что определение корректно: если $\varphi = \sum_{i=1}^{r} c_i \mathbb{1}_{A_i} = \sum_{j=1}^{s} \alpha_j \mathbb{1}_{B_j}$, то $\sum_{i=1}^{r} c_i W(A_i) = \sum_{j=1}^{s} \alpha_j W(B_j)$, так как W - мера) может быть продолжено до непрерывного линейного оператора $\tilde{\mathcal{J}}: L^2(m) \to L^2(\Omega, \mathcal{B}, P)$. Это продолжение $\tilde{\mathcal{J}}(\varphi)$ мы обозначим как $\int \varphi\, dW$ - стохастический интеграл от (неслучайной) функции φ по мере W.

Введем некоторые операции над стохастическими мерами. Если $\Lambda: L^2(F) \to L^2(F)$ - непрерывный линейный оператор, а W - стохастическая мера, непрерывная в среднеквадратичном относительно вероятностной меры F, то определим по W новую стохастическую меру, которую назовем суперпозицией W и Λ:

$$W(A) = (W_0 \Lambda(A) = \int \Lambda(\mathbb{1}_A)\, dW.$$

В частности, когда Λ -ортопроектор на замкнутое линейное подпространство $\mathcal{N} \subset L^2(F)$, то V мы будем называть проекцией W на $\mathcal{N}$.

О п р е д е л е н и е 2.2. Стохастическая мера W называется гауссовской, если для любых $A_1, \ldots, A_n$ совместное распределение $(W(A_1), \ldots, W(A_n))$ - гауссовское.

О п р е д е л е н и е 2.3. Неопределенным интегралом от стохастической меры W называется случайная функция

$$w(x) = \int \mathbb{1}_{<x}(y)\, W(dy) \quad , \text{ где}$$

$$\mathbb{1}_{<x}(y) = \begin{cases} 1, \text{ если все } & y_i < x_i, \\ 0 \text{ в противном случае.} \end{cases}$$

О п р е д е л е н и е 2.4. Пусть для $x^0 \in \mathcal{R}^k$ $A_{<x^0}$ $(A_{\leq x^0})$ обозначает подмножество в $\mathcal{R}^k$, состоящее из таких y , что $y_j < x^0_j$ $(y_j \leq x^0_j)$, $A_{>x^0} = \mathcal{R}^k \setminus A_{\leq x^0}$, $A_{\geqslant x^0} = \mathcal{R}^k \setminus A_{<x^0}$.
Пространство Д состоит из вещественных ограниченных функций f: $\mathcal{R}^k \to \mathcal{R}$, таких, что

1) $\forall_j$ существует $f_j(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k) =$

$$= \lim_{x_j \to +\infty} f(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_k)$$

и

$$\lim_{x_j \to -\infty} f(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_k) = 0 ;$$

2) для всякого $x^0 \in \mathcal{R}^k$ существуют пределы

$$\lim_{\substack{x \to x^0 \\ x \in A_{<x^0}}} f(x) = \lim_{\substack{x \to x^0 \\ x \in A_{\leq x^0}}} f(x) = f(x^0)$$

и $\lim\limits_{\substack{x \to x^0 \\ x \in A_{\geqslant x^0}}} f(x)$;

3) среди множества плоскостей в $\mathcal{R}^k$ вида $\{x_j = x^0_j\}$ лишь при переходе через не более чем счетное число плоскостей функция f терпит разрыв.

В пространстве Д мы введем $\sup$ -норму $\|f\| = \sup\limits_{x \in \mathcal{R}^k} |f(x)|$.
Траектории неопределенного интеграла w от стохастической меры W принадлежат пространству Д , но отображение F_n : $((\mathcal{R}^k)^n, (\mathcal{R}^k)) \to$ Д, ставящее в соответствие выборке $X_1, \ldots, X_n$ неопределенный интеграл от эмпирической меры (приписывающей вес $1/n$ точкам $X_1, \ldots, X_n$ и только им), не будет измеримым отображением, если в Д рассматривать σ - алгебру $\mathcal{D}$ всех борелевских множеств. Дадли [1, 2] предложил вместо $\mathcal{D}$ рассматривать

σ - алгебру $\mathcal{D}_0$, порожденную шарами в Д. Так как Д несепарабельно, то $\mathcal{D}_0$ существенно меньше $\mathcal{D}$. Следующее определение сходимости стохастических мер по распределению принадлежит Дадли [I].

Определение 2.5. Пусть W, W_n - последовательность среднеквадратичнонепрерывных относительно F стохастических мер; w, w_n - их неопределенные интегралы; P_n - вероятностные меры на измеримом пространстве*(Д, $\mathcal{D}_0$), соответствующие w_n; P - борелевская мера на (Д, $\mathcal{D}$), соответствующая w (т.е. P сосредоточена на сепарабельном подпространстве в Д). Мы скажем, что последовательность мер P_n слабо сходится к P (соответственно $w_n \to w$ и $W_n \to W$ по распределению), если для всякой непрерывной ограниченной функции $g: Д \to \mathfrak{R}$

$$\int^* g\,dP_n \to \int g\,dP, \quad \int_* g\,dP_n \to \int g\,dP, \text{ где}$$

$$\int^* g\,dP_n = \inf\{\int g_0\,dP_n, \quad g_0 \text{ измерима относительно } Д_0, g_0 \geq g\},$$

$$\int_* g\,dP_n = \sup\{\int g_0\,dP_n, \quad g_0 \text{ измерима относительно } Д_0, g_0 \leq g\}.$$

В частности [I, 2], если F_n - эмпирическая вероятностная мера, $W_n = \sqrt{n}\,(F_n - F)$, то W_n сходится по распределению к гауссовской стохастической мере с нулевым средним и ковариациями $E_p[W(A)W(B)] = F(A \cap B) - F(A)F(B)$. Это мера P к тому же сосредоточена на непрерывных функциях, если F непрерывна.

Обычными рассуждениями (помимо [I], см. также [8], гл. I) можно показать, что вместо непрерывной ограниченной функции g: Д $\to \mathfrak{R}$ можно брать ограниченную функцию, множество точек разрыва которой имеет P -меру нуль.

§ 3. Симметричные распределения в $\mathfrak{R}^k$

Пусть $\mathcal{G}_c$ - группа некоторых движений $\mathfrak{R}^k$, оставляющая неподвижной точку $c \in \mathfrak{R}^k$; $\mathcal{F}_c$ состоит из всех таких $F \in \mathcal{F}$, что для всякого измеримого множества $A \in \mathfrak{R}^k$ и для всякого преобразования $g \in \mathcal{G}_c$ $F[g(A)] = F[A]$. Плотность $f(x) = dF/dx\,(x)$

обладает свойством инвариантности в следующей форме: $f(g(x))=f(x)$ для всех $g\in\mathcal{G}_c$ и для почти всех $x\in\mathcal{R}^k$.

Пусть $\mathcal{F}^{(s)}=\bigcup_{c\in\mathcal{R}^k}\mathcal{F}_c$; выясним строение касательного пространства $T_F\mathcal{F}^{(s)}$ к $\mathcal{F}^{(s)}$ в точке F.

Л е м м а 3.I. Пусть H – гильбертово пространство; $\mathcal{G}$ – некоторая группа унитарных операторов в H, конечная или компактная; Δ – инвариантная мера Хаара на $\mathcal{G}$; $H_1=\{\xi\in H:\ g(\xi)=\xi$ для всех $g\in\mathcal{G}\}$;

$$A_\Delta : H\to H\ ;\ A_\Delta\xi=\int_{\mathcal{G}} g(\xi)\Delta(dg)\ ;$$
$$H_2=\{\xi\in H\ ;\ A_\Delta\xi=0\}.$$

Тогда H разлагается в прямую сумму ортогональных подпространств $H=H_1+H_2$, а A_Δ совпадает с ортопроектором на H_1.

Д о к а з а т е л ь с т в о. $A_\Delta^2\xi=A_\Delta(A_\Delta\xi)=A_\Delta\xi$; A_Δ – самосопряженный, так как $g\in\mathcal{G}$ – унитарные операторы; A_Δ действует на H_1 тождественным образом, а H_2 переводит в нуль. Если $\xi\in H$, то $\xi-A_\Delta\xi\in H_2$.

Л е м м а 3.2. Пространства $L^2(\lambda)$ и $L^2(F)$ могут быть представлены в виде прямой суммы подпространства $\mathcal{G}$ – инвариантных функций ($\varphi(g(x))=\varphi(x)$ для всех $g\in\mathcal{G}$, для λ – почти всех $x\in\mathcal{R}^k$) и подпространства антиинвариантных функций ($\int_G\varphi(g(x))\Delta(dg)=0$ для λ – почти всех x.

Доказательство непосредственно следует из леммы 3.I, поскольку из $\mathcal{G}$ – инвариантности мер F и λ вытекает $\mathcal{G}$ – инвариантность плотности $f=dF/d\lambda$; а также унитарность операторов g^*, $(g^*\varphi)(x)=\varphi(g(x))$, как в $L^2(\lambda)$, так и в $L^2(F)$.

Пример 3.I. Пусть группа $\mathcal{G}$ в $\mathcal{R}^k$ состоит из отражений относительно всех координатных плоскостей. Тогда проекция на подпространство инвариантных функций имеет вид $(A_\Delta\varphi)(x)=\frac{1}{2^k}\sum\varphi(\varepsilon x)$, где $\varepsilon=(\varepsilon_1,\dots,\varepsilon_k)$ – набор величин, принимающих значения ±1, $\sum$ берется по всем ε, $\varepsilon x=(\varepsilon_1x_1,\varepsilon_2x_2,\dots,\varepsilon_kx_k)$.

При $k=1$ $(A_\Delta\varphi)(x)=\frac{1}{2}[\varphi(x)+\varphi(-x)]$.

При $k=2$, если группа состоит из отражений относительно осей $x_1+x_2=0$ и $x_1-x_2=0$, то

$$(A_\Delta\varphi)(x)=\frac{1}{4}\left[\varphi(x_1,x_2)+\varphi(x_2,x_1)+\varphi(-x_1,-x_2)+\varphi(-x_2,-x_1)\right],$$

а если $\mathcal{G}$ содержит кроме тождественного отображения только одно отражение относительно оси $x_1 - x_2 = 0$, то

$$(A_\Delta \varphi)(x) = \frac{1}{2}\left[\varphi(x_1, x_2) + \varphi(x_2, x_1)\right].$$

§ 4. Обобщение оценок максимального правдоподобия на непараметрический случай. Оценивание функции распределения

Если $\mathcal{F}$ – конечнопараметрическое семейство вероятностных мер, абсолютно непрерывных относительно меры Лебега, удовлетворяющее условиям регулярности, а $\theta = (\theta_1, \dots, \theta_p)$ – локальные координаты в окрестности $F^0 \in \mathcal{F}$ ($\theta = 0$ соответствует F^0), то для получения МП – оценки (оценки максимального правдоподобия) $F_n^*[\cdot]$ надо взять $F[\cdot \mid \theta_n^*]$, т.е. подставить вместо неизвестного значения параметра θ его МП – оценку θ_n^*. При выполнении условий регулярности θ_n^* определяется как решение системы уравнений

$$\sum_i \xi_j(X_i ; \theta_n^*) = 0 \quad , \text{ где } \xi_j(x ; \theta) = \frac{\partial f(x ; \theta)}{\partial \theta_j} / f(x ; \theta).$$

Нетрудно видеть, что с точностью до коэффициента I/2 функция $\xi_j(\cdot ; \theta)$ совпадает с касательным вектором к пути, направленному по координате θ_j ; условие $\sum_i \xi(X_i ; \theta) = 0$ означает, что $\int \xi(x ; \theta) F_n(dx) = 0$, где F_n – эмпирическая вероятность, построенная по выборке $X_1, \dots, X_n$. Если бы мы знали $\xi_j(\cdot ; \theta^0)$, то оценка θ_n^* при этих $\xi_j(\cdot ; \theta^0)$ определяла бы проекцию стохастической меры F_n на $T_{F^0}\mathcal{F}$. На самом же деле оценка $F_n^*[\cdot] = F[\cdot \mid \theta_n^*]$ является проекцией F_n на касательное пространство $T_{F_n^*}\mathcal{F}$ к семейству $\mathcal{F}$ в близкой точке F_n^*. При $n \to \infty$ $F_n^* \to F^0$ и мы вправе ожидать, поскольку $\mathcal{K}_{F_n^*} \to \mathcal{K}_{F^0}$ в S -топологии, что проекции стохастических мер $W_n = \sqrt{n}(F_n - F^0)$ на $T_{F_n^*}\mathcal{F}$ будут в пределе иметь то же распределение, что и проекция предельной (гауссовски распределенной) стохастической меры W на предельное касательное пространство $T_{F^0}\mathcal{F}$.

Определение 4.I. Оценкой максимального правдоподобия (МП-оценкой) функции распределения $F \in \mathcal{F}$ мы называем оценку, получаемую проектированием эмпирической вероятности F_n на $T_F\mathcal{F}$, или асимптотически эквивалентную ей при $n \to \infty$.

Оговорка в определении 4.I "или асимптотически эквивалентную ей" вызвана тем, что в ряде случаев вместо проектора $\mathcal{T}_F$ подставляется его состоятельная (в S -потологии) оценка и полученная проекция эмпирической вероятности не будет совпадать с $F_n \circ \mathcal{T}_F$ но будет эквивалентна ей при $n \to \infty$.

Известно, что $W_n = \sqrt{n}\,(F_n - F)$ сходится к гауссовской стохастической мере W (в смысле Дадли, см. выше определение 2.5) с нулевым средним и ковариациями $E[W(A)W(B)] = F(A \cap B) - F(A)F(B)$. (Это верно в пространстве Д [I]). Рассмотрим некоторые специальные проекции W, часто встречающиеся в статистических задачах.

I) $\mathcal{F}_1$ - конечномерное семейство вероятностных мер;

2) $\mathcal{F}$ выделяется из семейства всех абсолютно непрерывных мер конечным числом условий типа $\{\Phi_i(F) = 0\,;\ i = 1, \dots, p\}$;

3) $\mathcal{F}$ выделяется из семейства всех абсолютно непрерывных мер условиями симметрии.

В случае I) при выполнении условий регулярности для оценки F_n^* получаем известную теорему о предельном распределении оценки максимального правдоподобия: θ_n^* асимптотически нормальна со средним θ^0 и асимптотической ковариационной матрицей $\frac{1}{n}(I(\theta^0))^{-1}$. Соответственно F_n^* асимптотически нормально и асимптотическое распределение $V_n = \sqrt{n}\,(F_n^* - F)$ совпадает с проекцией W на $T_F\mathcal{F}_1$.

Учитывая, что $U_n = W_n - V_n = \sqrt{n}(F_n - F_n^*)$ некоррелирована с V_n, $W_n \to W$, $V_n \to V$ по распределению, U сходится по распределению к $U = W - V$.

Последний факт был установлен Дурбином [3] для случая k = I.

Случай 2) - так называемое условное оценивание [5, 6]. Пусть $\Phi_1, \dots, \Phi_p$ - функции на семействе $\mathcal{F}$ (например, $\mathcal{F}$ - все абсолютно непрерывные вероятностные меры). Предположим, что отображение $\Phi : \mathcal{F} \to \mathcal{R}^p$, $\Phi(F) = (\Phi_1(F), \dots, \Phi_p(F))$ дифференцируемо и регулярно в смысле следующего определения.

Определение 4.2. $\Phi : \mathcal{F} \to \mathcal{R}^p$ называется дифференцируемым в точке $F \in \mathcal{F}$, если существует непрерывный линейный оператор $T_F\Phi : T_F\mathcal{F} \to \mathcal{R}^p$ такой, что для всякого а.н.д. пути γ, проходящего через F при $h = 0$, суперпозиция отображений $\Phi \circ \gamma$: $h \mapsto \gamma(h) \to \Phi(\gamma(h))$ дифференцируема в нуле и производная $(\Phi \circ \gamma)'(0)$ равна $T_F\Phi(\gamma'(0))$. Φ называется дифференцируемым (на $\mathcal{F}$), если Φ дифференцируемо в каждой точке $F \in \mathcal{F}$. Φ называется регулярным в точке F, если $T_F\Phi$ отображает $T_F\mathcal{F}$ на

$\mathfrak{R}^p$. Φ называется регулярным (на $\mathcal{F}$), если Φ регулярно в каждой точке $F \in \mathcal{F}$.

Рассмотрим регулярное отображение $\Phi : \mathcal{F} \to \mathfrak{R}^p, \mathcal{F}_1 = \{F \in \mathcal{F} : \Phi(F) = 0\}$. Если в $\mathfrak{R}^p$ выбрана ситема координат и $e_1, \ldots, e_p$ - базис $\mathfrak{R}^p$, то производную $T_F \Phi$ можно записать в виде

$$T_F \Phi(\xi) = \alpha_1(\xi) e_1 + \ldots + \alpha_p(\xi) e_p \qquad \text{, где } \xi \in T_F \mathcal{F}; \alpha_1, \ldots, \alpha_p -$$

непрерывные линейные функционалы на $T_F \mathcal{F}$. Так как $T_F \mathcal{F}$ - подпространство гильбертова пространства, а поэтому само - гильбертово, то существуют $\varphi_1, \ldots, \varphi_p \in T_F \mathcal{F}$ такие, что $\alpha_i(\xi) = \int \xi \varphi_i \, dF$. Пусть а.н.д. путь γ лежит в $\mathcal{F}_1$. Тогда его производная $\gamma'(0) = \xi$ такова, что $\alpha_i(\xi) = 0$, $i = 1, \ldots, p$. Поэтому касательное пространство к $\mathcal{F}_1$ в точке F совпадает с множеством таких $\xi \in T_F \mathcal{F}$, что $\alpha_i(\xi) = 0$, т.е. ξ ортогонально φ_i .

В этом случае можно по выборке $X_1, \ldots, X_n$ и эмпирической вероятности F_n построить МП - оценку функции распределения $F_n^* = F_n \circ \mathcal{T}_F$, где $\mathcal{T}_F$ - ортопроектор из $\hat{L}^2(F)$ на $T_F \mathcal{F}_1$. Для $V_n = \sqrt{n}(F_n^* - F)$ предельное распределение такое же, как условное распределение W при условии $\{\int \varphi_i \, dW = 0, \ i = 1, \ldots, p\}$, т.е. проекция W на $T_F \mathcal{F}_1$ - ортогональное дополнение к линейному подпространству в $\hat{L}^2(F)$, порожденному $\varphi_1, \ldots, \varphi_p$.

Случай I) и 2) в известном смысле дуальны - в случае I) $T_F \mathcal{F}_1$ имеет конечную размерность, в случае 2) $T_F \mathcal{F}_1$ - конечную коразмерность (в $\hat{L}^2(F)$).

Рассмотрим теперь случай 3).

Пусть сперва $k = 1$, т.е. наблюдается одномерная случайная величина X . Мы предполагаем, что группа $\mathcal{G}_0$ состоит из двух движений прямой - тождественного $g_0(x) = x$ и отражения $g_1(x) = -x$; семейство $\mathcal{F}_0$ состоит из обычных симметричных распределений (с четными плотностями), а $\mathcal{F}^{(s)}$ - из всевозможных сдвигов распределений из $\mathcal{F}_0$.

В этом случае можно усилить лемму 3.I.

Л е м м а 4.I. I) Проекция $\mathcal{T}_F$ из $\hat{L}^2(F)$ на $T_F \mathcal{F}_0$ определяется так:

$$(\mathcal{T}_F \varphi)(x) = \frac{1}{2}(\varphi(x) + \varphi(-x)); \ (\varphi - \mathcal{T}_F \varphi)(x) = \frac{1}{2}(\varphi(x) - \varphi(-x)) -$$

проекция на ортогональное дополнение к $T_F \mathcal{F}$ в $\hat{L}^2(F)$.

2) Если w – неопределенный интеграл от стохастической меры W, т.е. (поскольку k = I)

$$w(x) = W(A_{<x}) = \int_{-\infty}^{\infty} \mathbb{1}_{A<x}(y)\, W(dy),$$

где

$$A_{<x} = \{ y \in \mathfrak{R} : y < x \},$$

то легко можно получить неопределенные интегралы для V -проекции W на $T_F \mathfrak{F}$ и $U = W - V$:

$$v(x) = \frac{1}{2}\left[w(x) - w(-x)\right]; \quad u(x) = \frac{1}{2}\left[w(x) + w(-x)\right].$$

Доказательство. Первая часть леммы немедленно следует из леммы 3.I. Чтобы доказать вторую часть, заметим следующее: пусть $A_{\geqslant x} = \{ y \in \mathfrak{R};\ y \geqslant x \}$. Тогда $\mathbb{1}_{A<x} + \mathbb{1}_{A \geqslant x} = 1$ – функция, тождественно равная I.

С другой стороны, $\mathcal{K}_F(\mathbb{1}_{A<x})(y) = \frac{1}{2}\left[\mathbb{1}_{A<x}(y) + \mathbb{1}_{A<x}(-y)\right] =$

$$= \frac{1}{2}\left[\mathbb{1}_{A<x}(y) + \mathbb{1}_{A>-x}(y)\right],$$

поскольку стохастическая мера W абсолютно непрерывна в среднеквадратическом относительно меры Лебега, то $\mathbb{1}_{A>x}$ и $\mathbb{1}_{A\geqslant x}$ при интегрировании по W дают одинаковый результат. Таким образом,

$$\int \frac{1}{2}\left[\mathbb{1}_{A<x}(y) + \mathbb{1}_{A>-x}(y)\right] W(dy) = \frac{1}{2}\left[w(x) + \int_{-x}^{\infty} dw(y)\right] =$$

$$= \frac{1}{2}\left[w(x) - w(-x)\right],$$

что и требовалось доказать.

Для изучения семейства $\mathfrak{F}^{(s)}$ важен следующий результат.

Лемма 4.2. Касательное пространство $T_F \mathfrak{F}^{(s)}$ разлагается в прямую сумму двух взаимно ортогональных подпространств – одномерного подпространства, определяемого касательным вектором к пути, вдоль которого варьируется параметр сдвига c, а сама плотность f только сдвигается, т.е. $f(x; c(h)) = f(x - c(h)); f(x, c(0)) = \frac{dF}{dx}$; и пространства симметричных относительно $c(0) = c$ функций $\varphi \in \hat{L}^2(F)$

$\psi(x) = \varphi(2c - x)$.

Д о к а з а т е л ь с т в о. Рассмотрим сперва путь в множестве плотностей вида $f(x;h) = f(x - c(F) - h)$; $f(x - c(F)) = dF/d\lambda$; f - плотность, симметричная относительно 0. Тогда производная пути γ (если только γ является дифференцируемым при $h = 0$) имеет вид

$$\frac{\partial\sqrt{f(x;h)}}{\partial h} = \frac{1}{2}\,\frac{\partial f(x - c(F))}{\partial x}\cdot\frac{1}{\sqrt{f(x - c(F)}}\ .$$

Если плотность $f(\cdot)$ достаточно гладкая и $\sqrt{f}$ дифференцируема, то указанный путь будет дифференцируем. Он будет даже а.н.д., если $\frac{1}{f}\,\frac{\partial f}{\partial x} \in L^2(F)$. Нетрудно видеть, что если f четная, то $\partial\sqrt{f}/\partial x$ - нечетная функция и поэтому ортогональная пространству всех четных функций как в $L^2(\lambda)$, так и в $\hat{L}^2(F)$ (точнее, $\frac{\partial f}{\partial x}/\sqrt{f}$ ортогональна четным функциям в $L^2(\lambda)$, а $\frac{\partial f}{\partial x}/f$ ортогональна четным функциям в $\hat{L}^2(F)$).

Наоборот, пусть путь γ имеет вид (в терминах плотностей

$$f(x;h) = \frac{d\gamma(h)}{d\lambda}(x))\ \sqrt{f(x;h)} = \varphi(x - c(F), h)$$

, где $\varphi(\cdot\,;h)$ - функция четная при любом h; если γ - а.н.д. путь, то его производная при $h = 0$ является симметричной функцией относительно $c(F)$, т.е. $\gamma'(x + c(F)) = \xi(x)$, ξ - четная функция.

Наконец, если γ - произвольный путь в $\mathcal{F}^{(s)}$, то предположим $c(h)$ - центр распределения $\gamma(h)$; тогда, дифференцируя по h функцию $\varphi(x - c(h); h) = \sqrt{f(x;h)}$, если, например, $c(h)$ дифференцируема по h, получаем соотношение

$$\left.\frac{\partial\sqrt{f(x;h)}}{\partial h}\right|_{h=0} = \left.\frac{\partial c}{\partial h}\right|_{h=0}\frac{\partial\varphi}{\partial x}(x - c(0);) + \frac{\partial\varphi}{\partial h}(x - c(0); 0)\ ,$$

т.е. производная а.н.д. пути представляется в виде суммы симметричной относительно $c(0)$ функции $\partial\varphi/\partial h$ и функции, пропорциональной $\partial\varphi/\partial x$, что и требовалось доказать.

С л е д с т в и е 4.I. Проекция из $\hat{L}^2(F)$ на $T_F\mathcal{F}^{(s)}$ имеет вид

$$(\pi_F\varphi)(x) = \frac{1}{2}\left[\varphi(x) + \varphi(2c(F) - x)\right] + \left[\int\varphi(y)\psi(y)f(y)\lambda(dy)\right]\cdot\psi(x),$$

где $c(F)$ – центр распределения F,

$$\psi(y)=\left[\frac{df(y)}{dy}\Big/f(y)\right]:\left(\int\left[\frac{df(y)}{dy}\cdot\frac{1}{f(y)}\right]^2 f(y)\lambda(dy)\right)^{1/2}$$

Здесь первое слагаемое есть проекция на $T_F\mathcal{F}_{c(F)}$, а второе – проекция на одномерное семейство с касательным вектором ψ, в котором варьируется только параметр сдвига $c(\cdot)$.

Следствие 4.2. Для стохастических мер: проекция из $\hat{L}^2(F)$ на $T_F\mathcal{F}$ применяется к стохастической мере W;

$$(\mathcal{K}_F W)(A)=\frac{1}{2}\left[W(A)+W(2c(F)-A)\right]+\left(\int\psi dW\right)\left[\int_A\psi(y)f(y)\lambda(dy)\right].$$

Поскольку для множества $A=A_{<x}$ множество $c-A$ есть $A_{>c-x}$, то окончательно для неопределенного интеграла

$$v(x)=(\mathcal{K}_F W)(A_{<x}) \qquad \text{имеем}$$

$$v(x)=\frac{1}{2}\left[w(x)-w(2c-x)\right]+\int_{-\infty}^{\infty}\psi(y)dw(y)\cdot\int_{-\infty}^{\infty}\psi(y)f(y)\lambda(dy).$$

Рассмотрим теперь одномерную выборку $X_1,\dots,X_n$, где X_1 подчиняется неизвестному распределению $F\in\mathcal{F}^{(s)}$. Функция распределения F имеет вид $F(x)=F^0(x-c)$, где c – параметр сдвига, F^0 – дифференцируемая функция распределения, плотность которой, $f^0=dF^0/dx$ – четная функция.

Оценка функции распределения F строится в два этапа: I) оценивается параметр сдвига c; 2) производится симметризация F_n относительно оцененного значения

$$\tilde{F}_n(x)=\frac{1}{2}\left[F_n(x)+1-F_n(2c_n-x)\right], \text{ где } c_n \text{ – оценка } c.$$

Рассмотрим оценивание параметра сдвига. Неизвестная плотность имеет вид $f(x)=f^0(x-c)$. Рассмотрим путь в $\mathcal{F}^{(s)}$, вдоль которого изменяется только c, а f^0 остается неизменной. Производная такого пути (как элемент $L^2(\lambda)$) при $c=c^0$ есть

$$-\frac{1}{2}\frac{\partial f^0(x-c^0)}{\partial x}\cdot\frac{1}{\sqrt{f^0(x-c^0)}}$$, а соответствующий элемент $L^2(F)$ равен

$$-\frac{1}{2}\frac{\partial f^0(x-c^0)}{\partial x}\cdot\frac{1}{f^0(x-c^0)}.$$

Информационное количество Фишера для оценки c^0 (по семейству вида $\{f^0(x-c)\}$, где изменяется только c) есть $\mathcal{J}=\int_{-\infty}^{\infty} 1/f\,(f')^2 dx$. При этом [7], существует непараметрическая оценка c_n параметра сдвига c^0, асимптотическое распределение которой нормально со средним c^0 и дисперсией $1/n\mathcal{J}$.

Иными словами, если $\varphi(x)=-\dfrac{1}{f^0(x-c^0)}\dfrac{\partial f^0(x-c^0)}{\partial x}$, то $\mathcal{J}=\int \varphi^2 dx$, а асимптотическое распределение оценки c_n такое же, как распределение $c^0+\frac{1}{\sqrt{n}\mathcal{J}}\int\varphi(x)dw(x)$, где интеграл берется как стохастический интеграл Винера по процессу w, гауссовскому, с ковариациями $E[w(s)w(t)]=F(min(s,t))-F(s)F(t); Ew(t)=0$. В силу ортогональности φ всем четным функциям в $\hat{L}^2(F)$ (лемма 4.I) непараметрическая оценка функции распределения F будет асимптотически независима с оценкой c_n, так как их совместное предельное распределение такое же, как совместное распределение проекций гауссовской стохастической меры на подпространства $H_1=\{\alpha\varphi,\ \alpha\in\mathbb{R}\}$ и H_2, состоящие из функций вида $\psi(x-c^0)$, где ψ - четная, а эти подпространства ортогональны.

Перейдем теперь к построению оценки функции распределения. Положим $\widetilde{F}_n(x)=\frac{1}{2}\left[F_n(x)+1-F_n(2c_n-x)\right]$.

Т е о р е м а 4.I. Асимптотическое распределение стохастической меры V_n, неопределенный интеграл которой равен

$$v_n(x)=\sqrt{n}\,(\widetilde{F}_n(x)-F^0(x-c^0)),$$

совпадает с распределением проекции гауссовской стохастической меры W на подпространство в $\hat{L}^2(F)$, являющееся прямой суммой H_1 и H_2, т.е. на $T_F\mathcal{F}^{(s)}$. Соответсвенно $u_n(x)=\sqrt{n}\,(F_n(x)-\widetilde{F}_n(x))$ имеет асимптотическое распределение такое же, как проекция W на ортогональное дополнение к $T_F\mathcal{F}$ в $\hat{L}^2(F)$ и (v_n , u_n) асимптотически независимы.

Д о к а з а т е л ь с т в о. В одномерном случае нам более удобно работать с неопределенным интегралом. Поскольку $F_n(x)-F(x)=\dfrac{w_n(x)}{\sqrt{n}}$, то $F_n(x)=F^0(x-c^0)+\frac{1}{\sqrt{n}}w_n(x)$.

Таким образом, $$\widetilde{F}_n(x)=\frac{1}{2}\left[F^0(x-c^0)+1-F^0(2c_n-x-c^0)\right]+$$
$$+\frac{1}{2\sqrt{n}}\left[w_n(x)-w_n(2c_n-x)\right];$$

вычитая из $\widetilde{F}_n(x)$ теоретическую функцию распределения $F^0(x-c^0)$, имеем:

$$\widetilde{F}_n(x) - F^0(x-c^0) = \frac{1}{2}\left[1 - F^0(2c_n - x - c^0) - F^0(x-c^0)\right] + \\ + \frac{1}{2\sqrt{n}}\left[w_n(x) - w_n(2c_n - x)\right].$$

В силу симметрии теоретической функции распределения,

$$1 - F^0(2c_n - x - c^0) = F^0(x + 2(c^0 - c_n) - c^0).$$

Таким образом, получим

$$v_n(x) = \sqrt{n}\,(\widetilde{F}_n(x) - F(x)) = \frac{\sqrt{n}}{2}\left[F^0(x + 2(c^0 - c_n) - c^0) - \right. \\ \left. - F^0(x-c^0)\right] + \frac{1}{2}\left[w_n(x) - w_n(2c_n - x)\right].$$

Рассмотрим теперь на пространстве Д операторы h и h_n: Д $\to$ Д, для функции $w \in$ Д :

$$h_n(w)(x) = \frac{1}{2}\left[w(x) - w(2c_n - x)\right];\ h(w)(x) = \frac{1}{2}\left[w(x) - w(2c^0 - x)\right].$$

Нетрудно видеть, что когда $c_n \to c^0$, то $h_n(w)(\cdot) \to h(w)(\cdot)$ в равномерной метрике для непрерывных траекторий $w(\cdot)$.
Если $w_n \to w$, то $\sup\limits_x |w_n(x) - w(x)| \to 0$ при $n \to \infty$; поэтому

$$h_n(w_n)(x) - h(w)(x) = \left[h_n(w_n)(x) - h_n(w)(x)\right] + \left[h_n(w)(x) - h(w)(x)\right].$$

Второе слагаемое справа стремится к нулю при $n \to \infty$, так как c_n - состоятельная оценка c.

Поскольку $w_n \to w$, то $\sup\limits_x |w_n(x) - w(x)| < \varepsilon$ при $n > n_0$, поэтому $|w_n(2c_n - x) - w_n(2c^0 - x)| \le |w_n(2c_n - x) - w(2c_n - x)| + |w(2c_n - x) - w(2c^0 - x)| + |w(2c^0 - x) - w_n(2c^0 - x)|$.

Первое и третье слагаемые в правой части $< \varepsilon$, так как $w_n \to w$, а второе может быть сделано $< \varepsilon$ в силу непрерывности w . Поэтому $\frac{1}{2}\left[w_n(x) \pm w_n(2c_n - x)\right]$ имеет такое же асимптотическое распределение, как и $\frac{1}{2}\left[w_n(x) \pm w_n(2c^0 - x)\right]$, т.е. такое же, как $\frac{1}{2}\left[w(x) \pm w(2c^0 - x)\right]$.

Что касается $\frac{\sqrt{n}}{2}\left[F^0(x + 2(c^0 - c_n) - c^0) - F^0(x - c^0)\right]$, то здесь можно заметить, что предельное распределение этого слагаемого такое же, как проекции w на одномерное подпространство, натянутое на φ . Ввиду асимптотической независимости $\frac{1}{2}\left[w_n(x) - w_n(2c_n - x)\right]$ и $\frac{\sqrt{n}}{2}\left[F^0(x + 2(c^0 - c_n) - c^0) - F^0(x - c^0)\right]$ их сумма имеет такое же асимптотическое распределение, что и проекция W на $T_F\mathcal{F}$. Аналогичными рассуждениями доказывается, что и $u_n = w_n - v_n$ имеет нужное асимптотическое распределение.

Для случая $k > 1$ пусть, как это упоминалось ранее, $\mathcal{G}$ включает в себя все отражения относительно координатных плоскостей. Распределение F имеет вид $F(x) = F^0(x - c^0)$, где c^0 - k - мерный вектор сдвига, а F^0 имеет $\mathcal{G}$ - симметричную плотность

$$f^0(\pm x_1, \pm x_2, \dots, \pm x_k) = f^0(x_1, x_2, \dots, x_k).$$

По аналогии [7] можно построить асимптотически нормальную оценку c_n вектора c^0 с асимптотическими параметрами $(c^0, \frac{1}{n}\mathcal{J}^{-1})$, где

$$\mathcal{J}_{ij} = \int \frac{\partial f^0(x - c^0)}{\partial x_i} \frac{\partial f^0(x - c^0)}{\partial x_j} \frac{1}{f^0(x - c^0)} \lambda(dx).$$

Рассмотрим функции

$$\psi_{+1}(x_j) = x_j \ ; \ \psi^j_{-1,0}(x_j) = 2c^0_j - x_j \ ;$$

$$\psi^j_{-1}(x_j) = 2c_{(n)j} - x_j \ ;$$ пусть

$$\tilde{g}_{\varepsilon_1,\dots,\varepsilon_k}(x) = \tilde{g}_\varepsilon(x) = (\psi_{\varepsilon_1}(x_1), \psi_{\varepsilon_2}(x_2), \dots, \psi_{\varepsilon_k}(x_k)).$$

Кроме меры F_n, рассмотрим всевозможные меры, полученные применением $\tilde{g}_\varepsilon$ к F_n, т.е.

$$(\tilde{g}_\varepsilon F_n)[A] = F_n[\tilde{g}_\varepsilon(A)].$$

Теорема 4.2. Пусть $\tilde{F}_n(x) = \frac{1}{2^k}\sum_{\varepsilon}(\tilde{g}_\varepsilon F_n)(x)$. Тогда асимптотическое распределение $v_n(x) = \sqrt{n}\,(\tilde{F}_n(x) - F^0(x-c^0))$ такое же, как у проекции неопределенного интеграла w гауссовской стохастической меры W на $T_F\mathcal{F}^{(s)}$, а асимптотическое распределение $u_n(x) = \sqrt{n}\,[F_n(x) - \tilde{F}_n(x)]$ совпадает с распределением проекции w на ортогональное дополнение к $T_F\mathcal{F}^{(s)}$ в $\hat{L}^2(F)$. При этом u_n и v_n асимптотически независимы.

Доказательство теоремы 4.2. аналогично доказательству теоремы 4.I.

Изложенные приемы позволяют рассматривать и другие условия симметрии, сводящиеся к упоминавшимся выше путем несложного преобразования. Рассмотрим, например, задачу оценивания симметричной функции распределения G на $\mathcal{R}^2$ с условием симметрии, имеющим в терминах плотности вид

$$g(y_1, y_2) = g^0(y_1 - c, y_2), \text{ где } c \text{ - параметр сдвига,}$$

$$g^0(-y_1, y_2) = g(y_1, y_2).$$

Сделаем преобразование $\mathcal{R}^2$ (не зависящее от истинного распределения G): $y_1 = \frac{x_1 - x_2}{2}$; $y_2 = \frac{x_1 + x_2}{2}$.

Плотность g определяет плотность f (с точностью до коэффициента пропорциональности, так как преобразование не сохраняет двумерную меру Лебега, а изменяет ее в $(\sqrt{2})^2$ раз):

$$f(x_1, x_2) = g\left(\frac{x_1 - x_2}{2}, \frac{x_1 + x_2}{2}\right) = g^0\left(\frac{(x_1 - c) - (x_2 + c)}{2}, \frac{(x_1 - c) + (x_2 + c)}{2}\right).$$

Поскольку g инвариантна относительно изменения знака у первого аргумента, то f инвариантна относительно замены $(x_1 - c)$ на $(x_2 + c)$ и f имеет вид

$$f(x_1, x_2) = f^0(x_1 - c, x_2 + c), \text{ где } f^0(x_1, x_2) = f^0(x_2, x_1).$$

Иначе говоря, f инвариантна при отражении относительно оси

$x_1 - x_2 = 2c$. Тем самым, поскольку по выборке $Y_1 = (Y_1^1, Y_1^2), Y_2 = (Y_2^1, Y_2^2), \ldots, Y_n = (Y_n^1, Y_n^2)$ мы умеем строить симметричную оценку функции распределения G , то по выборке $X_1 = (X_1^1, X_1^2), X_2 = (X_2^1, X_2^2), \ldots, X_n = (X_n^1, X_n^2)$, производя преобразование $Y_i^1 = \frac{X_i^1 - X_i^2}{2}$, $Y_i^2 = \frac{X_i^1 + X_i^2}{2}$, можно построить оценку $\widehat{G}$ для G и обратным преобразованием перейти к $\widehat{F}_n$.

Несколько иная постановка задачи позволяет ввести двумерный вектор сдвига. Для этого рассмотрим снова преобразование

$y_1 = \frac{x_1 - x_2}{2}$, $y_2 = \frac{x_1 + x_2}{2}$, а для g предположим следующее условие симметрии:

$$g(y_1, y_2) = g^0(y_1 - c_1, y_2 - c_2),$$

$$g^0(\pm y_1, \pm y_2) = g^0(y_1, y_2).$$

Тогда f удовлетворяет условию симметрии

$$f(x_1, x_2) = f^0(x_1 - (c_1 + c_2), x_2 - (c_2 - c_1)),$$

$$f^0(x_1, x_2) = f^0(x_2, x_1) = f^0(-x_2, -x_1) = f^0(-x_1, -x_2).$$

Используя теоремы этого параграфа, можно записать оценку при условии симметрии для этих случаев.

Автор благодарен Ю.Н.Тюрину за внимание к работе.

Литература

1. Dudley R.M., Weak Convergence on Nonseparable Metric Spaces and Empirical Measures on Euclidean Spaces.- Illinois Journal of Math., 1966, v.10, n1, 109-126.

2. Dudley R.M. Gaussian Processes on Several Parameters.- Ann. Math. Stat., 1965, v.36, n3, 771-788.

3. Durbin J. Weak Convergence of the Sample Distribution Function when Parameters Are Estimated.- Ann. Math. Stat., 1973, v.1, n2, 279-290.

4. Т ю р и н Ю.Н. Линейная модель в многомерной непараметрической статистике. - В сб. Учен. зап. по статистике. М., Наука, 1974, т. XXVI, 7 - 24.

5. Л е в и т Б.Я. Условное оценивание линейных функционалов. - Проблемы передачи информации, 1975, т. XII, № 4, 39 - 54.

6. К о ш е в н и к Ю.А., Л е в и т Б.Я. О непараметрическом аналоге информационной матрицы. - Теория вероятн. и ее примен., 1976, т. XXI, № 4, 759 - 774.

7. Stone C.J. Adaptive Maximum Likelihood Estimators of a Location Parameter.- Annals of Statist., 1975, v.3, n 2, 267-284.

8. Б и л л и н г с л и П. Сходимость вероятностных мер. М., Наука, 1977.