Записки научных
семинаров ПОМИ
Том 431, 2014 г.

И. А. Ибрагимов

ОБ ОЦЕНИВАНИИ ПЛОТНОСТИ ИНТЕНСИВНОСТИ ПУАССОНОВСКОГО СЛУЧАЙНОГО ПОЛЯ В $\mathbb{R}^d$ ВНЕ ОБЛАСТИ НАБЛЮДЕНИЯ

§1. Постановка задачи. Результаты.

Пусть Π_ε – пуассоновское случайное поле в $\mathbb{R}^d$ с плотностью интенсивности относительно меры Лебега, равной $\frac{\lambda(x)}{\varepsilon}$, где $\varepsilon > 0$ – положительный малый параметр. Поле Π_ε наблюдается в некоторой области $G \subset \mathbb{R}^d$ (иными словами наблюдается пуассоновское случайное множество $\Pi_\varepsilon \cap G = \Pi_\varepsilon G$) в априорных предположениях $\lambda \in \mathbf{\Lambda}$, где $\mathbf{\Lambda}$ – известный заданный класс функций, а ε – известный параметр. Нас будут интересовать асимптотические результаты при $\varepsilon \to 0$.

Естественный пример такой задачи – наблюдения n независимых одинаково распределенных пуассоновских случайных полей $\Pi^1, \Pi^2, \dots,$ Π^n с плотностью интенсивности $\lambda(x)$. В задаче оценивания λ по наблюдениям Π^j объединение полей $\Pi_\varepsilon = \bigcup_1^n \Pi^j$ есть достаточная статистика. Поле Π_ε есть пуассоновское случайное поле с плотностью интенсивности $n\,\lambda(x)$, т.е. $\varepsilon = 1/n$.

Задача, которая исследуется ниже, заключается в следующем: поле Π_ε наблюдается в G, оценке подлежит значение $\lambda(x)$ функции λ в точке $x \notin G$. Решение сформулированной задачи не представляется возможным даже если $\mathbf{\Lambda}$ состоит из бесконечно дифференцируемых функций. Действительно, найдется сколько угодно бесконечно дифференцируемых плотностей λ, которые совпадают в области G и различаются вне любой окрестности G. С другой стороны, потеря наблюдений поля вне G мало скажется на точности оценивания $\lambda(x)$ внутри G.

Ситуация меняется, если $\mathbf{\Lambda}$ состоит из аналитических функций. В самом деле, если функция $f(z)$ аналитична в области $G_1 \subseteq \mathbb{C}^d$ и $G \subset G_1$, то $f(z)$ однозначно определяется в G_1 своими значениями в

Ключевые слова: процесс Пуассона, теорема единственности, непараметрические оценки, неравенство Крамера–Рао.

Работа выполнена при частичной поддержке гранта РФФИ 14-01-00856, гранта НШ-2504.2014.1 и Программы фундаментальных исследований РАН "Современные проблемы теоретической математики".

(вещественной) области G. Поэтому естественно возникают два следующих вопроса. Во-первых, возможно ли состоятельно оценить $\lambda(x)$ в точках x, отстоящих от G на расстояние r_ε, $r_\varepsilon \to \infty$, когда $\varepsilon \to 0$, и как r_ε зависит от ε. Во-вторых, в какой степени потеря наблюдений поля Π_ε вне G влияет на точность оценивания $\lambda(x)$ внутри G. (Аналогичные проблемы для других статистических задач оценивания исследовались автором в [2–4]).

Ответ на первый вопрос дают теоремы 1.1 и 1.2 ниже (ср. [2–4]). Обозначим $\mathbf{\Lambda} = \mathbf{\Lambda}(M, \sigma, \rho)$, $\sigma = (\sigma_1, \dots, \sigma_d)$, $\rho = (\rho_1, \dots, \rho_d)$, – класс целых аналитических плотностей интенсивности λ, таких что для всех $z = (z_1, \dots, z_d) \in \mathbb{C}^d$

$$\sup_{\substack{z:\, |z_i| \leqslant R_i,\\ i=1,\dots,d}} |\lambda(z)| \leqslant M \exp\Big\{ \sum_1^d \sigma_i R_i^{\rho_i} \Big\}. \tag{1.1}$$

Здесь $M > 0$, $\sigma_i > 0$, $\rho_i > 0$ – параметры, определяющие класс $\mathbf{\Lambda}$.

Если G есть область наблюдения поля Π_ε, то ниже через $G_\varepsilon(\tau)$, $\tau > 0$, условимся обозначать r_ε-окрестность G, где $r_\varepsilon = r_\varepsilon(\tau) = |\ln \varepsilon|^{\tau/\varrho}$, $\varrho = \max(\rho_1, \rho_2, \dots, \rho_n)$.

Теорема 1.1. *Пусть* $\mathbf{\Lambda} = \mathbf{\Lambda}(M, \sigma, \rho)$. *Существует такая оценка* $\lambda_\varepsilon(x)$ *плотности интенсивности* $\lambda(x)$, *построенная по наблюдениям* $\Pi_\varepsilon G$, *что для любых фиксированных* $\beta < \alpha < 1$ *выполняется неравенство*

$$\sup_{\lambda \in \mathbf{\Lambda}} \mathbf{E}_\lambda \Big\{ \sup_{G_\beta} |\lambda(x) - \lambda_\varepsilon(x)| \Big\} \leqslant C_{\alpha,\beta}\, \varepsilon^{\frac{(1-\alpha)}{2d}}. \tag{1.2}$$

Постоянные $C_{\alpha,\beta}$ *зависят также от* M, σ_j, ρ_j, d, G.

Следующая ниже теорема показывает, что размер области состоятельности $G_\varepsilon(\beta)$ предыдущей теоремы не может быть существенно увеличен.

Теорема 1.2. *Пусть* $\mathbf{\Lambda} = \mathbf{\Lambda}(M,\ \sigma,\ \rho)$. *Пусть* x *– произвольная точка внешности множества* $G_\varepsilon(\alpha)$, $\alpha > 1$. *Тогда*

$$\inf_{\widehat{\lambda}} \sup_{\lambda \in \mathbf{\Lambda}} \mathbf{E}_\lambda \big|\, \lambda(x) - \widehat{\lambda}\, \big|^2 \geqslant c_0 > 0. \tag{1.3}$$

Здесь нижняя грань берется по всем возможным оценкам $\widehat{\lambda}$ *величины* $\lambda(x)$, *постоянная* c_0 *зависит от* α, M, σ, ρ, G.

При ответе на второй вопрос ограничимся плотностями класса $\mathbf{\Lambda}(M, \sigma, \mathbf{1})$, т.е. плотностями, которые являются целыми функциями экспоненциального типа, предположив еще дополнительно, что

$$\int_{\mathbb{R}^d} \lambda(x)\,dx \leqslant 1. \tag{1.4}$$

Из последнего условия и свойств функций экспоненциального типа следует, что $\lambda \in L_p(\mathbb{R}^d)$, $p \geqslant 1$. В частности, функция $\lambda \in L_2(\mathbb{R}^d)$ и по теореме Пэли–Винера–Пойа она допускает представление вида (см., напр., [11])

$$\lambda(x) = \int_K e^{i(t,x)}\psi(t)\,dt, \quad \psi \in L_2(\mathbb{R}^d), \tag{1.5}$$

где K – ограниченное симметричное выпуклое тело в $\mathbb{R}^d$.

Всякая функция вида (1.5) принадлежит $\mathbf{\Lambda}(M, \sigma, \mathbf{1})$. Обозначим через $\mathbf{\Lambda}(K)$ класс плотностей интенсивности, представимых в виде (1.5) и удовлетворяющих дополнительному условию (1.4).

Теорема 1.3. *Пусть $\lambda \in \mathbf{\Lambda}(K)$. Существует оценка $\lambda_\varepsilon(x)$ плотности интенсивности $\lambda(x)$ по наблюдениям поля Π_ε во всем пространстве $\mathbb{R}^d$, такая что*

$$\mathbf{E}_\lambda\{\|\lambda - \lambda_\varepsilon\|_2^2\} \leqslant \varepsilon\frac{\operatorname{mes} K}{(2\pi)^d}. \tag{1.6}$$

Ср. последний результат с [7].

Теорема 1.4. *Пусть $\lambda \in \mathbf{\Lambda}(K)$. Пусть G – ограниченная область в $\mathbb{R}^d$. Имеют место неравенства*

$$c_1\varepsilon\left|\frac{\ln\varepsilon}{\ln|\ln\varepsilon|}\right|^d \leqslant \inf_{\widehat{\lambda}}\sup_{\lambda\in\mathbf{\Lambda}(K)}\int_G |\lambda(x) - \widehat{\lambda}(x)|^2\,dx \leqslant c_2\varepsilon\left|\frac{\ln\varepsilon}{\ln|\ln\varepsilon|}\right|^d. \tag{1.7}$$

Нижняя грань здесь берется по всем оценкам $\widehat{\lambda}$, построенным по $\Pi_\varepsilon G$, а константы $c_1 \leqslant c_2$ положительны и зависят от d, K, G.

Неравенства двух последних теорем показывают, насколько страдает точность оценивания от потери наблюдений, не принадлежащих области G.

Мы не ставим своей целью указать границы для констант в теоремах 1.1, 1.2, 1.4, что же касается порядка убывания по ε, то он не зависит от G. Поэтому естественно при доказательстве этих теорем выбрать в качестве G достаточно простые множества, скажем, кубы. Если теоремы будут доказаны для всех кубов, то, в силу того, что для любой ограниченной области G найдутся кубы Γ_1, Γ_2 со сторонами параллельными осям координат и такие, что $\Gamma_1 \subseteq G \subseteq \Gamma_2$, теоремы будут доказаны для всех G. При этом ясно, что достаточно рассмотреть какой-нибудь один куб. Поэтому ниже в ходе доказательства мы всегда предполагаем, не оговаривая этого особо, что $G = [-1, 1]^d$.

Заметим еще, что математические ожидания вычисляются по отношению к "истинной" плотности интенсивности λ и правильнее, как это и сделано выше, писать $\mathbf{E}_\lambda(\,\cdot\,)$. Но мы часто опускаем индекс λ, сохраняя его лишь в тех случаях, когда его отсутствие может привести к недоразумению.

§2. Верхние границы. Построение оценок. Доказательство теорем 1.1, 1.3

2.1. Доказательство теоремы 1.1. Ниже мы явно построим оценки, на которых достигаются нужные неравенства. В качестве подходящих кандидатов рассмотрим проекционные оценки Н. Н. Ченцова [13]. Именно, пусть $\{\varphi_j\}$ – какой-нибудь ортонормальный базис в $L_2(G)$ и

$$\lambda(x) = \sum_j a_j\,\varphi_j(x), \quad a_j = \int\limits_G \lambda(x)\,\varphi_j(x)\,dx,$$

– разложение $\lambda(x)$ в ряд Фурье по этому базису. Проекционные оценки для λ имеют следующий вид

$$\widehat{\lambda}(x) = \sum_0^N \widehat{a}_j\,\varphi_j(x),$$

где $\widehat{a}_j$ – оценки для a_j (их обычно легко найти, поскольку a_j линейно зависит от λ), а N выбирается в зависимости от априорной информации $\mathbf{\Lambda}$ о λ.

Так как мы приняли, что $G = [-1, 1]^d$, естественно в качестве функций φ_j выбрать классические полиномы $P_j(x)$ Лежандра (точнее – их многомерные аналоги), свойства которых хорошо изучены (см., напр., [10]). Именно, для $x = (x_1, \dots, x_d)$, $k = (k_1, \dots, k_d)$ положим

$$\varphi_k(x) = P_{k_1}(x_1)\cdots P_{k_d}(x_d),$$

где $P_{k_j}(x_j)$ – ортонормированные на $[-1,1]$ полиномы Лежандра. Полиномы Лежандра $P_j(y)$ образуют ортонормальный базис в $L_2([-1,1])$. Поэтому многомерные полиномы Лежандра образуют ортонормальный базис в $L_2(G)$. В частности, сужение плотности интенсивности λ на G разлагается в ряд Фурье по полиномам Лежандра:

$$\lambda(x) = \sum_k a_k\,\varphi_k(x), \tag{2.1}$$

причем коэффициенты Фурье–Лежандра

$$a_k = \int\limits_G \lambda(x)\,\varphi_k(x)\,dx = \int\limits_G \lambda(x)\,P_{k_1}(x_1)\cdots P_{k_d}(x_d)\,dx_1\cdots dx_d. \tag{2.2}$$

Условимся ниже посредством c, C с индексами или без обозначать константы (не обязательно одни и те же даже при полном графическом совпадении), зависящие лишь от M, σ, ρ, d.

Лемма 2.1 (см. [3]). *Если* $\lambda \in \mathbf{\Lambda}(M,\ \sigma,\ \rho)$, *то*

$$|a_k| \leqslant c\,M\,\exp\left\{-\frac{k_1}{\rho_1}\ln\frac{k_1}{e\,\rho_1\,\sigma_1} - \cdots - \frac{k_d}{\rho_d}\ln\frac{k_d}{e\,\rho_d\,\sigma_d}\right\}. \tag{2.3}$$

Лемма 2.2 (см. [10]). *Имеют место равенства*

$$\max_{|y|\leqslant 1}\ |P_j(y)| = |P_j(1)| = \sqrt{(2j+1)/2}\,. \tag{2.4}$$

Лемма 2.3 (см. [1], с. 74). *Если* $Q(z)$, $z \in \mathbb{C}^1$ *– полином степени* n, *то*

$$|Q(z)| \leqslant \max_{|x|\leqslant 1}\ |Q(x)|\cdot|z + \sqrt{z^2-1}|^n. \tag{2.5}$$

Из трех последних лемм следует, что ряд в правой части (2.1) сходится для всех $z \in \mathbb{C}^d$ к некоторой целой аналитической функции; в G этот ряд сходится к $\lambda(x)$ и, следовательно, для всех $z \in \mathbb{C}^d$

$$\lambda(z) = \sum_k a_k\,\varphi_k(z). \tag{2.6}$$

Рассмотрим в качестве оценок a_k статистики $\widehat{a}_k$, определяемые равенствами

$$\widehat{a}_k = \varepsilon \sum_{x\in\Pi_\varepsilon G} \varphi_k(x) = \varepsilon \int\limits_G \varphi_k(x)\,d\Pi_\varepsilon. \tag{2.7}$$

С вероятностью 1 поле $\Pi_\varepsilon G$ содержит лишь конечное число точек, так что статистики $\widehat{a}_k$ корректно определены. Определим статистики $\lambda_N(x)$ равенствами

$$\lambda_N(x) = \sum_{\substack{k=(k_1,\dots,k_d):\\ k_j \leqslant N}} \widehat{a}_k \, \varphi_k(x) \tag{2.8}$$

и исследуем их свойства. Для этого нам понадобится следующий известный результат (теорема Кэмпбелла, см. [8]).

Лемма 2.4. *Пусть* Π – *пуассоновское случайное поле в* $A \subseteq \mathbb{R}^d$ *с плотностью интенсивности относительно меры Лебега равной* μ *и пусть* $f(x)$ – *вещественнозначная функция на* A. *Положим*

$$S = \sum_{x \in \Pi} f(x) = \int_A f(x) \, d\Pi. \tag{2.9}$$

Если $\int_A \min\bigl\{|f(x)|,\, 1\bigr\}\, \mu(x)\, dx < \infty$, *ряд* S *с вероятностью* 1 *абсолютно сходится и, кроме того,*

$$\mathbf{E}(e^{\theta S}) = \exp\Bigl\{ \int_A (e^{\theta f(x)} - 1)\, \mu(x)\, dx \Bigr\} \tag{2.10}$$

для любого комплексного θ, *для которого сходится интеграл в правой части последнего равенства (в частности – для чисто мнимых* θ*). Кроме того,*

$$\mathbf{E} S = \int_A f(x)\, \mu(x)\, dx, \quad \mathbf{D} S = \int_A f^2(x)\, \mu(x)\, dx, \tag{2.11}$$

если сходятся интегралы справа.

В силу (2.11),

$$\mathbf{E}\, \widehat{a}_k = \int_G \varphi_k(x)\, \lambda(x)\, dx = a_k,$$

так что $\widehat{a}_k$ – несмещенная оценка a_k. Поэтому (см. леммы 2.1–2.3) для любого $R \geqslant 1$

$$\sup_{x:|x_i|\leqslant R} |\lambda(x) - \mathbf{E}\lambda_N(x)| \leqslant C \prod_{i=1}^{d} \Big(\sum_{j=0}^{\infty} R^j \exp\Big\{ -\frac{j}{\rho_i} \ln\Big(\frac{j}{e\,\sigma_i\,\rho_i} \Big) \Big\} \Big)$$
$$\times \sum_{i=1}^{d} \sum_{j=N+1}^{\infty} R^j \exp\Big\{ -\frac{j}{\rho_i} \ln\Big(\frac{j}{e\,\sigma_i\,\rho_i} \Big) \Big\} \tag{2.12}$$
$$\leqslant (CR)^{N+1} \exp\big\{ c(R^{\rho_1} + \cdots + R^{\rho_d}) \big\} \exp\Big\{ -\frac{N}{\varrho} \ln N \Big\}.$$

Снова в силу (2.11),

$$\mathbf{E}\,|a_k - \widehat{a}_k| \leqslant \sqrt{\mathbf{E}\,|a_k - \widehat{a}_k|^2} = \Big(\varepsilon \int_G \varphi_k^2(x)\,\lambda(x)\,dx \Big)^{1/2} \leqslant C\,\sqrt{\varepsilon}$$

и потому

$$\mathbf{E} \sup_{x:|x_i|\leqslant R} \big| \lambda_N(x) - \mathbf{E}\,\lambda_N(x) \big| \leqslant C\,\sqrt{\varepsilon} \sum_{k:\,k_j\leqslant N} \sup_x |\varphi_k(x)|$$
$$\leqslant C\,\sqrt{\varepsilon}\,N^d\,R^{(N+1)d}. \tag{2.13}$$

Таким образом,

$$\mathbf{E} \sup_{x:|x_i|\leqslant R} \big| \lambda(x) - \lambda_N(x) \big|$$
$$\leqslant (CR)^{(N+1)d} \Big(\exp\Big\{ -\frac{N}{\rho} \ln N + cR^{\varrho} \Big\} + \sqrt{\varepsilon} \Big). \tag{2.14}$$

Полагая далее $\widehat{\lambda}_\varepsilon(x)$ равным той оценке $\widehat{\lambda}_N$, для которой N минимизирует правую часть последнего неравенства, мы придем к утверждению теоремы 1.1. Теорема 1.1 доказана. □

Замечание. Нетрудно видеть, что утверждение теоремы 1.1 останется в силе, если в (1.2) брать верхнюю грань по комплексным окрестностям G. Этот результат, хотя и не имеет статистического смысла, интересен с точки зрения статистического взгляда на устойчивость теоремы единственности аналитических функций.

2.2. Доказательство теоремы 1.3. По условию теоремы

$$\lambda(x) = \int\limits_K e^{i(x,t)}\,\psi(t)\,dt = \int\limits_{\mathbb{R}^d} e^{i(x,t)}\,\psi_1(t)\,dt,$$

и, в свою очередь,

$$\psi_1(t) = \frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d} e^{-(x,t)}\lambda(x)\,dx.$$

Положим $k(x) = \int\limits_K e^{i(x,t)}\,dt$ и определим оценку $\lambda_\varepsilon(x)$ плотности интенсивности $\lambda(x)$ равенством

$$\lambda_\varepsilon(x) = \frac{\varepsilon}{(2\pi)^d}\sum_{y\in\Pi_\varepsilon} k(x-y) = \frac{\varepsilon}{(2\pi)^d}\int\limits_{\mathbb{R}^d} k(x-y)\,d\Pi_\varepsilon(y). \tag{2.15}$$

В силу (2.11) и равенства Парсеваля,

$$\mathbf{E}_\lambda(\lambda_\varepsilon(x)) = \frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d} k(x-y)\,\lambda(y)\,dy = \int\limits_K e^{i(x,t)}\psi(t)\,dt = \lambda(x), \tag{2.16}$$

так что $\lambda_\varepsilon(x)$ – несмещенная оценка $\lambda(x)$.

Далее, снова в силу (2.11),

$$\begin{aligned}\mathbf{E}_\lambda\,\|\lambda-\lambda_\varepsilon\|_2^2 &= \int\limits_{\mathbb{R}^d} \mathbf{E}_\lambda\big|\lambda_\varepsilon(x) - \mathbf{E}_\lambda\lambda_\varepsilon(x)\big|^2\,dx \\ &= \varepsilon\,\frac{1}{(2\pi)^{2d}}\int\limits_{\mathbb{R}^d} dx\int\limits_{\mathbb{R}^d} k^2(x-y)\,\lambda(y)\,dy \\ &= \varepsilon\,\frac{1}{(2\pi)^{2d}}\int\limits_{\mathbb{R}^d}\lambda(y)\,dy\int\limits_{\mathbb{R}^d} k^2(x)\,dx. \end{aligned}\tag{2.17}$$

На основании равенства Парсеваля

$$\frac{1}{(2\pi)^{2d}}\int\limits_{\mathbb{R}^d} k^2(x)\,dx = \frac{1}{(2\pi)^d}\int\limits_K dt = \frac{\operatorname{mes} K}{(2\pi)^d},$$

так что

$$\mathbf{E}_\lambda\,\|\lambda-\lambda_\varepsilon\|_2^2 \leqslant \varepsilon\,\frac{\operatorname{mes} K}{(2\pi)^d}\int\limits_{\mathbb{R}^d}\lambda(x)\,dx \leqslant \varepsilon\,\frac{\operatorname{mes} K}{(2\pi)^d}.$$

Теорема доказана. $\square$

2.3. Доказательство теоремы 1.4. Здесь мы докажем правое неравенство (1.7), указав оценки λ_ε, на которых оно достигается. Как и выше, определим оценки λ_N равенствами (2.8). Имеем

$$\mathbf{E}\,\|\lambda-\lambda_N\|_2^2=\sum_{k\colon k_j\leqslant N}\mathbf{E}|a_k-\widehat{a}_k|^2+\sum_{k\colon \max k_j>N}|a_k|^2. \tag{2.18}$$

Первое слагаемое справа равно

$$\varepsilon\int\limits_G\Big(\sum_{k_j\leqslant N}\varphi_k^2(x)\Big)\lambda(x)\,dx. \tag{2.19}$$

В силу условий теоремы $\lambda(x)\leqslant C$ и потому слагаемое (2.19) не превзойдет $\varepsilon\, CN^d$. Второе слагаемое справа в (2.18) не больше $C\,e^{-cN\ln N}$. Таким образом,

$$\mathbf{E}\,\|\lambda-\lambda_N\|_2^2\leqslant C(\varepsilon N^d+e^{-cN\ln N}).$$

Выбирая здесь $N\sim\frac{1}{c}\frac{\ln\frac{1}{\varepsilon}}{\ln\ln\frac{1}{\varepsilon}}$, придем к неравенству теоремы. □

§3. Доказательство теоремы 1.2.

Чтобы упростить изложение, мы не будем доказывать теорему для класса $\mathbf{\Lambda}(M,\sigma,\rho)$ с заданными M, σ, а проведем доказательство для класса $\mathbf{\Lambda}(M,\sigma\,\rho)$ с заданным ρ и какими-то M и σ. Конструкцию доказательства легко дополнить так, чтобы рассмотреть заданные M,σ.

Изложим общую схему доказательства теоремы 1.2. Без потери общности можно предположить, что оценивается значение плотности $\lambda(x)$ в точке $x^T=(T,0,\ldots,0)$, $T>0$. Рассмотрим однопараметрическое, зависящее от параметра $\theta\in[-\delta,\delta]$, семейство плотностей интенсивности $\lambda(x,\theta)$ вида

$$\lambda(x,\theta)=1+\theta\,g(x-T). \tag{3.1}$$

Здесь $g(x)$ – это неотрицательная или равномерно ограниченная на всей вещественной гиперплоскости функция класса $\mathbf{\Lambda}(M,\sigma,\rho)$, $\rho=(\rho_1,\ldots,\rho_d)$, $\rho_1=\max_i\rho_i=\varrho$. Функции $\lambda(x,\theta)$ действительно представляют собой плотности интенсивности, во всяком случае для достаточно малых δ. В силу ограниченности g, для малых δ выполняется неравенство $\lambda(x,\theta)\geqslant 0$. При всех значениях параметра $\theta\in[-\delta,\delta]$ плотности $\lambda(x,\theta)\in\mathbf{\Lambda}$ и потому

$$\inf_{\widehat\lambda}\sup_{\lambda\in\mathbf{\Lambda}}\big|\lambda(x^T)-\widehat\lambda\big|\geqslant\inf_{\widehat\lambda}\sup_{\theta\in[-\delta,\delta]}\big|\lambda(x^T,\theta)-\widehat\lambda\big|. \tag{3.2}$$

Рассмотрим теперь наряду с нашей основной задачей следующую вспомогательную задачу оценивания: оценить параметр θ по наблюдениям $\Pi_\varepsilon G$ с плотностью интенсивности Π_ε равной $\frac{1}{\varepsilon}\lambda(x,\theta)$, $\lambda(x,\theta)$ из семейства (3.1) (при этом мы, как и выше, предполагаем, что множество $G = [-1,1]^d$). Если $\widehat{\lambda}$ – какая-нибудь оценка значения $\lambda(x^T)$ плотности $\lambda(x)$ в основной задаче, то статистика

$$\widehat{\theta} = \frac{\widehat{\lambda} - 1}{g(0)}$$

может быть принята в качестве оценки параметра θ вспомогательной задачи. При этом

$$\mathbf{E}_{\lambda(\cdot\,,\theta)}|\widehat{\theta} - \theta|^2 = \frac{\mathbf{E}_\lambda|\widehat{\lambda} - \lambda|^2}{g(0)}\,. \tag{3.3}$$

Поэтому если g выбрано так, что $g(0) = 1$, то

$$\inf_{\widehat{\theta}}\ \mathbf{E}\,|\widehat{\theta} - \theta|^2 \leqslant \mathbf{E}\,\bigl|\widehat{\lambda} - \lambda(x^T,\theta)\bigr|^2$$

и потому

$$\inf_{\widehat{\lambda}}\sup_{\lambda\in\mathbf{\Lambda}}\mathbf{E}_\lambda\bigl|\lambda(x^T) - \widehat{\lambda}\bigr|^2 \geqslant \inf_{\widehat{\theta}}\sup_{\theta\in[-\delta,\delta]}\mathbf{E}_\theta|\theta - \widehat{\theta}|^2, \tag{3.4}$$

причем слева нижняя грань берется по всем оценкам $\widehat{\lambda}$ основной задачи, а справа – по всем оценкам $\widehat{\theta}$ вспомогательной задачи. Таким образом, чтобы доказать теорему 1.2, нам достаточно показать, что для всех малых ε правая часть в (3.4) больше положительной постоянной:

$$\inf_{\widehat{\theta}}\sup_{\theta\in[-\delta,\delta]}\mathbf{E}_\theta|\theta - \widehat{\theta}|^2 \geqslant c_0 > 0. \tag{3.5}$$

Это последнее неравенство мы и будем доказывать ниже, опираясь на неравенство Крамера–Рао и подбирая функцию $g(x)$ в (3.1) так, чтобы сделать информационное количество Фишера вспомогательной задачи как можно меньше. Последней цели можно добиться, как это видно из (3.1), минимизируя значения $|g(x)|$ на множестве $G - T$.

Лемма 3.1. *Пусть P_λ и P_μ – распределения пуассоновского поля* Π *в $A \subseteq \mathbb{R}^d$ с плотностями интенсивности λ и μ соответственно. Тогда*

$$\frac{dP_\mu}{dP_\lambda} = \exp\Bigl\{\int\limits_A \ln\frac{\mu(x)}{\lambda(x)}\,d\Pi - \int\limits_A[\mu(x) - \lambda(x)]\,dx\Bigr\}. \tag{3.6}$$

Доказательство леммы см. в [9], хотя там доказательство приведено для процессов с одномерным временем, переход к $d > 1$ не представляет затруднений.

Обозначим через P_θ распределение вероятностей наблюдений вспомогательной задачи, отвечающее значению параметра θ. В силу лемм 2.4 и 3.1, количество информации Фишера во вспомогательной задаче

$$I(\theta) = \mathbf{E}_\theta \Big| \frac{d}{d\theta} \ln \frac{dP_\theta}{dP_0}\big(\Pi_\varepsilon G\big)\Big|^2 = \frac{1}{\varepsilon} \int\limits_G \frac{g^2(x - x^T)}{1 + \theta\, g(x - x^T)}\, dx. \tag{3.7}$$

Заметим теперь, что семейство плотностей распределения $\frac{dP_\theta}{dP_0}$ гладко зависит от θ, смещение $b(\theta, \widehat{\theta}) = b(\theta)$ разумных оценок $\widehat{\theta}$ непрерывно дифференцируемо и если $I(\theta) \leqslant 1$, то в силу неравенства Крамера–Рао для любой оценки параметра θ во вспомогательной задаче выполняется неравенство (см., напр., [6])

$$\sup_\theta \mathbf{E}_\theta |\widehat{\theta} - \theta|^2 \geqslant \sup_\theta \big((b'(\theta) - 1)^2 + b^2(\theta)\big) \geqslant c_0,$$

где c_0 зависит только от δ.

Теперь для окончания доказательства теоремы достаточно построить функции g в (3.1) так, чтобы заданная равенством (3.7) информация Фишера удовлетворяла бы неравенству $I(\theta) \leqslant 1$. Подобные конструкции детально разобраны в [2–4]. Поэтому, отсылая читателя за деталями к этим работам, мы проведем доказательство лишь для случая, когда $\max \rho_j = \varrho = k$ есть целое число и когда $\varrho < 1$. Наша задача даже проще, чем рассмотренная в цитированных работах, поскольку нет ограничений на $\int_{\mathbb{R}^d} \lambda(x)\, dx$. Итак, положим $g(x) = e^{-x_1^k}$, если k – четное, и $g(x) = e^{x_1^k}$, если k – нечетное. В обоих случаях в силу (3.7), если $T \geqslant 2$,

$$I(\theta) \leqslant \frac{2^d}{\varepsilon} \exp\{-T^k/2^k\},$$

так что $I(\theta) \leqslant 1$ для $T \geqslant 2\big(d + \ln \frac{1}{\varepsilon}\big)^{1/k}$.

Для нецелых $\varrho \geqslant 1$ как и в [4] с использованием канонических произведений строятся функции $g(x)$, убывающие на полуоси $(0 \leqslant x_1 < \infty)$ как $e^{-x_1^\varrho}$, см. детали в [4], так что опять информация Фишера $I(\theta) \leqslant C\varepsilon^{-1} e^{-cT^\varrho}$, $c, C > 0$.

Если $1 > \varrho > 1/2$, мы положим $g(x) = \prod_{n=1}^\infty \big(1 - \frac{x_1}{n^{1/\varrho}}\big)$. Несложно показать, что для $x_1 > 0$ функция $g_T(x_1, 0, \ldots, 0) \leqslant Ce^{-x_1^\varrho}$, $c > 0$.

Ситуация немного усложняется, если $\varrho \leqslant 1/2$. Поскольку целые функции с $\varrho < 1/2$ не могут быть ограничены ни на каком луче, мы вынуждены выбирать $g(x)$ зависящим от T. Именно, положим $g_T(x) = \prod_{n>T^\varrho}^{\infty}\left(1 - \frac{x_1}{n^{1/\varrho}}\right)$. Тогда для $T-2 \leqslant x_1 \leqslant T+2$

$$\left|g_T(x_1, 0, \ldots, 0)\right| \leqslant \exp\Big\{-x_1 \sum_{n>T^\varrho} \frac{1}{n^{1/\varrho}}\Big\} \leqslant C\exp\{-cT^\varrho\},$$

где постоянная $c > 0$ зависит только от ϱ. Теорема доказана. □

Доказательство правого неравенства теоремы 1.3 в идейном плане сходно с доказательством теоремы 1.3 из [4], нужно лишь дополнительно использовать результаты из [5], и мы опустим его.

Литература

1. С. Н. Бернштейн, *Экстремальные свойства полиномов.* ОНТИ, М.-Л., 1937.
2. И. А. Ибрагимов, *Об экстраполяции целых функций наблюдаемых в гауссовском белом шуме.* — Украинский матем. журн. **21**, No. 3 (2000), 58–69.
3. И. А. Ибрагимов, *Об оценивании аналитической плотности распределения по цензурированной выборке.* — Зап. научн. семин. ПОМИ **311** (2004), 147–160.
4. И. А. Ибрагимов, *Об оценке многомерной аналитической плотности распределения по цензурированной выборке.* — Теория вероятн. и ее примен. **51**, No. 1, (2006), 95–108.
5. И. А. Ибрагимов, *О нижних границах точности непараметрического оценивания.* — Зап. научн. семин. ПОМИ **396** (2011), 102–110.
6. И. А. Ибрагимов, Р. З. Хасьминский, *Асимптотическая теория оценивания.* Наука, М., 1979.
7. И. А. Ибрагимов, Р. З. Хасьминский, *Об оценке плотности распределения, принадлежащей одному классу целых функций.* — Теория вероятн. и ее примен. **27**, No. 3 (1982), 514–524.
8. Дж. Кингман, *Пуассоновские процессы*, МЦНМО, М., 2007.
9. Yu. A. Kutoyants, *Statistical Inference for Spatial Poisson Processes*, Lecture Notes Statist., **134**, Springer, 1998.
10. Г. Сеге, *Ортогональные многочлены.* Физматгиз, М., 1962.
11. И. Стейн, Г. Вейс, *Введение в гармонический анализ на евклидовых пространствах.* Мир, М., 1974.
12. R. Hasminskii, I. Ibragimov, *On density estimation in the view of Kolmogorov's ideas in approximation theory.* — Ann. Statist. **18**, No. 3 (1990), 999–1010.
13. Н. Н. Ченцов, *Оценка неизвестной плотности распределения по наблюдениям.* — Докл. АН СССР **147**, 1 (1962), 45–48.

Ibragimov I. A. On the estimation of the intensity density function of Poisson random field outside of the observation region.

A Poisson random field with the intensity density function $\frac{\lambda(x)}{\varepsilon}$ is observed in a bounded region $G \subseteq \mathbb{R}^d$. It is supposed that the unknown function λ belongs to a known class of entire functions. The parameter ε is supposed to be known. The problem is to estimate the value $\lambda(x)$ at the points $x \notin G$. We consider an asymptotic setup of the problem when $\varepsilon \to 0$.

С.-Петербургское отделение
Математического института
им. В. А. Стеклова РАН,
Фонтанка 27, Санкт-Петербург 191023
и С.-Петербургский
государственный университет,
Университетский пр., 28,
Петродворец,
Санкт-Петербург 198504, Россия
E-mail: ibr32@pdmi.ras.ru

Поступило 26 ноября 2014 г.