Сибирский математический журнал
Май—июнь, 2003. Том 44, № 3

УДК 519.21

СКОРОСТЬ СХОДИМОСТИ РАСПРЕДЕЛЕНИЙ НОРМИРОВАННЫХ ОЦЕНОК МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ ДЛЯ НЕРЕГУЛЯРНЫХ ПАРАМЕТРИЧЕСКИХ СЕМЕЙСТВ

И. С. Борисов, Д. В. Миронов

Аннотация: Исследована скорость сходимости распределений нормированных оценок максимального правдоподобия, построенных по параметрическому семейству разрывных многомерных плотностей в случае векторного параметра.

Ключевые слова: оценки максимального правдоподобия, процесс отношения правдоподобия, обобщенное пуассоновское поле, асимптотическая факторизация распределений

§ 1. Введение и формулировка основного результата

Пусть $X_1, \dots, X_n$ — выборка, состоящая из независимых случайных векторов, распределенных в пространстве $\mathbb{R}^d$ с общей плотностью $f(x, \theta)$ (относительно меры Лебега), зависящей от неизвестного параметра $\theta \in \Theta \subseteq \mathbb{R}^m$. Предположим, что плотность непрерывна по $x = (x_1, \dots, x_d)$, за исключением точек множества K_θ. Предположим также, что K_θ при всех $\theta \in \Theta$ — гладкое многообразие размерности $d - 1$ и определены такие множества Ω_θ^1 и Ω_θ^2, которые, во-первых, вместе с K_θ образуют разбиение пространства $\mathbb{R}^d$, а во-вторых, в точках $y \in K_\theta$ плотность $f(x, \theta)$ имеет разрывы 1-го рода по направлениям, задаваемым множествами Ω_θ^1 и Ω_θ^2:

$$0 \neq q(y, \theta) = \lim_{x \to y|_{\Omega_\theta^1}} f(x, \theta), \quad 0 \neq p(y, \theta) = \lim_{x \to y|_{\Omega_\theta^2}} f(x, \theta),$$

где запись $x \to y|_{\Omega_\theta^j}$ означает, что $x \to y$ и при этом $x \in \Omega_\theta^j$.

Считая далее истинное значение параметра θ_0 фиксированным, введем случайное поле (логарифмический процесс отношения правдоподобия)

$$Y_n(u) = \sum_{i \le n} \ln \frac{f(X_i, \theta_0 + u/n)}{f(X_i, \theta_0)}.$$

Определим оценки максимального правдоподобия $\hat{\theta}_n$ следующим соотношением:

$$\limsup_{\theta \to \hat{\theta}_n} \prod_{i \le n} f(X_i, \theta) = \sup_{\theta \in \Theta} \prod_{i \le n} f(X_i, \theta).$$

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 03–01–00459).

Понятно, что такие оценки определяются, вообще говоря, не единственным образом. Так что всюду в дальнейшем в качестве оценки максимального правдоподобия будет выступать любая статистика, удовлетворяющая вышеприведенному соотношению. Отметим также, что нормированные оценки $\hat{u}_n = n(\hat{\theta}_n - \theta_0)$ являются точками максимума случайного поля $Y_n(u)$.

В [1] доказано, что при определенных предположениях $Y_n(u)$ можно аппроксимировать суммой обобщенного пуассоновского поля с линейным сносом и не зависящего от него некоторого вырожденного центрированного случайного поля. Иными словами, имеет место асимптотическая факторизация распределения $Y_n(u)$. Наша цель — показать, как это обстоятельство позволяет получать весьма точные оценки близости распределений нормированных оценок $\hat{u}_n$ к распределению точки максимума упомянутого обобщенного пуассоновского поля. Отметим, что впервые подобный феномен для несколько иного объекта был изучен в [2]. Этот феномен может быть кратко описан следующим образом. Предположим, что на некотором вероятностном пространстве нам удалось аппроксимировать с точностью $O(n^{-1})$ (скажем, по вероятности) изучаемый случайный элемент (например, нормированную сумму независимых случайных полей) суммой двух независимых случайных элементов, один из которых собственно и есть предельный, а другой, имеющий малость $O(n^{-1/2})$, центрирован. Тогда легко видеть, что точность приближения моментов ограниченных достаточно гладких в смысле Фреше функционалов от изучаемого случайного элемента соответствующими моментами от предельного элемента будет иметь порядок $O(n^{-1})$, поскольку первый член в разложении по формуле Тейлора разности гладкого функционала соответственно от упомянутой суммы двух независимых случайных элементов и предельного будет иметь нулевое математическое ожидание. По существу, этот же эффект присутствует в хорошо известном методе композиции (методе Линдеберга). Именно представление функции распределения нормированной оценки максимального правдоподобия в виде математического ожидания некоторого гладкого функционала от процесса $Y_n(u)$ и позволило использовать предложенные выше соображения при доказательстве нижеследующей теоремы 2. Отметим также, что в одномерной постановке этот план был ранее реализован в [3].

Прежде всего подробнее опишем изучаемые объекты и сформулируем необходимые условия.

1. Мы предполагаем, что многообразия K_θ ориентированы. Это означает, что в каждой точке $y \in K_\theta$ можно восстановить единичный вектор нормали $N(y,\theta)$, который как векторнозначная функция непрерывен по θ. Будем считать, что вектор $N(y,\theta)$ «обращен» в множество Ω^2_θ.

Многообразия K_θ представимы в параметрической форме

$$K_\theta = \{y(t,\theta);\ t \in [0,1]^{d-1}\}, \quad \theta \in \Theta,$$

где $y(t,\theta) = (y^1(t,\theta),\dots,y^d(t,\theta))$, а функции $y^i(t,\theta)$, $i = 1,\dots,d$, непрерывно дифференцируемы по параметрам $t = (t_1,\dots,t_{d-1})$ и $\theta = (\theta_1,\dots,\theta_m)$; соответствующие частные производные компонент параметризации обозначим через $y^i_{t_k}(t,\theta)$, $k = 1,\dots,d-1$, и $y^i_{\theta_j}(t,\theta)$, $j = 1,\dots,m$. Будем также предполагать, что вторые частные производные функций $y^i(t,\theta)$ по t_k ограничены равномерно по t и θ.

В силу компактности множества определения параметризации и приведенных условий ее гладкости многообразия K_θ с необходимостью ограничены.

Пусть, кроме того, для любого $\theta \in \Theta$ и любых $x' \in \Omega^1_\theta$, $x'' \in \Omega^2_\theta$ имеет место соотношение $[x', x''] \cap K_\theta \neq \varnothing$, где $[x', x''] = \{x' + h(x'' - x');\ h \in [0,1]\}$. Это означает, что при каждом $\theta \in \Theta$ множество K_θ представляет собой компактное многообразие без края (т. е. область $\Omega^1_\theta \cup \Omega^2_\theta$ не является односвязной).

Обозначим $\Theta_1(x) = \bigl\{\theta \in \Theta : x \in \Omega^1_\theta\bigr\}$, $\Theta_2(x) = \bigl\{\theta \in \Theta : x \in \Omega^2_\theta\bigr\}$.

2. Для всех $x \in \bigcup K_\theta$ в каждой из областей $\Theta_1(x)$ и $\Theta_2(x)$ существуют частные производные $l'_{\theta_j}(x,\theta)$ функции $l(x,\theta) = \ln f(x,\theta)$, ограниченные равномерно по всем θ и x и, кроме того, удовлетворяющие условиям

$$|l'_{\theta_j}(x,\theta+\delta) - l'_{\theta_j}(x,\theta)| \le L(x,\theta)|\delta|, \quad \mathbf{E}_\theta L^{2+\beta}(X_1,\theta) < \infty$$

при всех $\theta, \theta+\delta \in \Theta_i(x)$, $i = 1,2$, $1 \le j \le m$, и некотором $\beta > 0$, где $\mathbf{E}_\theta$ — математическое ожидание по распределению с плотностью $f(\cdot,\theta)$.

Здесь и всюду в дальнейшем символами c, C (с индексами или без) обозначаются различные постоянные, не зависящие от n, пространственных переменных и $\theta \in \Theta$, но зависящие в некоторых случаях от распределения X_1 и других параметров задачи. На протяжении всей статьи символом $|\cdot|$ будет обозначаться равномерная норма (покоординатный максимум) в том или ином конечномерном евклидовом пространстве, причем зависимость от размерности указываться не будет.

3. Частные производные $y^i_{\theta_j}(t,\theta)$ ограничены равномерно по t и θ и удовлетворяют условию Липшица по θ:

$$\left|y^i_{\theta_j}(t,\theta) - y^i_{\theta_j}(t,\theta+\delta)\right| \le c|\delta|, \quad \theta, \theta+\delta \in \Theta,\ t \in [0,1]^{d-1}.$$

4. Справедливы следующие неравенства:

$$|f(x,\theta) - q(y,\theta)| \le c|x-y|, \quad x \in \Omega^1_\theta,\ y \in K_\theta,$$

$$|f(x,\theta) - p(y,\theta)| \le c|x-y|, \quad x \in \Omega^2_\theta,\ y \in K_\theta,$$

$$|l(x,\theta) - \ln q(y,\theta')| \le c(|x-y| + |\theta-\theta'|), \quad x \in \Omega^1_\theta,\ y \in K_{\theta'},$$

$$|l(x,\theta) - \ln p(y,\theta')| \le c(|x-y| + |\theta-\theta'|), \quad x \in \Omega^2_\theta,\ y \in K_{\theta'}.$$

В частности, из двух последних неравенств следует, что $p(\cdot,\theta)$ и $q(\cdot,\theta)$ — непрерывные ограниченные функции на K_θ при всех $\theta \in \Theta$.

Процесс $Y(u)$, который является в известном смысле предельным для $Y_n(u)$, определяется следующим образом. Пусть W — случайная пуассоновская мера на $\mathbb{R} \times [0,1]^{d-1}$. На непересекающихся множествах она принимает независимые целочисленные значения, распределенные по закону Пуассона, и для любого $b \in \mathbb{R}$ и любого борелевского множества $B \subseteq [0,1]^{d-1}$ выполняются соотношения

$$\mathbf{E}W([0,b] \times B) = b \int\limits_B p(y(t,\theta_0),\theta_0) I(t,\theta_0)\,dt, \quad b > 0,$$

$$\mathbf{E}W([b,0] \times B) = -b \int\limits_B q(y(t,\theta_0),\theta_0) I(t,\theta_0)\,dt, \quad b < 0,$$

где $I(t,\theta_0) = \sqrt{\det \|(y_{t_i}(t,\theta_0), y_{t_j}(t,\theta_0))\|}$ — коэффициент искажения объема, возникающий при замене переменных в поверхностном интеграле.

Положим

$$D(t,\theta)=\begin{pmatrix} y^1_{\theta_1}(t,\theta) & \dots & y^d_{\theta_1}(t,\theta)\\ y^1_{\theta_2}(t,\theta) & \dots & y^d_{\theta_2}(t,\theta)\\ & \dots & \\ y^1_{\theta_m}(t,\theta) & \dots & y^d_{\theta_m}(t,\theta)\end{pmatrix},$$

$$\Pi(u)=\int\limits_{[0,1]^{d-1}}\ln\frac{q(y(t,\theta_0),\theta_0)}{p(y(t,\theta_0),\theta_0)}\int\limits_0^{(uD,N)}W(dv,dt),$$

где uD — произведение вектора u и матрицы $D=D(t,\theta_0)$, $N=N(y(t,\theta_0),\theta_0)$, а скобки в верхнем пределе внутреннего интеграла обозначают скалярное произведение. Определим процесс

$$Y(u)=\int\limits_{[0,1]^{d-1}}(uD(t,\theta_0),N(y(t,\theta_0),\theta_0))$$
$$\times[p(y(t,\theta_0),\theta_0)-q(y(t,\theta_0),\theta_0)]I(t,\theta_0)\,dt+\Pi(u).$$

Наконец, обозначим $l'_\theta(x,\theta)=(l'_{\theta_1}(x,\theta),l'_{\theta_2}(x,\theta),\dots,l'_{\theta_m}(x,\theta))$. В нижеследующих утверждениях условие Крамера для вектора $l'_\theta(X_1,\theta)$ означает, что существует окрестность нуля $U\subset\mathbb{R}^m$, для которой величина $\mathbf{E}_\theta\exp\{(u,l'_\theta(X_1,\theta))\}$ ограничена равномерно по $u\in U$ и $\theta\in\Theta$.

В [1] установлена асимптотическая факторизация распределений процессов $Y_n(u)$. В частности, из лемм 2–8 в [1] можно извлечь следующую удобную для нас версию этого результата.

Теорема 1. *Пусть выполнены приведенные выше условия, и, кроме того, пусть $\mathbf{E}_\theta|l'_\theta(X_1,\theta)|^\alpha<\infty$ при некотором $\alpha>2$. Тогда случайные процессы $Y_n(u)$ и $Y(u)$ можно так определить на одном вероятностном пространстве, что*

$$\mathbf{P}(\sup_{|u|\le T_n}|Y_n(u)-Y(u)-(u,\eta_n)|\ge c\gamma_n)\le c\gamma_n,$$

где $T_n=c_0\ln n$, $\gamma_n=(\ln n)^2n^{1/(1+\alpha)-1}$, $\eta_n=n^{-1}\sum\limits_1^n\xi_i$, а $\{\xi_i\}$ — заданные на том же вероятностном пространстве независимые одинаково распределенные центрированные случайные векторы с конечным моментом порядка α, не зависящие от $Y(u)$.

Если же для случайного вектора $l'_\theta(X_1,\theta)$ выполнено условие Крамера, то утверждение теоремы справедливо при $\gamma_n=(\ln n)^2n^{-1}$.

Замечание 1. В условиях работы [1] требовалось, чтобы функции $q(y,\theta)$ и $p(y,\theta)$ не принимали нулевых значений. Без выполнения этого условия не определена величина $\ln\frac{q(y,\theta)}{p(y,\theta)}$, которая используется при задании поля $Y(u)$. Однако в некоторых случаях это условие может быть ослаблено.

Пусть, например, $f(x,\theta)=0$ при всех $x\in\Omega^1_\theta$, $\theta\in\Theta$. Такому условию удовлетворяет, скажем, равномерное распределение в области. В этом случае $q(y,\theta)\equiv 0$. Случайные поля $Y_n(u)$ и $Y(u)$ кроме конечных значений будут принимать еще и значение $-\infty$. По построению они не могут принимать с положительной вероятностью значение $+\infty$. В случае, когда поля $Y_n(u)$ и $Y(u)$ одновременно тождественно в области определения равны $-\infty$, мы, очевидно, можем положить $Y_n(u)=Y(u)+(u,\eta_n)$ (или соответствующая разность полагается равной нулю) при любом способе задания случайного вектора η_n на одном вероятностном пространстве с указанными полями. Заметим, что согласно лемме 8 в [1] вероятность того, что при некотором значении u таком, что $|u| \leq T_n$, один из процессов $Y_n(u)$ или $Y(u)$ равен $-\infty$, а второй — нет, не превышает $c(\ln n)^2 n^{-1}$. Поэтому с дополнительной вероятностью $1 - c(\ln n)^2 n^{-1}$ обе траектории либо всюду конечны, либо тождественно равны $-\infty$. Так что при сделанном выше допущении касательно операций на расширенной числовой прямой утверждение теоремы 1 останется в силе и для приведенного примера. Отметим также, что рассмотренный пример имеет прямое отношение к известной проблеме построения статистических оценок носителя наблюдаемого распределения (см. [4]).

Обозначим через $\hat{u}$ случайную величину, удовлетворяющую соотношению

$$\limsup_{u \to \hat{u}} Y(u) = \sup_u Y(u).$$

В [5] при $d, m = 1$ и в [6] для произвольных d, m было показано, что распределение нормированной оценки максимального правдоподобия $\hat{u}_n$ при $n \to \infty$ слабо сходится к распределению $\hat{u}$ — точки максимума процесса $Y(u)$. Для случая $d, m = 1$ в [3] получены близкие к неулучшаемым оценки для равномерного расстояния между функциями распределения случайных точек $\hat{u}_n$ и $\hat{u}$.

Наша задача — получить оценку расстояния в равномерной метрике между функциями распределения случайных величин $\hat{u}_n$ и $\hat{u}$ при произвольных d и m. Для этого нам потребуются некоторые дополнительные предположения.

Обозначим через r вектор из $\mathbb{R}^m$ с координатами

$$r_i = \int\limits_{[0,1]^{d-1}} (ND^*)_i [p(y(t,\theta_0),\theta_0) - q(y(t,\theta_0),\theta_0)] I(t,\theta_0)\, dt,$$

где $(ND^*)_i$ — i-я координата вектора $N(y(t,\theta_0),\theta_0)D^*(t,\theta_0)$ (знак $*$ применяется для обозначения транспонированной матрицы):

$$(ND^*)_i = \sum_{j \leq d} y^j_{\theta_i}(t,\theta_0) N^j(y(t,\theta_0),\theta_0).$$

Тогда случайный процесс $Y(u)$ может быть представлен как $Y(u) = (u, r) + \Pi(u)$.

Обозначим $Y_{\eta_n}(u) = Y(u) + (u, \eta_n)$. Пусть u^*_δ — случайная величина, удовлетворяющая соотношению

$$\limsup_{u \to u^*_\delta} (Y(u) + (u,\delta)) = \sup_u (Y(u) + (u,\delta)).$$

Положим

$$h(\delta, x) = \mathbf{P}(u^*_\delta < x),$$

где $\delta = (\delta_1, \ldots, \delta_m)$ — произвольный m-мерный вектор, а знак $<$ соответствует стандартному (покоординатному) частичному порядку в $\mathbb{R}^m$. Обозначим через $h''_{\delta_i \delta_j}$ смешанную частную производную второго порядка (если таковая существует) функции $h(\delta, x)$ по соответствующим координатам вектора δ.

Предположим, что выполнены следующие условия.

A1. $r \neq 0$.

A2. Существует такое $\rho > 0$, что для любого δ из ρ-окрестности нуля в $\mathbb{R}^m$ и любого ненулевого вектора $\bar{l}$, ортогонального $\underline{r} + \delta$, множество точек t, в которых вектор $N(y(t,\theta_0),\theta_0)D^*(t,\theta_0)$ ортогонален $\bar{l}$, имеет нулевую лебегову меру.

В [6] доказано, что при выполнении условий А1 и А2 поле $Y(u)$ с вероятностью 1 имеет единственную точку максимума. Это, в частности, означает, что функция $h(\delta, x)$ определена корректно.

А3. Будем предполагать, что для всех $x \in \mathbb{R}^d$, а также i и j существует смешанная производная $h''_{\delta_i\delta_j}$, удовлетворяющая неравенству

$$\sup_{|\delta|<\rho} |h''_{\delta_i\delta_j}(\delta, x)| \leq c_1|x|^2 + c_2$$

с некоторыми положительными постоянными ρ, c_1, c_2.

Проверка последнего условия представляет основную трудность. При доказательстве основного утверждения в [3], по существу, установлено, что в случае $d, m = 1$ условие А3 следует из А1 и условий одномерного аналога теоремы 1. В § 3 настоящей работы мы покажем, что то же самое верно при $m = 1$ и произвольном d, а также установим справедливость А3 для конкретного примера семейства плотностей в случае $m, d > 1$. Отметим, что в многомерном случае по сравнению с [3] качественно меняется и усложняется методика соответствующих доказательств.

Теперь мы можем сформулировать основное утверждение настоящей работы.

Теорема 2. *Пусть выполнены условия теоремы* 1, *а также условия* А1–А3. *Тогда*

$$\sup_{x\in\mathbb{R}^d} |\mathbf{P}(\hat{u}_n < x) - \mathbf{P}(\hat{u} < x)| \leq c\gamma_n,$$

где $\gamma_n = (\ln n)^2 n^{1/(1+\alpha)-1}$ *в случае, когда у* $l'_\theta(X_1, \theta)$ *существует конечный момент порядка* α, *и* $\gamma_n = (\ln n)^2 n^{-1}$, *когда* $l'_\theta(X_1, \theta)$ *удовлетворяет условию Крамера.*

Замечание 2. Теорема 2 остается верной и в случае, когда мы исключаем условие $q(y, \theta)$, $p(y, \theta) \neq 0$, заменяя его такими требованиями, как, скажем, в замечании 1. В [3] рассмотрен соответствующий пример — равномерное распределение на отрезке и найден явный вид функций распределения $\hat{u}_n$ и $\hat{u}$. Равномерное расстояние между этими функциями распределения имеет порядок n^{-1}. Этот пример, в частности, показывает, что оценка в теореме 2 близка к неулучшаемой.

§ 2. Доказательство теоремы 2

Обозначим через u_n^* точку максимума случайного поля $Y_{\eta_n}(u) = Y(u) + (u, \eta_n)$, т. е. случайную величину, удовлетворяющую соотношению

$$\limsup_{u\to u_n^*} Y_{\eta_n}(u) = \sup_u Y_{\eta_n}(u). \tag{1}$$

Случайная величина, определяемая таким образом, вообще говоря, не единственна.

Для того чтобы доказать теорему, мы сначала получим оценки близости функций распределения $\hat{u}_n$ и u_n^*, а затем сравним функции распределения u_n^* и $\hat{u}$. Воспользуемся следующим неравенством:

$$\begin{aligned}|\mathbf{P}(u_n^* < x) - \mathbf{P}(\hat{u}_n < x)| &\leq \mathbf{P}(|u_n^* - \hat{u}_n| \geq c\gamma_n; |u_n^*| \leq T_n; |\hat{u}_n| \leq T_n)\\ &+ \mathbf{P}(u_n^* \in H_x^{c\gamma_n}; |u_n^*| \leq T_n) + \mathbf{P}(|u_n^*| > T_n) + \mathbf{P}(|\hat{u}_n| > T_n),\end{aligned}$$

где T_n и γ_n определены в теореме 1,

$$H_x^\varepsilon = \bigcup_{i=1}^m \{u = (u_1, \ldots, u_m) : u_i \in [x_i - \varepsilon, x_i + \varepsilon],\ -T_n \le u_j \le x_j + \varepsilon,\ j \ne i\};$$

иными словами, H_x^ε — это ε-окрестность (в смысле введенной выше нормы) «правой» границы множества $H_x = \{u : -T_n \le u_i < x_i,\ i = 1, \ldots, m\}$.

Лемма 1. *Пусть выполнены условия теоремы* 2. *Тогда максимум поля* $Y_{\eta_n}(u)$ *достигается в единственной точке* u_n^* *с вероятностью, не меньшей чем* $1 - c/n$. *Кроме того,*

$$\mathbf{P}(|u_n^* - \hat{u}_n| \ge c\gamma_n; |u_n^*| \le T_n; |\hat{u}_n| \le T_n) \le c\gamma_n.$$

Доказательство первого утверждения леммы практически не отличается от доказательства единственности точки максимума процесса $Y(u)$, приведенного в [6].

Итак, сначала отметим, что вне разрывов случайное поле $Y_{\eta_n}(u)$ имеет вид $(u, r + \eta_n) + b$, где b постоянна в каждой области непрерывности и зависит от разрывов, которые происходят вдоль некоторых случайных гиперплоскостей Γ_i. Найдем уравнения для Γ_i. Для этого представим поле $\Pi(u)$ при $|u| \le T_n$ в следующем виде:

$$\Pi(u) = \sum_{i=1}^{\pi_\lambda} \big(\mathbf{I}\{Y_i \in G^2(u)\} - \mathbf{I}\{Y_i \in G^1(u)\}\big) \ln \frac{q(\operatorname{pr} Y_i, \theta_0)}{p(\operatorname{pr} Y_i, \theta_0)}, \tag{2}$$

где

$$G^1(u) = \{\langle t, y\rangle \in [0,1]^{d-1} \times \mathbb{R} : t \in [0,1]^{d-1}, (uD(t,\theta_0), N(y(t,\theta_0),\theta_0))^- < y < 0\},$$

$$G^2(u) = \{\langle t, y\rangle \in [0,1]^{d-1} \times \mathbb{R} : t \in [0,1]^{d-1}, 0 < y < (uD(t,\theta_0), N(y(t,\theta_0),\theta_0))^+\},$$

$\langle t, y\rangle$ — d-мерный вектор, первая «координата» которого есть $(d-1)$-мерный вектор t; a^- и a^+ — отрицательная и положительная части числа a соответственно; $\operatorname{pr} x$ — ортопроекция x на $[0,1]^{d-1}$; π_λ — не зависящая от последовательности $\{Y_i\}$ пуассоновская случайная величина с параметром λ; $\{Y_i\}_{i=1}^\infty$ — последовательность независимых случайных величин, распределенных в $G = G^1 \cup G^2$, $G^1 = \bigcup\limits_{|u| \le T_n} G^1(u)$, $G^2 = \bigcup\limits_{|u| \le T_n} G^2(u)$ с плотностью

$$f_Y(x, \theta_0) = \frac{1}{\lambda} q(\operatorname{pr} x, \theta_0) I(\operatorname{pr} x, \theta_0), \quad x \in G^1,$$

$$f_Y(x, \theta_0) = \frac{1}{\lambda} p(\operatorname{pr} x, \theta_0) I(\operatorname{pr} x, \theta_0), \quad x \in G^2,$$

где величина λ выбирается так, чтобы

$$\int\limits_G f_Y(x, \theta_0)\, dx = 1.$$

Нетрудно убедиться в том, что $\lambda = A \ln n$, где константа A не зависит от n.

Поверхность Γ_i, соответствующая атому Y_i с координатами $\langle t_i, y_i\rangle$, проходит через точку u, если выполнено соотношение

$$y_i = (uD(t_i, \theta_0), N(y(t_i, \theta_0), \theta_0)).$$

Это соотношение задает в пространстве $\mathbb{R}^m$ случайную гиперплоскость с нормалью $N(y(t_i,\theta_0),\theta_0)D^*(t_i,\theta_0)$. В силу условия A1 экстремум поля $Y_{\eta_n}(u)$ достигается на разрывах, точнее на некотором многограннике, содержащемся в пересечении нескольких гиперплоскостей Γ_i. Предположим, что с вероятностью, большей чем c/n, экстремум достигается на многограннике, который состоит более чем из одной точки и содержится в пересечении гиперплоскостей Γ_i, $i=1,\dots,k$, $k<m$. Этот многогранник должен быть ортогонален вектору $r+\eta_n$. Следовательно, существует линейная комбинация нормалей к гиперплоскостям $\Gamma_1,\dots,\Gamma_k$, равная вектору $r+\eta_n$, т. е. справедливо соотношение

$$\sum_{i\le k}\alpha_i D^*(t_i,\theta_0)N(y(t_i,\theta_0),\theta_0)=r+\eta_n. \tag{3}$$

Зафиксируем векторы

$$D^*(t_1,\theta_0)N(y(t_1,\theta_0),\theta_0),\dots,D^*(t_{k-1},\theta_0)N(y(t_{k-1},\theta_0),\theta_0),\quad r+\eta_n$$

(случайная величина η_n не зависит от процесса $\Pi(u)$). На множестве, мера Лебега которого превышает c/n, вектор $D^*(t_k,\theta_0)N(y(t_k,\theta_0),\theta_0)$ должен лежать в некотором фиксированном подпространстве, порожденном векторами

$$D^*(t_1,\theta_0)N(y(t_1,\theta_0),\theta_0),\dots,D^*(t_{k-1},\theta_0)N(y(t_{k-1},\theta_0),\theta_0),\quad r+\eta_n.$$

Существует вектор $\bar{l}$, ортогональный этому подпространству, а значит, и вектору $r+\eta_n$. Осталось лишь заметить, что для случайной величины η_n верна оценка $\mathbf{P}(|\eta_n|>\rho)\le c/n$ (она легко получается с помощью неравенства Чебышева), и мы получаем противоречие с условием A2. Поскольку в точках пересечения любых двух наборов из m гиперплоскостей разрывов $Y_{\eta_n}(u)$ это поле принимает равные значения с нулевой вероятностью, его глобальный максимум с вероятностью, не меньшей чем $1-c/n$, достигается в единственной точке.

Перейдем теперь к доказательству второго утверждения леммы 1. Пусть $Y_n'(u)$ — вектор, составленный из частных производных соответствующего поля (дифференцирование производится потраекторно). В [1] было установлено, что, когда $|u|<c\ln n$, поля $Y_{\eta_n}(u)$ и $Y_n(u)$ можно построить таким образом, чтобы поверхности разрывов полей совпадали с вероятностью, не меньшей чем $1-c\gamma_n$. Из этого следует, что максимум поля $Y_n(u)$ достигается в той же области непрерывности, что и максимум поля $Y_{\eta_n}(u)$, с вероятностью, не меньшей чем $1-c\gamma_n$.

Далее, в [1] была получена оценка

$$\mathbf{P}\big(\sup_{|u|\le T_n}|Y_n'(u)-(u,r+\eta_n)|\ge c\gamma_n\big)\le c\gamma_n.$$

Эта оценка, а также условие A2 позволяют сделать вывод, что максимум поля $Y_n(u)$ с вероятностью, не меньшей чем $1-c\gamma_n$, будет достигаться в той же точке, что и максимум поля $Y_{\eta_n}(u)$. Лемма доказана.

Лемма 2. *При выполнении условий теоремы 2 справедлива следующая оценка:*

$$\mathbf{P}\big(u_n^*\in H_x^{c\gamma_n}\big)\le c\gamma_n. \tag{4}$$

Доказательство. Достаточно показать, что случайная величина u_n^* имеет абсолютно непрерывное распределение с ограниченной плотностью. Пусть

u_0 — произвольная точка из $\mathbb{R}^m$, а $U(\varepsilon, u_0)$ — ε-окрестность u_0. Оценим вероятность $\mathbf{P}(u_n^* \in U(\varepsilon, u_0))$. Обозначим

$$S(u) = \{\langle t, y\rangle \in \mathbb{R}^d : t \in [0,1]^{d-1}, y = (uD(t,\theta_0), N(y(t,\theta_0),\theta_0))\}.$$

Событие $\{u_n^* \in U(\varepsilon, u_0)\}$ содержится в событии {*множество* $U(\varepsilon, u_0)$ *содержит* m *поверхностей разрыва поля* $Y(u)$}. Последнее событие совпадает с событием {m *атомов* Y_i *попало в множество* $\bigcup\limits_{u\in U(\varepsilon,u_0)} S(u)$}.

Вероятность того, что множество $U(\varepsilon, u_0)$ содержит m поверхностей разрыва, не превосходит величину

$$\sum_{i\ge m} e^{-A\ln n}\frac{(A\ln n)^i}{i!}C_i^m \mathbf{P}\Big(Y_1 \in \bigcup_{u\in U(\varepsilon,u_0)} S(u)\Big)^m \le \frac{(A\ln n)^m}{m!}\mathbf{P}\Big(Y_1 \in \bigcup_{u\in U(\varepsilon,u_0)} S(u)\Big)^m.$$

Оценим вероятность $\mathbf{P}\big(Y_1 \in \bigcup\limits_{u\in U(\varepsilon,u_0)} S(u)\big)$. Имеем

$$\mathbf{P}(Y_1 \in \bigcup_{u\in U(\varepsilon,u_0)} S(u)) \le \frac{1}{A\ln n}\int\limits_{[0,1]^{d-1}} (p+q)(\sup_{u\in U(\varepsilon,u_0)}(uD,N) - \inf_{u\in U(\varepsilon,u_0)}(uD,N))I(t,\theta_0)\,dt, \quad (5)$$

где мы для краткости обозначили

$$D = D(t,\theta_0), \quad N = N(y(t,\theta_0),\theta_0), \quad p = p(y(t,\theta_0),\theta_0), \quad q = q(y(t,\theta_0),\theta_0).$$

Проводя вычисления, аналогичные вычислениям в лемме 1 из [1], получаем, что

$$\sup_{u\in U(\varepsilon,u_0)}(uD,N) - \inf_{u\in U(\varepsilon,u_0)}(uD,N) = \sup_{u\in U(\varepsilon,0)}(uD,N) - \inf_{u\in U(\varepsilon,0)}(uD,N) = 2\varepsilon(\operatorname{sign}(D^*N), D^*N).$$

Подставляя полученное выражение в (5), приходим к следующей оценке:

$$\mathbf{P}(Y_1 \in \bigcup_{u\in U(\varepsilon,u_0)} S(u)) \le \frac{2\varepsilon}{A\ln n}\int\limits_{[0,1]^{d-1}} (p+q)(\operatorname{sign}(D^*N), D^*N)I(t,\theta_0)\,dt = 2\varepsilon/\ln n.$$

Таким образом, мы доказали, что

$$\mathbf{P}(u_n^* \in U(\varepsilon, u_0)) \le \frac{(2A)^m}{m!}\varepsilon^m.$$

Из полученной оценки следует, что распределение u_n^* абсолютно непрерывно с плотностью, не превосходящей $(2A)^m/m!$. Лемма доказана.

Лемма 3. *Пусть выполнены условия теоремы* 2. *Тогда*

$$\mathbf{P}(\sup_{|u|>T_n} Y_{\eta_n}(u) > 0) \le c/n.$$

Доказательство. В случае $d, m = 1$ утверждение леммы 3 доказано в [3]. При доказательстве общего случая мы воспользуемся методом, предложенным в [5, с. 349–355].

Мы уже отмечали, что для положительных ρ имеет место оценка $\mathbf{P}(|\eta_n| > \rho) \leq c/n$. Поэтому достаточно показать, что при некотором $\rho > 0$

$$\mathbf{P}(\sup_{|u|>T_n}(Y(u) + |u|\rho) > 0) \leq c/n.$$

Из представления поля $Y(u)$ получаем (здесь, как и в предыдущей лемме, мы для краткости опускаем аргументы у функций q, p, D, N)

$$\mathbf{E}_{\theta_0}\exp\{bY(u)\} = \exp\biggl\{(u,r)b + \int\limits_{[0,1]^{d-1}} p((q/p)^b - 1)(uD,N)^+ I(t,\theta_0)\,dt + \int\limits_{[0,1]^{d-1}} q((p/q)^b - 1)(uD,N)^- I(t,\theta_0)\,dt\biggr\}.$$

Обозначим через $\phi(u,b)$ функцию, стоящую под знаком экспоненты в правой части этого равенства. Тогда

$$\mathbf{E}_{\theta_0}\exp\{bY(u)\} = \exp\{|u|\phi(u/|u|, b)\}.$$

Легко видеть, что $\phi(u,0) = 0$ для всех u. Кроме того,

$$\phi'_b(u,0) = \int\limits_{[0,1]^{d-1}} \left(p - q + p\ln\frac{q}{p}\right)(uD,N)^+ I(t,\theta_0)\,dt + \int\limits_{[0,1]^{d-1}} \left(p - q - q\ln\frac{q}{p}\right)(uD,N)^- I(t,\theta_0)\,dt,$$

где штрих обозначает производную по параметру b. Применяя неравенство $\ln x \leq x - 1$, нетрудно проверить, что существует число $a > 0$, для которого $\phi'_b(u/|u|, 0) \leq -a$.

Далее, имеем

$$\phi''_{bb}(u,0) = \int\limits_{[0,1]^{d-1}} p(\ln(q/p))^2 (q/p)^b (uD,N)^+ I(t,\theta_0)\,dt + \int\limits_{[0,1]^{d-1}} q(\ln(q/p))^2 (q/p)^b (uD,N)^- I(t,\theta_0)\,dt.$$

Функция $\phi''_{bb}(u/|u|, b)$ ограничена равномерно по u, когда b принадлежит конечному интервалу. Следовательно, найдутся такие положительные b, b_1 и ρ, что

$$\mathbf{E}_{\theta_0}\exp\{bY(u) + |u|\rho\} \leq \exp\{-|u|b_1\}. \tag{6}$$

Пусть z — точка с целочисленными координатами в пространстве $\mathbb{R}^m$. Обозначим $A(z) = \{u = (u_1, \ldots, u_m) : z_j \leq u_j \leq z_j + 1,\ j = 1, \ldots, m\}$. Разобьем куб $A(z)$ на k^m равных кубов с длиной ребра k^{-1} (целое число k мы уточним позднее). Символом Δ_m будем обозначать любой такой куб. Рассмотрим также точки u из $A(z)$, координаты которых представимы в следующем виде:

$u_j = z_j + lk^{-2^{m-1}}$, $l = 0, \dots, k^{2^{m-1}}$. Нас интересуют те из этих точек, которые принадлежат $(m-1)$-мерным граням кубов Δ_m. Далее такие точки мы будем обозначать через $\{u(i)\}$.

Введем событие S, состоящее в том, что в $A(z)$ найдется по крайней мере одна точка u такая, что для любой точки $u(i)$ отрезок $[u, u(i)]$ содержит точки разрыва поля $Y(u)$. Для произвольной вещественнозначной функции $F(u)$, заданной в $\mathbb{R}^m$, определим величину

$$\omega_h^z(F) = \sup |F(u') - F(u'')|,$$

где супремум берется по всем точкам u' и u'' из $A(z)$, для которых $|u' - u''| \le h$ и отрезок $[u', u'']$ не содержит точек разрыва функции $F(u)$. Тогда при достаточно малом ρ имеем

$$\begin{aligned}\mathbf{P}(\sup_{u \in A(z)} (Y(u) + |u|\rho) > 0) \le \mathbf{P}(\max_i (Y(u(i)) + |u(i)|\rho) \\ > -c|z|/k) + \mathbf{P}(\omega_h^z(Y(u) + |u|\rho) > c|z|/k) + \mathbf{P}(S), \quad (7)\end{aligned}$$

где $h = k^{-1}$. Первое слагаемое в правой части (7) оценим с помощью соотношения (6):

$$\mathbf{P}(\max_i (Y(u(i)) + |u(i)|\rho) > -c|z|/k) \le c(k+1)^{m2^{m-1}} \exp\{-|z|b_2\} \quad (8)$$

с некоторыми константами c и $b_2 > 0$. Поскольку траектории поля $Y(u)$ представляют собой кусочно-линейные функции, для второго слагаемого в правой части (7) при подходящей c и достаточно малом ρ верна оценка

$$\mathbf{P}\big(\omega_h^z(Y(u) + |u|\rho) > c|z|/k\big) = 0. \quad (9)$$

Осталось оценить последнее слагаемое в (7). Покажем, что верна оценка

$$\mathbf{P}(S) \le c/k. \quad (10)$$

Введем несколько обозначений. Символом Δ_j будем обозначать j-мерный куб из $A(z)$ с ребром длины $k^{-2^{m-j}}$; ребра этого куба параллельны осям координат, а вершины совпадают с некоторыми из точек $\{u(i)\}$ (определение Δ_j при $j = m$ уже давалось выше). Например, Δ_j при $j = 1$ — это отрезок, при $j = 2$ — квадрат; в первом случае под ребром понимается сам отрезок Δ_j, во втором — сторона квадрата. Пусть Y^{Δ_j} — проекция поля Y на подпространство размерности j, содержащее Δ_j. Событие S^{Δ_j} состоит в том, что по крайней мере одна область непрерывности поля Y^{Δ_j} в кубе Δ_j не содержит ни одной из точек $u(i)$.

Докажем, что для любого $j \le m$ и любого Δ_j верна оценка

$$\mathbf{P}(S^{\Delta_j}) \le ck^{-(j+1)2^{m-j}}. \quad (11)$$

Доказательство проведем индукцией по j. При $j = 1$ из оценок, полученных в лемме 10, следует, что для любого Δ_1 выполняется соотношение

$$\mathbf{P}(S^{\Delta_1}) \le ck^{-2^m},$$

где c не зависит от Δ_1.

Предположим, что оценка (11) верна для некоторого $j \ge 1$. Покажем, что она верна и для $j+1$. Обозначим $h = k^{-2^{m-j-1}}$. Рассмотрим произвольный куб Δ_{j+1}. Каждую из его j-мерных граней можно представить в виде объединения кубов с длиной ребра h^2 (число таких кубов равно h^{-j}). Пусть Δ_j — один из таких кубов. По предположению индукции имеем

$$\mathbf{P}(S^{\Delta_j}) \le ch^{2j+2}.$$

Для того чтобы произошло событие $S^{\Delta_{j+1}}$, необходимо, чтобы куб Δ_{j+1} содержал более $j+1$ поверхностей разрыва поля $Y^{\Delta_{j+1}}$ (обозначим это событие через $S_1^{\Delta_{j+1}}$) или хотя бы для одного Δ_j из разбиения граней куба Δ_{j+1} произошло событие S^{Δ_j} (это событие мы обозначим через $S_2^{\Delta_{j+1}}$). Следовательно, для произвольного Δ_{j+1} имеем

$$\mathbf{P}(S^{\Delta_{j+1}}) \le \mathbf{P}\big(S_1^{\Delta_{j+1}}\big) + \mathbf{P}\big(S_2^{\Delta_{j+1}}\big).$$

Используя оценки из леммы 2, для первой вероятности правой части этого неравенства получаем

$$\mathbf{P}\big(S_1^{\Delta_{j+1}}\big) \le ch^{j+2}.$$

Для оценки второй вероятности воспользуемся предположением индукции:

$$\mathbf{P}\big(S_2^{\Delta_{j+1}}\big) \le ch^{j+2}.$$

Таким образом, $\mathbf{P}(S^{\Delta_{j+1}}) \le ch^{j+2}$, и оценка (11) верна при всех $j \le m$. В частности, при $j = m$ имеем

$$\mathbf{P}(S^{\Delta_m}) \le ck^{-(m+1)}.$$

Поскольку $A(z)$ можно представить в виде объединения k^m кубов вида Δ_m, то окончательно

$$\mathbf{P}(S) \le c/k.$$

Положим

$$k = \left[\exp\left\{\frac{b_2|z|}{1+m2^{m-1}}\right\}\right] + 1,$$

где квадратными скобками мы обозначаем целую часть числа. Тогда справедливо неравенство

$$\mathbf{P}\big(\sup_{u\in A(z)} (Y(u) + |u|\delta) > 0\big) \le c\exp\left\{-\frac{b_2}{1+m2^{m-1}}|z|\right\}.$$

В результате имеем

$$\mathbf{P}\big(\sup_{|u|>T_n} (Y(u) + |u|\delta) > 0\big) \le \sum_{|z|>c\ln n-1} \mathbf{P}\big(\sup_{u\in A(z)} (Y(u) + |u|\delta) > 0\big) \le c/n,$$

если c достаточно велико. Лемма 3 доказана.

Лемма 4. *Пусть выполнены условия теоремы* 2. *Тогда*

$$\mathbf{P}\big(\sup_{|u|>T_n} Y_n(u) > 0\big) \le c/n.$$

Доказательство этой леммы почти не отличается от доказательства леммы 3, поэтому мы не будем приводить его полностью. Достаточно лишь получить для поля $Y_n(u)$ оценки, аналогичные (6), (9) и (10). Оценка, подобная (10), по существу, следует из [6, с. 102] и получается точно так же, как и в лемме 3.

Оценим $\omega_h^z(Y_n)$. Для произвольных u и u' имеем

$$|Y_n(u) - Y_n(u')| \le \frac{|u-u'|}{n} \sup_u \Big| \sum_{i \le n} l'_\theta(X_i, \theta_0 + u/n) \Big|.$$

В силу условия 2 с вероятностью 1 выполняется неравенство

$$\omega_h^z(Y_n) \le ch.$$

Теперь оценим снизу расстояние Хеллингера

$$r_2^2(\theta, \theta') = \int\limits_{\mathbb{R}^d} (f^{1/2}(x,\theta) - f^{1/2}(x,\theta'))^2\, dx.$$

Если $|\theta - \theta'|$ достаточно мало, то

$$\begin{aligned} r_2^2(\theta,\theta') &\ge \int\limits_{\Omega^1_\theta \cap \Omega^2_{\theta'}} (f^{1/2}(x,\theta) - f^{1/2}(x,\theta'))^2\, dx + \int\limits_{\Omega^2_\theta \cap \Omega^1_{\theta'}} (f^{1/2}(x,\theta) - f^{1/2}(x,\theta'))^2\, dx \\ &\ge \frac{|\theta-\theta'|}{2} \int\limits_{[0,1]^{d-1}} |p(y(t,\theta),\theta) - q(y(t,\theta),\theta)| |(\theta, N(y(t,\theta),\theta) D^*(t,\theta))|\, dt. \end{aligned}$$

Из этой оценки и соответствующего результата в [5, с. 76] следует существование такого положительного b, что при любом u верна оценка

$$\mathbf{E}_{\theta_0} \exp\{(1/2) Y_n(u)\} \le \exp\{-b|u|\}.$$

Лемма 4 доказана.

Последний шаг в доказательстве теоремы 2 — получение оценки для величины $|\mathbf{P}(u_n^* < x) - \mathbf{P}(\hat{u} < x)|$.

Лемма 5. *При выполнении условий теоремы* 2 *верна следующая оценка:*

$$\sup_x |\mathbf{P}(u_n^* < x) - \mathbf{P}(\hat{u} < x)| \le c \ln^2 n / n.$$

Доказательство. В силу независимости $Y(u)$ и η_n имеем

$$\mathbf{P}(u_n^* < x) = \mathbf{E}_{\theta_0} h(\eta_n, x).$$

Следовательно,

$$|\mathbf{P}(u_n^* < x) - \mathbf{P}(\hat{u} < x)| = |\mathbf{E}_{\theta_0}(h(\eta_n, x) - h(0,x))|.$$

Используя формулу Тейлора и равенство $\mathbf{E}\eta_n = 0$, получаем

$$\begin{aligned} &|\mathbf{E}_{\theta_0}(h(\eta_n,x) - h(0,x))| \\ &\qquad \le \big| \mathbf{E}_{\theta_0}\big(h'_\delta(0,x), \eta_n\big) + \mathbf{E}_{\theta_0}\big(\eta_n h''_{\delta\delta}(\xi, x), \eta_n\big) \big| \le n^{-1}(c_1|x|^2 + c_2), \end{aligned}$$

где $h''_{\delta\delta}$ — матрица вторых смешанных производных функции $h(\delta, x)$ по всем координатам вектора δ, а ξ — некоторый случайный вектор из $\mathbb{R}^m$. В итоге получаем оценку

$$\sup_{|x| \le T_n} \big| \mathbf{P}(u_n^* < x) - \mathbf{P}(\hat{u} < x) \big| \le c \ln^2 n / n.$$

Аналогичная оценка при $|x| > T_n$ следует из леммы 3. Лемма 5 доказана.

Утверждение теоремы 2 вытекает из лемм 1–5.

§ 3. Проверка условия A3

Сначала проведем проверку условия A3 при $m = 1$. При этом будем предполагать выполненными условия работы [7]. Кроме того, в заключение этого параграфа приведем пример семейства распределений с многомерным параметром, которое также удовлетворяет условию A3.

Положим $Y_\delta(u) = Y(u) + u\delta$. Для $m = 1$ функцию $h(\delta, x)$ можно определить следующим соотношением:

$$h(\delta, x) = \mathbf{P}(\sup_{u<x} Y_\delta(u) > \sup_{u\geq x} Y_\delta(u)).$$

Пусть $x > 0$. В определении функции $h(\delta, x)$ мы можем заменить $Y_\delta(u)$ похожим процессом $Y^*_\delta(u)$, значения которого в соответствующие моменты скачков будут совпадать со значениями $Y_\delta(u)$, а линейный коэффициент не будет зависеть от параметра δ. Напомним, что процесс $Y_\delta(u)$ в случае $m = 1$ выглядит как

$$(r + \delta)u - \Pi^+(\lambda u) + \Pi^-(-\mu u), \tag{12}$$

где

$$r = \int\limits_{[0,1]^{d-1}} [p(y(t,\theta_0),\theta_0) - q(y(t,\theta_0),\theta_0)] I(t,\theta_0)\, dt, \tag{13}$$

$$\lambda = \int\limits_{[0,1]^{d-1}} p(y(t,\theta_0),\theta_0) I(t,\theta_0)\, dt, \tag{14}$$

$$\mu = \int\limits_{[0,1]^{d-1}} q(y(t,\theta_0),\theta_0) I(t,\theta_0)\, dt, \tag{15}$$

$\Pi^+(u)$ и $\Pi^-(u)$ — независимые обобщенные пуассоновские процессы, имеющие вид

$$\Pi^+(u) = \sum_{i=1}^{\pi^+(u)} \xi_i, \quad \Pi^-(u) = \sum_{i=1}^{\pi^-(u)} \zeta_i. \tag{16}$$

Здесь $\pi^+(u)$, $\pi^-(u)$ — независимые стандартные пуассоновские процессы, равные нулю на отрицательной полуоси (суммы в (16) полагаются равными нулю, когда $\pi^+,\pi^- = 0$), $\{\xi_i\}$, $\{\zeta_i\}$ — независимые последовательности независимых одинаково распределенных (в каждой последовательности) случайных величин. Они не зависят также и от процессов $\pi^+(u)$, $\pi^-(u)$ и имеют следующие распределения:

$$\mathscr{L}(\xi_i) = \mathscr{L}\left(\ln\frac{p(\xi,\theta_0)}{q(\xi,\theta_0)}\right), \tag{17}$$

$$\mathscr{L}(\zeta_i) = \mathscr{L}\left(\ln\frac{p(\zeta,\theta_0)}{q(\zeta,\theta_0)}\right), \tag{18}$$

где ξ и ζ — случайные векторы из $[0,1]^{d-1}$, распределения которых задаются соотношениями

$$\mathbf{P}(\xi \in B) = \lambda^{-1} \int\limits_B p(y(t,\theta_0),\theta_0) I(t,\theta_0)\, dt, \tag{19}$$

$$\mathbf{P}(\zeta \in B) = \mu^{-1} \int\limits_B q(y(t,\theta_0),\theta_0) I(t,\theta_0)\, dt \tag{20}$$

для любых борелевских множеств $B \subseteq [0,1]^{d-1}$. Так как функции $p(y(t,\theta_0),\theta_0)$ и $q(y(t,\theta_0),\theta_0)$ непрерывны, строго положительны и заданы на ограниченном множестве, случайные величины ξ_i и ζ_i ограничены с вероятностью 1.

Будем считать для определенности, что $r > 0$. Положим

$$Y^*_\delta(u) = ru - \Pi^+\left(\frac{r\lambda}{r+\delta}u\right) + \Pi^-\left(\frac{r\mu}{r+\delta}u\right).$$

Нам потребуется знание свойств функций распределения случайных величин $\sup\limits_{u\ge 0} Y^*_\delta(u)$ и $\sup\limits_{u<0} Y^*_\delta(u)$. Обозначим эти функции распределения как $F_+(\delta,x)$ и $F_-(\delta,x)$ соответственно. Рассмотрим первый из этих супремумов. Обозначим

$$S_1 = r\tau_1, \quad S_j = r\tau_1 + \sum_{i=2}^{j}(r\tau_i - \xi_{i-1}), \quad j \ge 2,$$

где τ_i — время, прошедшее между $(i-1)$-м и i-м скачками процесса $\pi^+(\frac{r\lambda}{r+\delta}u)$ (τ_1 — момент 1-го скачка); случайные величины $r\tau_i$ имеют экспоненциальное распределение с параметром $\lambda/(r+\delta)$ и не зависят от последовательности $\{\xi_i\}$. Отметим, что распределение $\sup\limits_{u\ge 0} Y^*_\delta(u)$ совпадает с распределением $\max\limits_{j\ge 1}\{S_j\}$.

Рассмотрим еще одну последовательность частичных сумм:

$$\widetilde{S}_0 = 0, \quad \widetilde{S}_j = \sum_{i=1}^{j}(r\tau_i - \xi_{i-1}), \quad j \ge 1.$$

Функцию распределения $\max\limits_{j\ge 0}\{\widetilde{S}_j\}$ обозначим через $\widetilde{F}_+(\delta,x)$. С функцией $F_+(\delta,x)$ она связана соотношением

$$\begin{aligned} F_+(\delta,x) = (1 - e^{-\frac{\lambda}{r+\delta}x})(\widetilde{F}_+(\delta,0+0) - \widetilde{F}_+(\delta,0)) \\ + \frac{\lambda}{r+\delta}\int_0^\infty (\widetilde{F}_+(\delta,x-y) - \widetilde{F}_+(\delta,0+0))e^{-\frac{\lambda}{r+\delta}y}\,dy. \end{aligned} \tag{21}$$

Отметим, что функция $F_+(\delta,x)$ абсолютно непрерывна, а распределение с функцией $\widetilde{F}_+(\delta,x)$ имеет атом в нуле. Согласно [8, с. 686] характеристическая функция, соответствующая $\widetilde{F}_+(\delta,x)$, имеет вид

$$\widetilde{\Phi}_\delta(t) = \exp\left\{\sum_{k=1}^\infty \frac{1}{k}\int_0^\infty (e^{itz}-1)g_\delta^{k*}(z)\,dz\right\},$$

где g_δ — плотность распределения случайной величины $(r\tau_1 - \xi_1)$, g_δ^{k*} — k-кратная свертка плотности с самой собой.

Теперь получим представление для функции $\widetilde{F}_+(\delta,x)$. Обозначим

$$a_\delta = \lim_{t\to\infty}\widetilde{\Phi}_\delta(t).$$

Отметим, что

$$a_\delta = \widetilde{F}(\delta,0+0) = \exp\left\{-\sum_{k=1}^\infty \frac{1}{k}\mathbf{P}(\widetilde{S}_k > 0)\right\}.$$

Если функция $t^{-1}(\widetilde{\Phi}_\delta(t) - a_\delta)$ абсолютно интегрируема при всех δ, то верно следующее представление для функции $\widetilde{F}_+(\delta, x)$:

$$\widetilde{F}_+(\delta, x) = a_\delta + \frac{1}{2\pi} \int\limits_{\mathbb{R}} \frac{1 - e^{-ixt}}{it} (\widetilde{\Phi}_\delta(t) - a_\delta)\, dt, \quad x > 0. \tag{22}$$

Для того чтобы доказать требуемую интегрируемость, достаточно убедиться в существовании такого c_δ, для которого при любых t и δ верна оценка

$$\left| \sum_{k=1}^{\infty} \frac{1}{k} \int\limits_0^{\infty} e^{itz} g_\delta^{k*}(z)\, dz \right| \le \frac{c_\delta}{t}. \tag{23}$$

Получить такую оценку можно с помощью формулы интегрирования по частям (ее применение законно в силу гладкости $g_\delta^{k*}(z)$ по z):

$$\int\limits_0^{\infty} e^{itz} g_\delta^{k*}(z)\, dz = -\frac{1}{it} \left(g_\delta^{k*}(0) - \int\limits_0^{\infty} e^{itz} \frac{\partial}{\partial z} (g_\delta^{k*}(z))\, dz \right).$$

Для доказательства сходимости ряда $\sum\limits_{k=1}^{\infty} \frac{1}{k} g_\delta^{k*}(0)$ можно воспользоваться, например, асимптотическими разложениями в локальной предельной теореме (см. [9]). Оценим $\frac{\partial}{\partial z} g_\delta^{k*}(z)$. Имеем

$$g_\delta^{k*}(z) = \int\limits_{-\infty}^{z} e^{-\frac{\lambda}{r+\delta}(z-y)} \tilde{g}^{k*}(y)\, dy,$$

где $\tilde{g}^{k*}$ — плотность распределения случайной величины $\widetilde{S}_{k-1} - \xi_{i-1}$. Отсюда нетрудно получить соотношение

$$\frac{\partial}{\partial z} g_\delta^{k*}(z) = \tilde{g}^{k*}(z) - \frac{\lambda}{r+\delta} g_\delta^{k*}(z),$$

из которого следует сходимость ряда

$$\sum_{k=1}^{\infty} \frac{1}{k} \int\limits_0^{\infty} e^{itz} \frac{\partial}{\partial z} (g_\delta^{k*}(z))\, dz.$$

Используя полученное представление для $\widetilde{F}_+$, докажем, что эта функция распределения дважды непрерывно дифференцируема по параметру δ и ее первая и вторая производные ограничены равномерно по x и δ при достаточно малом δ.

Сначала докажем дифференцируемость a_δ. Применяя формулу обращения для представления g_δ^{k*} при $k \ge 2$ (в этом случае характеристическая функция распределения с плотностью g_δ^{k*} является абсолютно интегрируемой), можно установить соотношения

$$\frac{d}{d\delta} \mathbf{P}(\widetilde{S}_k > 0) = \frac{k}{r+\delta} (-\mathbf{P}(\widetilde{S}_k > 0) + \mathbf{P}(\widetilde{S}_k + r\chi_{k+1} > 0)),$$

$$\frac{d^2}{d\delta^2} \mathbf{P}(\widetilde{S}_k > 0) = \frac{k(k+1)}{(r+\delta)^2} (\mathbf{P}(\widetilde{S}_k > 0) - 2\mathbf{P}(\widetilde{S}_k + r\chi_{k+1} > 0) + \mathbf{P}(\widetilde{S}_k + r\chi_{k+1} + r\chi_{k+2} > 0)),$$

где $k \geq 2$. Используя экспоненциальное неравенство Чебышева для оценки вероятностей, входящих в вышеприведенные соотношения, нетрудно установить, что все члены двух следующих рядов:

$$\sum_{k=1}^{\infty} \frac{1}{k} \frac{d^j}{d\delta^j} \mathbf{P}(\widetilde{S}_k > 0), \quad j = 1, 2,$$

допускают равномерные в некоторой окрестности каждой точки δ суммируемые мажоранты, а значит, a_δ можно дважды дифференцировать по δ. При этом первая и вторая производные будут ограничены равномерно по δ, когда δ принадлежит некоторой окрестности нуля.

Далее, нам нужно установить существование производной второго порядка по параметру δ функции

$$\int\limits_{\mathbb{R}} \frac{1 - e^{-ixt}}{it} (\widetilde{\Phi}_\delta(t) - a_\delta)\, dt.$$

Для этого достаточно показать, что функция

$$\sum_{k=1}^{\infty} \frac{1}{k} \int\limits_0^{\infty} e^{itz} g_\delta^{k*}(z)\, dz \tag{24}$$

дважды непрерывно дифференцируема по δ и что существует такая абсолютная постоянная c, что для любых t и достаточно малых δ верны оценки

$$\left| \frac{\partial^j}{\partial \delta^j} \sum_{k=1}^{\infty} \frac{1}{k} \int\limits_0^{\infty} e^{itz} g_\delta^{k*}(z)\, dz \right| \leq \frac{c}{t}, \quad j = 1, 2. \tag{25}$$

Снова применяя формулу обращения для представления g_δ^{k*}, получим

$$\frac{\partial}{\partial \delta} g_\delta^{k*}(z) = \frac{k}{r+\delta} \big((g_\delta^{k*} * E_\delta)(z) - g_\delta^{k*}(z) \big),$$

$$\frac{\partial^2}{\partial \delta^2} g_\delta^{k*}(z) = \frac{k(k+1)}{(r+\delta)^2} \big(g_\delta^{k*}(z) - 2(g_\delta^{k*} * E_\delta)(z) + \big(g_\delta^{k*} * E_\delta^{2*}\big)(z) \big),$$

где $E_\delta(z)$ — плотность экспоненциального распределения с параметром $\lambda/(r+\delta)$, знак $*$, как и ранее, используется для обозначения свертки плотностей. Из этих соотношений следует, что производные $\frac{\partial}{\partial \delta} g_\delta^{k*}(z)$ и $\frac{\partial^2}{\partial \delta^2} g_\delta^{k*}(z)$ имеют равномерные по δ (в достаточно малой окрестности нуля) абсолютно интегрируемые мажоранты. Используя экспоненциальное неравенство Чебышева, можно показать, что члены рядов

$$\sum_{k=1}^{\infty} \frac{1}{k} \int\limits_0^{\infty} e^{itz} \frac{\partial^j}{\partial \delta^j} g_\delta^{k*}(z)\, dz, \quad j = 1, 2,$$

также допускают равномерные по δ суммируемые мажоранты. Следовательно, функцию (24) можно дважды непрерывно дифференцировать по δ. Оценки (25) получаются точно таким же способом, каким была получена оценка (23).

Итак, мы установили, что функция распределения $\widetilde{F}_+(\delta, x)$ дважды непрерывно дифференцируема по параметру δ и соответствующие производные ограничены равномерно по δ, x. Из соотношения (21) следует, что то же самое верно и для функции распределения $F_+(\delta, x)$. Доказательство для функции $F_-(\delta, x)$ ничем не отличается от доказательства для $\widetilde{F}_+(\delta, x)$.

Далее, поскольку $Y^*_\delta(u)$, $u \geq 0$, — однородный процесс с независимыми приращениями, справедливо соотношение (см. [3])

$$h(\delta, x) = \mathbf{P}(\sup_{u<x} Y^*_\delta(u) - Y^*_\delta(x) > \sup_{u\geq x} Y^*_\delta(u) - Y^*_\delta(x))$$
$$= 1 - \int\limits_0^\infty \omega(\delta, x, z) \frac{\partial}{\partial z}(F_+(\delta, z))\, dz,$$

где $\omega(\delta, x, z) = \mathbf{P}\big(\sup\limits_{u<x} Y^*_\delta(u) < z + Y^*_\delta(x)\big)$. Найдем представление для $\omega(\delta, x, z)$. Имеем

$$\omega(\delta, x, z) = \sum_k \mathbf{P}\left(\sup_{u<x} Y^*_\delta(u) < z + rx - \sum_{j=1}^k \xi_j; \pi^+(\frac{r\lambda}{r+\delta}x) = k\right).$$

Каждая из вероятностей, стоящих под знаком суммы, представима в виде

$$\int\limits_{\mathbb{R}^k} \mathbf{P}\left(\sup_{u<x} Y^*_\delta(u) < z + rx - \sum_{j=1}^k y_j; \pi^+\left(\frac{r\lambda}{r+\delta}x\right) = k \mid \right.$$
$$\left. \xi_1 \in dy_1; \ldots; \xi_k \in dy_k \right) dF_{\xi_1}(y_1) \ldots dF_{\xi_k}(y_k),$$

где F_{ξ_j}, $j = 1, \ldots, k$, — функции распределения случайных величин ξ_j. Вероятность, стоящую под знаком интеграла, можно записать как

$$F_-\left(\delta, z + rx - \sum_{j=1}^k y_j\right) G_k\left(\delta, x, z + rx - \sum_{j=1}^k y_j\right),$$

где, напомним,

$$F_-\left(\delta, z + rx - \sum_{j=1}^k y_j\right) = \mathbf{P}\left(\sup_{u<0} Y^*_\delta(u) < z + rx - \sum_{j=1}^k y_j\right),$$

$$G_k\left(\delta, x, z + rx - \sum_{j=1}^k y_j\right)$$
$$= \mathbf{P}\left(\sup_{0\leq u<x} Y^*_\delta(u) < z + rx - \sum_{j=1}^k y_j; \pi^+\left(\frac{r\lambda}{r+\delta}x\right) = k \mid \xi_1 \in dy_1; \ldots; \xi_k \in dy_k\right).$$

Найдем представление для функции G_k. Пусть T_j — момент j-го скачка процесса $\pi^+(\frac{r\lambda}{r+\delta}u)$. Тогда

$$G_k(\delta, x, v) = \mathbf{P}\left(rT_1 < v; rT_2 < v + y_1; \ldots; rT_k < v + \sum_{j=1}^{k-1} y_j; rT_k < rx; rT_k > rx\right).$$

Положим

$$x_1 = v,\ x_2 = v + y_1, \ldots, x_k = \min\left\{rx, v + \sum_{j=1}^{k-1} y_j\right\}.$$

Имеем

$$G_k(\delta,x,v)=\left(\frac{\lambda}{r+\delta}\right)^k \exp\left\{-\frac{\lambda r}{r+\delta}x\right\} \times \int\limits_0^{x_1} \mathbf{I}(u_1\le x_2)du_1 \int\limits_{u_1}^{x_2} \mathbf{I}(u_2\le x_3)du_2 \cdots \int\limits_{u_{k-1}}^{x_k} du_k,$$

где $\mathbf{I}$ — индикатор события.

Для того чтобы условие A3 было выполнено, функция $\omega(\delta,x,z)$ должна быть дважды непрерывно дифференцируемой по параметру δ и ее первая и вторая производные должны удовлетворять следующим неравенствам (штрих означает частную производную по параметру δ):

$$|\omega'(\delta,x,z)|\le C_1x+C_0,\quad |\omega''(\delta,x,z)|\le C_1x^2+C_0.$$

Проверим, что это так. Выше мы установили, что

$$\omega(\delta,x,z)=\sum_k\int\limits_{\mathbb{R}^k} F_-\left(\delta,z+rx-\sum_{j=1}^k y_j\right) \times G_k\left(\delta,x,z+rx-\sum_{j=1}^k y_j\right)dF_{\xi_1}(y_1)\dots dF_{\xi_k}(y_k).$$

Производные функции ω можно схематично записать как

$$\omega'=\sum\int\big(F_-'G_k+F_-G_k'\big),\quad \omega''=\sum\int\big(F_-''G_k+2F_-'G_k'+F_-G_k''\big).$$

Мы уже упоминали, что первая и вторая производные функции $F_-(\delta,y)$ по параметру δ ограничены равномерно по δ и y. Путем дифференцирования функций G_k нетрудно установить следующие соотношения:

$$G_k'(\delta,x,y)=(C_3(\delta)x+C_4(\delta))G_k(\delta,x,y),$$

$$G_k''(\delta,x,y)=(C_5(\delta)x^2+C_6(\delta)x+C_7(\delta))G_k(\delta,x,y),$$

где функции $C_j(\delta)$ ограничены при достаточно малых абсолютных значениях δ. Для того чтобы закончить проверку условия A3, остается лишь заметить, что

$$\sum_k\int\limits_{\mathbb{R}^k} G_k\left(\delta,x,z+rx-\sum_{j=1}^k y_j\right)dF_{\xi_1}(y_1)\dots dF_{\xi_k}(y_k)\le 1.$$

Теперь мы покажем на конкретном примере семейства плотностей, как условие A3 проверяется в случае многомерного параметра.

Пусть $m=d>1$ и поверхность разрыва плотности $f(x,\theta)$ (т. е. многообразие K_θ) представляет собой параллелепипед, определяемый следующими соотношениями:

$$K_\theta=\bigcup_{j=1}^d K_\theta^j,$$

$$K_\theta^j=\{(x_1,\dots,x_d)\in\mathbb{R}^d: x_j\in\{-b_j-\theta_j,b_j+\theta_j\}; x_i\in[-b_i-\theta_i,b_i+\theta_i], i\ne j\},$$

где θ_j — j-я координата d-мерного параметра θ, а $\{b_j\}$ — набор положительных чисел.

Введенное многообразие не является гладким, однако это несущественно. Гладкость многообразия K_θ была необходима в [1] лишь для того, чтобы корректно определить ортопроекцию на K_θ. В свою очередь, ортопроекция использовалась в [1] для задания случайного поля $\Pi_n(u)$ и в доказательстве леммы 8 для задания плотности распределения атомов случайной пуассоновской меры Λ_2. Определим ортопроекцию на введенное выше многообразие K_θ следующим образом:

$$\operatorname{pr} x = \{y \in K_\theta : \|x-y\|_2 = \inf_{z \in K_\theta} \|x-z\|_2\},$$

где $\|\cdot\|_2$ — евклидова норма в $\mathbb{R}^d$. Введенная ортопроекция неоднозначна для точек, равноудаленных от граней параллелепипеда K_θ. Однако, поскольку множество таких точек в $\mathbb{R}^d$ имеет нулевую лебегову меру, неоднозначность не мешает нам корректно задать поле $\Pi_n(u)$ и случайную пуассоновскую меру Λ_2. Так что теорема 1 остается в силе и для модифицированного определения ортопроекции.

При этих предположениях поле $Y(u)$ будет иметь вид

$$Y(u) = \sum_{j=1}^{d} Y(j, u_j) \equiv \sum_{j=1}^{d} \big(u_j r_j - \Pi_j^+(\lambda_j u_j) + \Pi_j^-(-\mu_j u_j)\big). \tag{26}$$

Здесь u_j — координаты параметра u,

$$r_j = e_j \int\limits_{K_{\theta_0}^j} [p(y,\theta_0) - q(y,\theta_0)]\, dy,$$

e_j — единичный орт системы координат в $\mathbb{R}^d$, $\Pi_j^+(u_j)$ и $\Pi_j^-(u_j)$ — независимые обобщенные пуассоновские процессы, имеющие вид

$$\Pi_j^+(u_j) = \sum_{i=1}^{\pi_j^+(u_j)} \xi_i^j, \quad \Pi_j^-(u_j) = \sum_{i=1}^{\pi_j^-(u_j)} \zeta_i^j,$$

$\pi_j^+(u_j)$, $\pi_j^-(u_j)$ — независимые стандартные пуассоновские процессы, равные нулю на отрицательной полуоси,

$$\lambda_j = \int\limits_{K_{\theta_0}^j} p(y,\theta_0)\, dy, \quad \mu_j = \int\limits_{K_{\theta_0}^j} q(y,\theta_0)\, dy,$$

$\{\xi_i^j\}_{i=1}^\infty$, $\{\zeta_i^j\}_{i=1}^\infty$, $j = 1, \dots, d$, — независимые последовательности независимых случайных величин. Они не зависят также и от процессов $\pi_j^+(u_j)$, $\pi_j^-(u_j)$ и имеют распределения

$$\mathscr{L}(\xi_i^j) = \mathscr{L}\left(\ln \frac{p(\xi^j,\theta_0)}{q(\xi^j,\theta_0)}\right), \quad \mathscr{L}(\zeta_i^j) = \mathscr{L}\left(\ln \frac{p(\zeta^j,\theta_0)}{q(\zeta^j,\theta_0)}\right),$$

где ξ^j и ζ^j — случайные векторы, заданные на $K_{\theta_0}^j$, с распределениями

$$\mathbf{P}(\xi^j \in B) = \lambda_j^{-1} \int\limits_B p(y,\theta_0)\, dy, \quad \mathbf{P}(\zeta^j \in B) = \mu_j^{-1} \int\limits_B q(y,\theta_0)\, dy$$

для любых борелевских множеств $B \subseteq K^j_{\theta_0}$.

Соотношения, определяющие одномерные процессы $Y(j, u_j)$, вполне аналогичны соотношениям (12)–(20), с помощью которых мы задавали процесс $Y(u)$ в случае одномерного параметра.

Из представления (26) найдем, что функцию $h(\delta, x)$ можно записать в следующем виде:

$$h(\delta, x) = \prod_{j \le d} \mathbf{P}(\sup_{u_j < x_j} (Y(j, u_j) + u_j \delta_j) > \sup_{u_j \ge x_j} (Y(j, u_j) + u_j \delta_j)),$$

где δ_j — координаты δ. Такое представление функции $h(\delta, x)$ позволяет свести задачу о проверке условия A3 к случаю одномерного параметра, который уже рассмотрен выше.

ЛИТЕРАТУРА

1. *Борисов И. С., Миронов Д. В.* Асимптотическое представление отношения правдоподобия для нерегулярных семейств распределений в многомерном случае // Сиб. мат. журн. 2001. Т. 42, № 2. С. 275–208.
2. *Борисов И. С., Боровков А. А.* Аппроксимация второго порядка в принципе инвариантности Донскера — Прохорова // Теория вероятностей и ее применения. 1986. Т. 31, № 2. С. 179–202.
3. *Мосягин В. Е.* Оценка скорости сходимости распределений нормированных оценок максимального правдоподобия в случае разрывной плотности // Сиб. мат. журн. 1996. Т. 37, № 4. С. 895–903.
4. *Korostelev A. B., Tsybakov A. B.* Minimax theory of image reconstruction. New York: Springer, 1993. (Lecture Notes in Stat.; 82).
5. *Ибрагимов И. А., Хасьминский Р. З.* Асимптотическая теория оценивания. М.: Наука, 1979.
6. *Ермаков М. С.* Асимптотическое поведение статистических оценок параметров многомерной разрывной плотности // Зап. науч. семинаров ЛОМИ. 1977. Т. 74. С. 88–107.
7. *Борисов И. С., Миронов Д. В.* Асимптотическое представление отношения правдоподобия для многомерных выборок с разрывными плотностями // Теория вероятностей и ее применения. 2000. Т. 45, № 2. С. 345–356.
8. *Феллер В.* Введение в теорию вероятностей и ее приложения. М.: Мир, 1984.
9. *Петров В. В.* Суммы независимых случайных величин. М.: Наука, 1972.

Статья поступила 26 ноября 2002 г.

Борисов Игорь Семенович
Институт математики им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4, Новосибирск 630090
sibam@math.nsc.ru

Миронов Дмитрий Витальевич
Новосибирский гос. университет, механико-математический факультет,
ул. Пирогова, 2, Новосибирск 630090
d.mironov@s7.ru