УДК 519.21

Э.А. НАДАРАЯ, Г.А. СОХАДЗЕ, А.Д. ШАТАШВИЛИ

О СТАТИСТИЧЕСКОМ ОЦЕНИВАНИИ ЛОГАРИФМИЧЕСКОЙ ПРОИЗВОДНОЙ МЕРЫ В ГИЛЬБЕРТОВОМ ПРОСТРАНСТВЕ

Ключевые слова: *логарифмическая производная меры, независимые наблюдения, статистический анализ, непараметрические оценки*.

Исследования в теории вероятностных распределений [1, 2], проводимые в настоящее время, показали, что наряду с моментными функционалами особую важность приобретает исследование логарифмической производной меры, которая фигурирует во многих вопросах и особенно в задачах, где исследуются гладкости мер. В связи с этим возникает необходимость уметь оценивать логарифмическую производную меры по независимым наблюдениям. Эта задача является важной еще и потому, что проблема оценивания, как известно из результатов работы [3], является сложной ввиду отсутствия прямого аналога теоремы Гливенко–Кантелли в бесконечномерном линейном пространстве.

В настоящей статье частично разрешена поставленная выше задача как для конечномерного, так и для бесконечномерного пространства. При этом применяется подход так называемого непараметрического оценивания, развитого в работе [4]. Полученные результаты авторы намереваются применять далее для оценки функции правдоподобия в бесконечномерном пространстве.

1. Рассмотрим одномерный вариант задачи. Всюду ниже $\{\Omega, \Im, P\}$ — некоторое фиксированное полное вероятностное пространство, $X_1, X_2, \ldots, X_n$ — выборка независимых, одинакого распределенных случайных величин. Предположим, что неизвестное распределение имеет всюду положительную гладкую плотность $p(x) \in C^1(R)$. Требуется по данной выборке оценить логарифмическую производную $l(x)$ распределения $\mu(A) = \int_A p(x)dx$, $A \in B(R)$. В этом случае логарифмическая производная существует и равна $p'(x)/p(x)$. Критерием оценки будем считать равномерную норму. Поэтому необходимо построить статистику $T_n(x, X_1, X_2, \ldots, X_n)$ такую, чтобы $\lim_{n\to\infty} \sup_{x\in R} |T_n(x, X_1, X_2, \ldots, X_n) - l(x)| = 0 \pmod P$. Как известно [4–6], для оценки непосредственно самой плотности распределения эффективными являются так называемые ядерные оценки. Так, пусть $K(x)$ — всюду положительная, равномерно непрерывная на R четная функция, при этом $\int_R K(x)dx = 1$. Для таких функций будем считать $K \in CL$. Оценку плотности $p(x)$ представим в виде $\hat{p}_n(x) = \frac{\lambda_n}{n}\sum_{i=1}^{n} K(\lambda_n(x - X_i))$ с последовательностью чисел $\lambda_n \to \infty$ и $\lambda_n \cdot n^{-1} \to 0$. По теореме Э. Надарая из работы [4] $\lim_{n\to\infty} \sup_{x\in R} |\hat{p}_n - p(x)| = 0 \pmod P$. Естественно, что для оценки производной плотности применяется (см. [4, 6–8]) оценка $\hat{p}'_n(x) = \frac{\lambda_n^2}{n}\sum_{i=1}^{n} K'(\lambda_n(x - X_i))$. Эффективность этой оценки является следствием следующего утверждения.

Лемма 1. Пусть $\{Y, \Im, \mu\}$ — измеримое пространство с мерой; $\{B, \Xi, \|\cdot\|_B\}$ — сепарабельное, вещественное банахово пространство с борелевской σ-алгеброй Ξ и нормой $\|\cdot\|_B$; A — замкнутый (не обязательно линейный и не обязательно ограниченный) оператор в B с плотной областью определения $D(A) \subset B$. Пусть дано семейство измеримых отображений $F_{y_1, y_2, \ldots, y_k}(x) \colon Y^k \times B \to B$, $y_k \in Y$, $k = 1, 2, \ldots$, $x \in B$. Тогда, если семейство $\{F_{y_1, y_2, \ldots, y_k}(x)\}_{n=1}^{\infty}$ сходится по норме $\|\cdot\|_B$ к функции $F(x)$ при $n \to \infty$ и $\exists\, n \in N$, $\exists\, y_1, y_2, \ldots, y_n \in Y$, при которых $F_{y_1, y_2, \ldots, y_n} \in D(A)$, $F \in D(A)$, то семейство $\{AF_{y_1, y_2, \ldots, y_n}(x)\}_{n=1}^{\infty}$ сходится по норме $\|\cdot\|_B$ к $AF(x)$.

Доказательство вытекает из определения замкнутого оператора.

Чтобы применить это утверждение к рассматриваемому случаю, следует учитывать, что оператор A, действующий в $C(R)$ по правилу $(Ap)(x) = \dfrac{p'(x)}{p(x)} = l(x)$, определен на плотном подмножестве $C(R)$ и является замкнутым. Кроме того, в качестве измеримого пространства будем использовать $\{R, \Im, p(x)dx\}$. Таким образом, с учетом оценки плотности из работы [5] и утверждения данной леммы можно убедиться в справедливости следующей теоремы.

Теорема 1. Пусть $K(x) \in CL \bigcap D(A)$, $\lambda_n \to \infty$, $\dfrac{l_n^2 \ln n}{n} \to 0$ при $n \to \infty$. Тогда последовательность

$$\hat{l}_n(x) = \frac{\lambda_n \sum\limits_{i=1}^{n} K'(\lambda_n (x - X_i))}{\sum\limits_{i=1}^{n} K(\lambda_n (x - X_i))}$$

сходится в $C(R)$ к $l(x)$ с вероятностью единица.

Замечание. Общность утверждения леммы 1 обеспечивает справедливость теоремы 1 и в случае, когда сходимость рассматривается относительно пространства $L_p(R)$, $p \ge 1$.

Это замечание относится и к утверждению пространства R^n (см. теорему 2).

2. Рассмотрим m-мерный случай. Предположим, что имеется выборка $X_j = (X_j^1, X_j^2, \ldots, X_j^m)$, $j = 1, 2, \ldots, n$, независимых и одинаково распределенных случайных векторов. Для оценки неизвестной плотности $p(x)$ применяется статистика из [5], когда $\hat{p}_n(x) = \dfrac{\lambda_n^m}{n} \sum\limits_{i=1}^{n} K(\lambda_n (x - X_i))$, где $K(x) = \prod\limits_{j=1}^{m} K_j(x_j)$, $x = (x_1, x_2, \ldots, x_m)$ и $K_j \in CL$, $j = 1, 2, \ldots, m$. Как известно, логарифмическая производная представляет вектор с компонентами $\dfrac{1}{p(x)} \dfrac{\partial p(x)}{\partial x_i}$. Следовательно, необходимо оценить выражение $l(x) = \dfrac{1}{p(x)} \operatorname{grad} p(x)$.

В качестве статистики рассмотрим функцию

$$\hat{l}_n(x) = \frac{\sum\limits_{i=1}^{n} \operatorname{grad} \prod\limits_{j=1}^{m} K_j(\lambda_n (x_i - X_i^j))}{\sum\limits_{i=1}^{n} \prod\limits_{j=1}^{m} K_j(\lambda_n (x_j - X_i^j))}. \tag{1}$$

Для применения леммы 1 в качестве измеримого пространства возьмем пространство $\{R^m, \Im^m, p(x)dx\}$. Оператор A, действующий в $C(R^n)$ по правилу $(Ap)(x) = \dfrac{\text{grad } p(x)}{p(x)} = l(x)$, определен на плотном подмножестве $D(A) \subset C(R^n)$ и является замкнутым. Применяя утверждение леммы 1 с учетом соответствующего результата из [5], легко убедиться в справедливости следующей теоремы.

Теорема 2. Пусть заданы $K_j(x) \in CL \bigcap D(A)$, $j = 1, 2, \ldots$, и последовательность чисел λ_n таких, что $\lambda_n \to \infty$, $\dfrac{\lambda_n^2 \ln n}{n} \to 0$ при $n \to \infty$. Тогда (1) сходится в $C(R^n)$ к $l(x)$ с вероятностью единица.

3. Перейдем к изложению основного результата статьи, состоящего в изучении оценок логарифмической производной в гильбертовом пространстве.

Пусть H — сепарабельное, вещественное гильбертово пространство со скалярным произведением $(\cdot, \cdot)_H$ и нормой $||\cdot||_H$; $\mathcal{B}$ — борелевская σ-алгебра в H. Пусть далее для некоторого случайного элемента $X \in H$ имеется выборка $X_1, X_2, \ldots, X_n$ независимых, одинаково распределенных случайных элементов со значениями в H. Обозначим μ распределение (меру), соответствующее элементу X. Допустим, что μ обладает логарифмической производной вдоль постоянного направления $a \in H$. Необходимо найти оценку неизвестной логарифмической производной $l(x, a)$ по выборке $X_1, X_2, \ldots, X_n$.

Напомним, что мера μ обладает логарифмической производной вдоль постоянного направления $a \in H$ в том случае, если она дифференцируема по направлению a и $\mu'_a << \mu$. Тогда $l(x, a) = \dfrac{d\mu'_a}{d\mu}(x)$. При этом предполагается, что мера μ обладает производной вдоль a, если существует такая счетно-аддитивная мера μ'_a, что для любой ограниченной и ограниченно дифференцируемой функции $f(x): H \to R$ имеет место соотношение

$$\int_H (f'(x), a)_H \, \mu(dx) = -\int_H f(x) \mu'_a(dx),$$

которое называется формулой интегрирования по частям.

В случае существования логарифмической производной вдоль a справедлива формула

$$\int_H (f'(x), a)_H \, \mu(dx) = -\int_H f(x) l(x, a) \mu(dx),$$

которая иногда принимается для определения логарифмической производной вдоль a.

Пусть $\{L_m\}$ — такая возрастающая последовательность конечномерных подпространств пространства H, что их объединение $\bigcup_{m=1}^{\infty} L_m$ плотно в H. Обозначим P_m конечномерный проектор пространства H на L_m. Пусть $\mu_m = \mu \circ P_m^{-1}$, $a_m = P_m a$, $x_m = P_m x$, $l_m^{a_m}(x_m) = P_m l(P_m x, P_m a)$. Тогда $l_m^{a_m}(x_m)$ является логарифмической производной меры μ_m в пространстве L_m, и при условии, что $p_m(x_m)$ — всюду положительная и дифференцируемая плотность меры μ_m, имеем

$$l_m^{a_m}(x_m) = \frac{(\text{grad } p_m(x_m), a_m)_{L_m}}{p_m(x_m)}.$$

Такую логарифмическую производную можно оценивать по выборке $P_m X_1 = X_1^m$, $P_m X_2 = X_2^m, \ldots, P_m X_n = X_n^m$.

Для каждого m по выборке $X_1^m, X_2^m, \ldots, X_n^m$ построим эффективные оценки $\hat{l}_n^m = \hat{l}_m^{a_m}(x_m)_n$ для $l_m^{a_m}(x_m)$ и покажем их сходимость к $l(x,a)$ при n, $m \to \infty$. В этом случае $\hat{l}_n^m$ можно считать оценкой $l(x,a)$.

Для упрощения записи рассмотрим одну функцию $K(x) \in CL$. В условиях теоремы 2 при фиксированном m статистика

$$\hat{l}_n^m(x_m) = \frac{\lambda_n \sum_{i=1}^{n} \sum_{s=1}^{m} a_n^s K'(\lambda_n (x_m^s - X_{is}^m)) \prod_{\substack{j=1 \\ j \neq s}}^{m} K(\lambda_n (x_m^j - X_{ij}^m))}{\sum_{i=1}^{n} \prod_{j=1}^{m} K(\lambda_n (x_m^j - X_{ij}^m))} \quad (2)$$

сходится равномерно при $n \to \infty$ к $l_m^{a_m}(x_m)$ с вероятностью единица. Известно [1], что статистика $l_m^{a_m}(x_m)$ является мартингалом относительно системы измеримых пространств $\{L_m, \mathcal{B}_m, \mu_m\}$ и сходится к $l(x,a)$ тогда и только тогда, когда она является равномерно интегрируемой функцией относительно меры μ, что в данном случае заведомо выполнено. Таким образом, имеет место следующая теорема.

Теорема 3. Пусть $K(x) \in CL$, $0 < K(x) \le 1$, $\lambda_n \to \infty$, $\frac{\lambda_n^2 \ln n}{n} \to 0$. Тогда статистика $\hat{l}_n^m(x_m)$, определенная по формуле (2), сходится в $C(H)$ к $l(x,a)$ с вероятностью единица.

Доказательство. Для заданного $\varepsilon > 0$ выберем сначала натуральное число N_1 такое, что $\sup_{x \in H} | l_m^{a_m}(P_m x) - l(x,a) | < \frac{\varepsilon}{2}$ при $m > N$. Это возможно, поскольку $l_m^{a_m}(P_m x)$ является равномерно ограниченной функцией. Далее выбираем такое натуральное число N_2, чтобы имело место соотношение $\sup_{x \in H} | \hat{l}_n^m(P_m x) - l_m^{a_m}(P_m x) | < \frac{\varepsilon}{2}$ при $n > N_2$ с вероятностью единица. Следовательно, объединяя эти два момента, получаем

$$\sup_{x \in H} | \hat{l}_n^m(P_m x) - l(x,a) | < \varepsilon \text{ при } m, n > N = \max\{N_1, N_2\}.$$

Для полного доказательства теоремы 3 остается показать, что функция $\hat{l}_n^m(x_m)$ является равномерно интегрируемой по m относительно меры μ с вероятностью единица.

Для заданного $\varepsilon > 0$ необходимо указать такое не зависящее от m число $\delta > 0$, чтобы при $\mu(A) < \delta$, $A \in \mathcal{B}$ для любого m с вероятностью единица имело место соотношение

$$\left| \int_A \hat{l}_n^m(x_m) \mu(dx) \right| < \varepsilon .$$

Заметим, что $l_m^{a_m}(x_m)$ является равномерно интегрируемой функцией по мере μ. Поэтому для $\frac{\varepsilon}{\ln n}$ найдется не зависящее от m такое число $\delta > 0$, чтобы при $\mu(A) < \delta$, $A \in \mathcal{B}$ для любого m имело место соотношение

$$\left| \int_A l_m^{a_m}(x_m) \mu(dx) \right| < \frac{\varepsilon}{\ln n} .$$

Далее можно записать

$$\left|\int_A \hat{l}_n^m(x_m)\mu(dx)\right| = \left|\int_A \left(\left[\ln\sum_{i=1}^n\prod_{j=1}^m K(\lambda_n(x_m^j - X_{ij}^m))\right]', a_m\right)_{L_m} \mu(dx)\right| =$$

$$= \left|\int_A \left(\left[\ln\sum_{i=1}^n\prod_{j=1}^m K(\lambda_n(x_m^j - X_{ij}^m))\right]', a_m\right)_{L_m} \mu_m(dx_m)\right| =$$

$$= \left|\int_A \left[\ln\sum_{i=1}^n\prod_{j=1}^m K(\lambda_n(x_m^j - X_{ij}^m))\right] l_m^{a_m}(x_m)\mu_m(dx_m)\right| \le$$

$$\le \ln n \left|\int_A l_m^{a_m}(x_m)\mu(dx)\right| < \ln n \;\frac{\varepsilon}{\ln n} = \varepsilon .$$

Теорема доказана.

Полученные результаты можно применить, например, в случае, когда в гильбертовом пространстве H рассматривается случайный процесс $x(t)$, в качестве наблюдений имеем траектории этого случайного процесса и требуется оценить гладкость распределения такого процесса. Этому вопросу будет посвящена отдельная работа.

СПИСОК ЛИТЕРАТУРЫ

1. Скороход А.В. Интегрирование в гильбертовом пространстве. — М.: Наука, 1975. — 232 с.
2. Далецкий Ю.Л., Фомин С.В. Меры и дифференциальные уравнения в бесконечномерных пространствах. — М.: Наука, 1983. — 383 с.
3. Сазонов С.В. К теореме Гливенко–Кантелли // Теория вероятностей и ее применения. — 1963. — **8**, вып. 3. — С. 299–303.
4. Nadaraya E. Nonparametric estimation of probability densities and regression curves. — Kluver Academic Publishers, 1989. — 214 p.
5. Мания Г.М. Статистическое оценивание распределения вероятностей. — Тбилиси: Тбилисский университет, 1974. — 240 с.
6. Schuster E. Estimation of a probability density function and its derivatives // Ann. Math. Stat. — 1969. — **40**, N 4. — P. 1187–1195.
7. Надареишвили М.М. Непараметрическая оценка функции регрессии и плотности распределения для одной модели // Сообщ. АН ГССР. — 1981. — **104**, № 2. — С. 293–296.
8. Надарая Э.А. О непараметрических оценках производных плотности вероятности и функции регрессии // Там же. — 1969. — **55**, № 1. — С. 29–32.

Поступила 21.01.2009