Math-Net.Ru

Общероссийский математический портал

Ю. Ю. Линке, Об уточнении одношаговых оценок Фишера в случае медленно сходящихся предварительных оценок, *Теория вероятн. и ее примен.*, 2015, том 60, выпуск 1, 80–98

DOI: <https://doi.org/10.4213/tvp4606>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 35.174.16.151

26 апреля 2023 г., 15:31:43

© 2015 г.

ЛИНКЕ Ю. Ю.*

**ОБ УТОЧНЕНИИ ОДНОШАГОВЫХ ОЦЕНОК ФИШЕРА В СЛУЧАЕ
МЕДЛЕННО СХОДЯЩИХСЯ ПРЕДВАРИТЕЛЬНЫХ ОЦЕНОК¹⁾**

Приводится алгоритм построения асимптотически эффективных одношаговых оценок неизвестного параметра в случае, когда имеется предварительная n^β -состоятельная оценка. Известные ранее одношаговые оценки, предложенные Фишером как приближения для состоятельных оценок максимального правдоподобия, позволяют за один шаг находить асимптотически эффективные оценки лишь в случае, когда $\beta \geq 1/4$. Новые одношаговые оценки специально ориентированы на медленно сближающиеся с параметром предварительные оценки и позволяют за одну итерацию n^β -состоятельную оценку при $\beta < 1/4$ улучшать до асимптотически эффективной. Найдены достаточно общие условия, при которых эти одношаговые оценки могут быть асимптотически эффективными даже в случаях, когда оценки максимального правдоподобия не существуют, либо существуют, но не являются состоятельными.

Ключевые слова и фразы: одношаговые оценки, оценки максимального правдоподобия, предварительные оценки, точность предварительной оценки, приближенный поиск оценок максимального правдоподобия, метод Ньютона.

1. Введение.

1.1. Рассматривается задача оценивания числового параметра $\theta \in \Theta \subset \mathbf{R}$ по выборке объема n из последовательности $X_1, X_2, \dots$ независимых одинаково распределенных случайных величин произвольной природы, распределение которых имеет плотность $f_\theta(\cdot)$ относительно некоторой σ -конечной меры μ . В широких условиях некоторым эталоном точности при оценивании неизвестного параметра θ принято считать оценку максимального правдоподобия $\hat{\theta}_n$. Хорошо известно, что при выполнении некоторых условий регулярности имеет место следующая сходимость:

$$\sqrt{n}(\hat{\theta}_n - \theta) \implies \mathcal{N}(0, I^{-1}(\theta)), \quad (1)$$

где $I(t)$ — информация Фишера, соответствующая плотности $f_t(x)$, а запись вида $\eta_n \Rightarrow \mathcal{N}(0, \sigma^2)$ означает слабую сходимость распределений η_n

*Институт математики им. С. Л. Соболева, Новосибирский государственный университет, Новосибирск, Россия; e-mail: linke@math.nsc.ru

¹⁾ Работа выполнена при финансовой поддержке РФФИ (грант № 14-01-00220).

при $n \rightarrow \infty$ к нормальному закону с нулевым средним и дисперсией σ^2 . В дальнейшем оценки, удовлетворяющие соотношению (1), будем называть *асимптотически эффективными* для параметра θ .

Известны примеры распределений, когда отыскать собственно саму оценку максимального правдоподобия $\hat{\theta}_n$ затруднительно (например, в случае существования большого числа локальных максимумов у логарифмической функции правдоподобия), но можно вычислить (например, по методу моментов) некоторую достаточно хорошую (скажем, состоятельную) оценку θ_n^* параметра θ . Для таких ситуаций Р. Фишером [1] была предложена процедура оценивания, основанная на методе Ньютона приближенного вычисления решения уравнения $L_n(\hat{\theta}_n) = 0$, где $L_n(t)$ — логарифмическая функция правдоподобия:

$$L_n(t) = \sum_{i=1}^n l(t, X_i) \quad \text{при } l(t, x) = \ln f_t(x). \quad (2)$$

Следуя работам Р. Фишера, в качестве приближений для состоятельной оценки максимального правдоподобия $\hat{\theta}_n$ обычно используют одну из следующих двух статистик:

$$\theta_{n1}^{**} = \theta_n^* - \frac{L'_n(\theta_n^*)}{L''_n(\theta_n^*)} \quad \text{или} \quad \theta_{n2}^{**} = \theta_n^* + \frac{L'_n(\theta_n^*)}{nI(\theta_n^*)}, \quad (3)$$

где θ_n^* — некоторая предварительная оценка параметра θ .

1.2. Исследованию оценок вида (3) посвящен ряд публикаций (см., например, [2]–[10] и библиографию в этих работах). Отметим, что в литературе на протяжении нескольких последних десятилетий (см. подробности в [9]) для обозначения таких по сути *двухшаговых* оценок используется термин *одношаговые* оценки в том смысле, что при наличии некоторой *предварительной* оценки (существование которой зачастую лишь постулируется) именно *за один шаг* (метода Ньютона или некоторых его модификаций) происходит качественное улучшение этой начальной статистики. Следуя устоявшейся терминологии оценки такого типа будем называть *одношаговыми*.

Достаточные условия асимптотической эффективности оценок θ_{n1}^{**} и θ_{n2}^{**} получены в [3]–[5], [10].

Помимо оценивания статистиками θ_{n1}^{**} и θ_{n2}^{**} собственно самого параметра θ , зачастую изучается сближение этих оценок с оценкой максимального правдоподобия. Например, в [2] доказано (см. также библиографию в [7]), что при $n \rightarrow \infty$

$$\sqrt{n}(\theta_{nj}^{**} - \hat{\theta}_n) \xrightarrow{P} 0, \quad j = 1, 2. \quad (4)$$

В [7] и [6] установлено, что скорость сходимости в (4) при некоторых дополнительных предположениях может быть улучшена. В частности,

в этих работах доказано, что

$$n(\theta_{n1}^{**} - \hat{\theta}_n) = O_p(1) \quad \text{и} \quad n(\theta_n^{**} - \hat{\theta}_n) = O_p(1),$$

где θ_n^{**} — некоторая оценка, в известном смысле аналогичная введенным в (3). В [8] найдено предельное распределение для $n(\theta_n^{**} - \hat{\theta}_n)$ (не являющееся нормальным).

Стоит отметить, что если асимптотическую эффективность оценок (3) извлекать из (4), то с неизбежностью нужно предполагать существование оценки максимального правдоподобия, а также выполнение условий, обеспечивающих сходимость (1). Тем не менее одношаговые оценки (3) обладают несколько неожиданным свойством: эти оценки могут быть асимптотически эффективными даже в случаях, когда *оценки максимального правдоподобия не существуют, либо существуют, но не являются состоятельными*. Этот интересный эффект впервые был отмечен, по-видимому, Л. Ле Камом [5] и обсуждается, например, в [4] (см. также замечание 3.4). В [10] найдены новые достаточно общие условия, при которых оценки (3) являются асимптотически эффективными без каких-либо предположений о поведении $\hat{\theta}_n$.

Подчеркнем, что практически во всех известных автору работах при исследовании оценок Фишера в качестве предварительной оценки θ_n^* используются либо асимптотически нормальные оценки, либо чуть более широкий класс так называемых $\sqrt{n}$ -ограниченных оценок θ_n^* , т.е. оценок, удовлетворяющих соотношению $\sqrt{n}(\theta_n^* - \theta) = O_p(1)$. Исключение составляют лишь работы [3] и [10]. Будем говорить, что некоторая оценка θ_n^* является n^β -состоятельной, если

$$n^\beta(\theta_n^* - \theta) \xrightarrow{p} 0 \quad \text{при} \quad n \rightarrow \infty. \quad (5)$$

В [3] (теорема 5.5.4) при доказательстве асимптотической эффективности θ_{n1}^{**} предполагается выполненным условие $n^{1/4}$ -состоятельности оценки θ_n^* . В [10] асимптотическая эффективность θ_{n1}^{**} и θ_{n2}^{**} доказана при более широком спектре ограничений на точность предварительной оценки, начиная с $\sqrt{n}$ -ограниченных оценок θ_n^* и заканчивая $n^{1/4}$ -состоятельными оценками.

Но можно ли ослабить условие $n^{1/4}$ -состоятельности предварительной оценки при использовании оценок Фишера для отыскания асимптотически эффективной оценки? Оказывается, при одношаговом применении этих оценок — нет: при широких ограничениях сходимость (5) при $\beta = 1/4$ является *необходимым* или близким к необходимому ограничением для асимптотической эффективности оценок (3) и справедливости (4) (см. [10] и замечание 3.1). Другой подход, связанный с многошаговым оцениванием, обсуждается в замечании 3.5.

Тем не менее существуют параметрические семейства распределений (например, распределения с так называемыми «тяжелыми» хвостами; см. замечание 3.2), для которых легко вычисляемая предварительная оценка является лишь n^β -состоятельной при $\beta < 1/4$.

В завершении обзора отметим, что методология одношагового оценивания успешно используется в различных специальных статистических задачах в случае разнораспределенной выборки (см. подробную библиографию в [9]). В работе автора [9] изучается асимптотическое поведение одношаговых M -оценок, построенных по выборке независимых не обязательно одинаково распределенных случайных величин в схеме серий и являющихся явными приближениями для состоятельных M -оценок (эти одношаговые оценки обобщают оценки Фишера (3)). В [9] найдены достаточно общие условия асимптотической нормальности одношаговых M -оценок при широком спектре ограничений на точность предварительной оценки.

1.3. Цель данной работы — исследовать новый класс одношаговых асимптотически эффективных оценок, специально рассчитанных на начальные приближения параметра, сближающиеся с θ медленнее, чем $n^{-1/4}$.

В работе предполагается, что при некотором натуральном k нам известна $n^{1/(2(k+2))}$ -состоятельная оценка $\theta_{n,k}^*$. В этом случае в качестве одношаговой асимптотически эффективной оценки $\theta_{n,k}^{**}$ предлагается использовать статистику

$$\theta_{n,k}^{**} = \theta_{n,k}^* + L'_n(\theta_{n,k}^*) \left(\sum_{i=1}^n \lambda_k(\theta_{n,k}^*, X_i) \right)^{-1}.$$

При этом функция $\lambda_k(\theta, x)$ имеет следующую структуру:

$$\lambda_k(\theta, x) = I(\theta) + c_{k1,\theta} \frac{f'_\theta(x)}{f_\theta(x)} + c_{k2,\theta} \frac{f''_\theta(x)}{f_\theta(x)} + \dots + c_{kk,\theta} \frac{f_\theta^{(k)}(x)}{f_\theta(x)},$$

где $c_{k1,\theta}, \dots, c_{kk,\theta}$ — решения некоторой системы линейных уравнений, которую мы определим позже; дифференцирование производится по переменной θ , а $f_\theta^{(k)}(x)$ обозначает k -ю производную плотности $f_\theta(x)$.

В частности, при $k = 1$ (т.е. при выполнении условия $n^{1/6}$ -состоятельности предварительной оценки) одношаговая оценка определяется соотношением

$$\theta_{n,1}^{**} = \theta_{n,1}^* + L'_n(\theta_{n,1}^*) (nI(\theta_{n,1}^*) + L'_n(\theta_{n,1}^*) c(\theta_{n,1}^*))^{-1}, \quad (6)$$

где

$$c(\theta) := \frac{1}{2I(\theta)} \int \frac{f'_\theta(x) f''_\theta(x)}{f_\theta(x)} \mu(dx).$$

В работе получены достаточные условия асимптотической эффективности одношаговых оценок $\theta_{n,k}^{**}$. Подчеркнем, что эти условия не предполагают состоятельность и даже существование оценки максимального правдоподобия (см. замечание 3.4). Как видно из приведенных выше определений, предлагаемые оценки $\theta_{n,k}^{**}$ отличаются от оценки Фишера θ_{n2}^{**} в (3) наличием некоторых дополнительных слагаемых в знаменателе соответствующей дроби, при этом количество этих слагаемых определяется скоростью сходимости предварительной оценки $\theta_{n,k}^*$ (т.е. задается значением k). Кроме того, от значения k зависят и коэффициенты $c_{k1,\theta}, \dots, c_{kk,\theta}$, задающие эти аддитивные добавки. Тем самым, с ростом k не только добавляется соответствующее количество слагаемых, но при каждом новом k заново пересчитываются и коэффициенты $c_{k1,\theta}, \dots, c_{kk,\theta}$.

Таким образом, в настоящей работе построены одношаговые оценки для параметра θ , позволяющие n^β -состоятельную оценку при любом $\beta < 1/4$ улучшать за один шаг до асимптотически эффективной оценки.

2. Основные результаты.

2.1. Всюду в работе, даже если это не оговорено особо, считаем выполненными следующие два предположения.

(А) Случайные величины $X_1, X_2, \dots$, принимающие значения в произвольном измеримом пространстве $\mathcal{X}$, независимы и одинаково распределены с плотностью $f_\theta(x)$ относительно некоторой σ -конечной меры μ на $\mathcal{X}$, где $\theta \in \Theta \subset \mathbf{R}$ и Θ — открытое множество.

(В) Логарифмическая плотность $l(t, x) = \ln f_t(x)$ для почти всех (по мере μ) значений x дважды непрерывно дифференцируема по t при всех $t \in \Theta$, и имеют место соотношения

$$\mathbf{E}_\theta l'(\theta, X_1) = 0, \quad 0 < I(\theta) := \mathbf{E}_\theta (l'(\theta, X_1))^2 < \infty, \quad \mathbf{E}_\theta l''(\theta, X_1) = -I(\theta), \quad (7)$$

где символ $\mathbf{E}_\theta$ обозначает математическое ожидание по распределению с плотностью $f_\theta(x)$.

При некотором фиксированном $k \geq 1$ нам также потребуется следующее условие.

(С_k) Имеется некоторая $n^{1/(2(k+2))}$ -состоятельная оценка $\theta_{n,k}^* = \theta_{n,k}^*(X_1, \dots, X_n)$ параметра θ .

Пусть выполнены предположения (А), (В) и (С_k). Задача состоит в том, чтобы построить оценку $\theta_{n,k}^{**} = \theta_{n,k}^{**}(\theta_{n,k}^*)$, для которой имеет место сходимость

$$\sqrt{n}(\theta_{n,k}^{**} - \theta) \implies \mathcal{N}(0, I^{-1}(\theta)) \quad \text{при } n \rightarrow \infty, \quad (8)$$

т.е. за один шаг улучшить предварительную оценку $\theta_{n,k}^*$ параметра θ сразу до асимптотически эффективной оценки $\theta_{n,k}^{**}$.

Одношаговую оценку $\theta_{n,k}^{**}$ мы предлагаем искать в виде

$$\theta_{n,k}^{**} = \theta_{n,k}^* + L'_n(\theta_{n,k}^*) \left(\sum_{i=1}^n \lambda_k(\theta_{n,k}^*, X_i) \right)^{-1}, \quad (9)$$

где функция $\lambda_k(t, x)$ определяется скоростью сближения предварительной оценки $\theta_{n,k}^*$ и параметра θ , заданной в условии (C_k) . В (9) и далее сохранено обозначение $L_n(t)$ для логарифмической функции правдоподобия: $L_n(t) = \sum_{i=1}^n l(t, X_i)$.

Как будет показано далее, для наших целей нужно, чтобы функция $\lambda_k(t, x)$ удовлетворяла соотношениям

$$\mathbf{E}_\theta(m\lambda_k^{(m-1)}(\theta, X_1) + l^{(m+1)}(\theta, X_1)) = 0 \quad \text{при всех } 1 \leq m \leq k+1. \quad (10)$$

Такие функции $\lambda_k(t, x)$ будут найдены в п. 2.2 (теорема 2.2). Здесь же поясним идею построения одношаговых оценок $\theta_{n,k}^{**}$, приводящую к выбору в качестве функции $\lambda_k(t, x)$ решения системы интегродифференциальных уравнений (10). Воспользуемся следующим тождеством, которое нетрудно извлечь из (9):

$$\sqrt{n}(\theta_{n,k}^{**} - \theta) = \frac{\sum_{i=1}^n g_k(\theta_{n,k}^*, X_i)/\sqrt{n}}{\sum_{i=1}^n \lambda_k(\theta_{n,k}^*, X_i)/n}, \quad (11)$$

$$\text{где } g_k(t, x) = (t - \theta)\lambda_k(t, x) + l'(t, x).$$

Поскольку при широких ограничениях

$$\frac{1}{n} \sum_{i=1}^n \lambda_k(\theta_{n,k}^*, X_i) \xrightarrow{p} \mathbf{E}_\theta \lambda_k(\theta, X_1) \quad \text{при } n \rightarrow \infty, \quad (12)$$

то асимптотическое поведение разности $\sqrt{n}(\theta_{n,k}^{**} - \theta)$ определяется поведением числителя в правой части первого равенства в (11). При соответствующей гладкости функции $g_k(t, x)$ по первому аргументу имеем

$$\begin{aligned} \frac{\sum_{i=1}^n g_k(\theta_{n,k}^*, X_i)}{\sqrt{n}} &= \frac{\sum_{i=1}^n g_k(\theta, X_i)}{\sqrt{n}} \\ &+ \sum_{m=1}^{k+1} \frac{(\theta_{n,k}^* - \theta)^m}{m!} \frac{\sum_{i=1}^n g_k^{(m)}(\theta, X_i)}{\sqrt{n}} \\ &+ \sqrt{n}(\theta_{n,k}^* - \theta)^{k+2} G_{n,k}, \end{aligned} \quad (13)$$

где $g_k^{(m)}(t, x)$ — m -я производная функции $g_k(t, x)$ по t , а $G_{n,k}$ — остаточный член формулы Тейлора, который мы определим далее в лемме 4.2. Кроме того, при некоторых естественных ограничениях имеют место соотношения

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n g_k(\theta, X_i) \equiv \frac{1}{\sqrt{n}} \sum_{i=1}^n l'(\theta, X_i) \Rightarrow \mathcal{N}(0, I(\theta)) \quad \text{и} \quad G_{n,k} = O_p(1). \quad (14)$$

Второе соотношение в (14) и условие (C_k) позволяют заключить, что последнее слагаемое в правой части (13) сходится по вероятности к нулю. Допустим еще, что функция $\lambda_k(\theta, X_1)$ такова, что выполнено следующее ключевое соотношение:

$$\mathbf{E}_\theta g_k^{(m)}(\theta, X_1) = 0 \quad \text{при } 1 \leq m \leq k+1. \quad (15)$$

В этом случае при соответствующих моментных ограничениях распределения сумм $\sum_{i=1}^n g_k^{(m)}(\theta, X_i)/\sqrt{n}$ сходятся при $n \rightarrow \infty$ к нормальному распределению и из первого соотношения в (14) и разложения (13) немедленно следует, что $\sum_{i=1}^n g_k(\theta_{n,k}^*, X_i)/\sqrt{n} \Rightarrow \mathcal{N}(0, I(\theta))$. Тем самым, выполнено (8), если еще учесть (11), (12) и считать, что справедливо равенство $\mathbf{E}_\theta \lambda_k(\theta, X_1) = I(\theta)$.

Нам остается заметить, что в силу представления (11) для любого натурального m имеет место равенство

$$g_k^{(m)}(t, x) = m\lambda_k^{(m-1)}(t, x) + l^{(m+1)}(t, x) + (t - \theta)\lambda_k^{(m)}(t, x), \quad (16)$$

а потому соотношение (15) эквивалентно (10).

Приведем теперь точное утверждение об асимптотических свойствах оценок $\theta_{n,k}^{**}$. Нам потребуется следующее предположение.

(D_k) 1) Функция $\lambda_k(t, x)$ для почти всех значений x (по мере μ) является $k+1$ раз дифференцируемой по t при всех $t \in \Theta$ и

$$\begin{aligned} \mathbf{E}_\theta \lambda_k(\theta, X_1) &= I(\theta), \quad \mathbf{E}_\theta \left| \lambda_k^{(m)}(\theta, X_1) \right|^2 < \infty \quad \text{при } m = 0, \dots, k, \\ \mathbf{E}_\theta \sup_{t \in \Theta: |t-\theta| \leq \delta_1} |\lambda_k^{(k+1)}(t, X_1)| &< \infty \quad \text{при некотором } \delta_1 > 0. \end{aligned} \quad (17)$$

2) Функция $l(t, x)$ для почти всех значений x (по мере μ) является $k+2$ раза дифференцируемой по t при всех $t \in \Theta$, $\mathbf{E}_\theta |l^{(m+2)}(\theta, X_1)|^2 < \infty$ при $0 \leq m \leq k$ и

$$\mathbf{E}_\theta \sup_{t \in \Theta: |t-\theta| < \delta_2, t \neq \theta} \frac{|l^{(k+2)}(t, X_1) - l^{(k+2)}(\theta, X_1)|}{|t - \theta|} < \infty \quad (18)$$

при некотором $\delta_2 > 0$.

Имеет место следующее утверждение.

Теорема 2.1. Пусть выполнены предположения (A), (B) и при некотором $k \geq 1$ — предположения (C_k) и (D_k). Кроме того, пусть функция $\lambda_k(\theta, x)$ удовлетворяет системе уравнений (10). Тогда одношаговая оценка $\theta_{n,k}^{**}$, определенная в (9), асимптотически эффективна.

2.2. В этом пункте мы найдем функции $\lambda_k(t, x)$, удовлетворяющие системе (10) для любого целого $k \geq 1$. Нам потребуется несколько дополнительных обозначений. В пространстве $L_2 = \{g(\cdot): \mathbf{E}_\theta g^2(X_1) < \infty\}$

функций с конечным вторым моментом определим скалярное произведение $\langle \alpha(\cdot), \beta(\cdot) \rangle$ соотношением

$$\langle \alpha(\cdot), \beta(\cdot) \rangle := \mathbf{E}_\theta \alpha(X_1) \beta(X_1), \quad \alpha(\cdot), \beta(\cdot) \in L_2. \quad (19)$$

Положим

$$e_{m,\theta}(\cdot) := \frac{f_\theta^{(m)}(\cdot)}{f_\theta(\cdot)}, \quad a_{m,\theta} := \frac{1}{m+1} \langle e_{1,\theta}(\cdot), e_{m+1,\theta}(\cdot) \rangle, \quad (20)$$

где $m \geq 0$. При любом фиксированном $k \geq 1$ обозначим через $c_{k1,\theta}, \dots, c_{kk,\theta}$ решения следующей системы k линейных уравнений с k неизвестными:

$$c_{k1,\theta} \langle e_{1,\theta}(\cdot), e_{j,\theta}(\cdot) \rangle + \dots + c_{kk,\theta} \langle e_{k,\theta}(\cdot), e_{j,\theta}(\cdot) \rangle = a_{j,\theta}, \quad j = 1, \dots, k. \quad (21)$$

Отметим, что матрица, определяющая систему (21) (матрица Грама, построенная по элементам $\{e_{m,\theta}(\cdot); m \leq k\}$), является положительно определенной в случае линейной независимости функций $\{e_{m,\theta}(\cdot); m \leq k\}$ и, тем самым, система (21) однозначно задает элементы $c_{k1,\theta}, \dots, c_{kk,\theta}$. Теперь определим функцию $\lambda_k(\theta, x)$ равенством

$$\lambda_k(\theta, x) = I(\theta) + c_{k1,\theta} e_{1,\theta}(x) + c_{k2,\theta} e_{2,\theta}(x) + \dots + c_{kk,\theta} e_{k,\theta}(x). \quad (22)$$

Почему функция $\lambda_k(\theta, x)$ из (22) удовлетворяет системе (10)? Поясним кратко идею отыскания этой функции. Оказывается, что при достаточно общих условиях имеет место следующее ключевое утверждение (см. лемму 4.4): если функция $\lambda_k(\theta, x)$ является решением системы уравнений

$$\langle \lambda_k(\theta, \cdot), e_{m-1,\theta}(\cdot) \rangle = a_{m-1,\theta}, \quad m = 1, \dots, k+1, \quad (23)$$

то функция $\lambda_k(\theta, x)$ удовлетворяет системе (10). Тем самым, поиск подходящей функции $\lambda_k(\theta, x)$ сводится к отысканию решения системы (23). В предположении линейной независимости $e_{0,\theta}(x), \dots, e_{1,\theta}(x)$ система (23) задает проекции $a_{0,\theta}, \dots, a_{k,\theta}$ искомой функции $\lambda_k(\theta, x)$ на элементы базиса $e_{0,\theta}(x), \dots, e_{1,\theta}(x)$. Известно, что в этом случае линейная комбинация базисных элементов $\sum_{j=0}^k c_{kj,\theta} e_{j,\theta}(x)$ является решением (23), где коэффициенты $\{c_{kj,\theta}\}$ определяются системой линейных уравнений с матрицей Грама, построенной по элементам базиса $\{e_{j,\theta}(x)\}$ (см. доказательство леммы 4.5).

Если же функции $f_\theta(\cdot), f'_\theta(\cdot), \dots, f_\theta^{(k)}(\cdot)$ линейно зависимы (т.е. линейно зависимы функции $\{e_{m,\theta}(\cdot); m \leq k\}$), то, как хорошо известно, либо не существует решений системы линейных уравнений (21), либо их будет бесконечно много. Иными словами, в этих условиях либо невозможно представить искомую функцию $\lambda_k(\theta, x)$ в виде (22), либо таких

функций бесконечно много, и алгоритм их поиска сводится к известной процедуре линейной алгебры.

Чтобы сформулировать основное утверждение этого раздела, нам потребуется следующее предположение.

(**E_k**) Выполнено условие (**D_k**) без ограничений (17) и (18), функции $f_\theta(\cdot), f'_\theta(\cdot), \dots, f_\theta^{(k)}(\cdot)$ линейно независимы и $e_{2,\theta}(\cdot), \dots, e_{k+1,\theta}(\cdot) \in L_2$. Кроме того, $\int f_\theta^{(j)}(x) \mu(dx) = 0$, $j = 3, \dots, k+1$, и можно менять порядок дифференцирования по переменной θ и интегрирования по переменной x в следующих случаях:

$$\left(\int \lambda_k(\theta, x) f_\theta^{(m-1)}(x) \mu(dx) \right)^{(j)} = \int \left(\lambda_k(\theta, x) f_\theta^{(m-1)}(x) \right)^{(j)} \mu(dx),$$

$$\left(\int l'(\theta, x) f_\theta^{(k+1)}(x) \mu(dx) \right)^{(j)} = \int \left(l'(\theta, x) f_\theta^{(k+1)}(x) \right)^{(j)} \mu(dx)$$

при $m = 1, \dots, k+1$ и $j = 1, \dots, k+1-m$.

Имеет место следующая теорема.

Теорема 2.2. Пусть выполнены предположения (**A**), (**B**) и условие (**E_k**) при некотором фиксированном $k \geq 1$. Тогда функция $\lambda_k(t, x)$, определенная в (22), является решением системы уравнений (10).

2.3. Из теорем 2.1 и 2.2 вытекает следующее утверждение.

Следствие 2.1. Пусть выполнены предположения (**A**) и (**B**), а при некотором фиксированном $k \geq 1$ — предположения (**C_k**), (**D_k**) и (**E_k**). Тогда одношаговая оценка $\theta_{n,k}^{**}$, определенная в (9) и (22), является асимптотически эффективной.

Выделим в отдельное утверждение следствие из теорем 2.1 и 2.2 для случая $k = 1$, т.е. для случая, когда выполнено условие $n^{1/6}$ -состоятельности предварительной оценки. Из определений (20), (21) и (22) нетрудно извлечь представление (6) для одношаговой статистики $\theta_{n,1}^{**}$. Справедливо следующее утверждение.

Следствие 2.2. Пусть выполнены условия (**A**), (**B**), (**C₁**) и следующие предположения:

1) функция $l(t, x)$ для почти всех значений x (по мере μ) трижды дифференцируема по t , функции $I(t)$ и $a_{1,t}$ — дважды дифференцируемы при всех $t \in \Theta$, а при дифференцировании интегралов $\int l'(t, x) f'_t(x) \mu(dx)$ и $\int l'(t, x) f''_t(x) \mu(dx)$ по t при всех $t \in \Theta$ можно менять порядок дифференцирования и интегрирования;

2) выполнено (18) при $k = 1$ и $\mathbf{E}_\theta |l''(\theta, X_1)|^2 < \infty$, $\mathbf{E}_\theta |l'''(\theta, X_1)|^2 < \infty$;

3) функции $f_\theta(\cdot)$ и $f'_\theta(\cdot)$ линейно независимы.

Тогда одношаговая оценка $\theta_{n,1}^{**}$, определенная в (6), асимптотически эффективна.

3. Заключительные комментарии.

З а м е ч а н и е 3.1. Поясним, не приводя подробных доказательств, появление необходимых условий на точность предварительной оценки θ_n^* , возникающих при использовании оценок Фишера (3).

Для производной логарифмической функции правдоподобия имеет место равенство $0 = L'_n(\hat{\theta}_n) = L'_n(\theta_n^*) + (\hat{\theta}_n - \theta_n^*)L''_n(\theta_n^*) + (\hat{\theta}_n - \theta_n^*)^2 L'''_n(\hat{\theta}_n)$, где $\hat{\theta}_n^*$ — промежуточная точка между θ_n^* и оценкой максимального правдоподобия $\hat{\theta}_n$. Используя это равенство и определение (3), получаем следующее тождество:

$$\begin{aligned}\sqrt{n}(\theta_{n1}^{**} - \hat{\theta}_n) &= \sqrt{n}(\theta_n^* - \hat{\theta}_n) - \sqrt{n} \frac{L'_n(\theta_n^*) - L'_n(\hat{\theta}_n)}{L''_n(\theta_n^*)} \\ &= \frac{1}{2} (n^{1/4}(\theta_n^* - \hat{\theta}_n))^2 \frac{L'''_n(\hat{\theta}_n^*)}{L''_n(\theta_n^*)},\end{aligned}\quad (24)$$

где в широких условиях $L'''_n(\hat{\theta}_n^*)/L''_n(\theta_n^*) \xrightarrow{P} -\mathbf{E}_\theta l'''(\theta, X_1)/I(\theta)$. Следовательно, если $\sqrt{n}(\theta_{n1}^{**} - \hat{\theta}_n) \xrightarrow{P} 0$ и $\mathbf{E}_\theta l'''(\theta, X_1) \neq 0$, то с необходимостью $n^{1/4}(\theta_n^* - \theta) \xrightarrow{P} 0$. Аналогичные рассуждения справедливы и для второй оценки θ_{n2}^{**} из (3).

В [10] доказано, что при некоторых ограничениях условие $n^{1/4}$ -состоятельности оценки θ_n^* является необходимым для сходимости

$$\sqrt{n}(\theta_{n1}^{**} - \theta) - L'_n(\theta)(I(\theta)\sqrt{n})^{-1} \xrightarrow{P} 0.$$

Таким образом, условие $n^{1/4}$ -состоятельности предварительной оценки в широких условиях оказывается *необходимым* для асимптотической эквивалентности (4) оценки максимального правдоподобия $\hat{\theta}_n$ и оценок Фишера (3) и для доказательства их асимптотической эффективности.

З а м е ч а н и е 3.2. Пусть распределение сосредоточено на положительной полуоси и плотность $f_\theta(x)$ при $x \geq 0$ задается соотношением

$$f_\theta(x) = \frac{h+1}{\theta(1+x/\theta)^{2+h}}, \quad h \in \left(0, \frac{1}{3}\right]. \quad (25)$$

Нетрудно проверить, что $\mathbf{E}_\theta X_1 = \theta/h$. Следовательно, $\theta_n^* = h \sum_{i=1}^n X_i/n$ — оценка по методу моментов для параметра θ . В п. 4.3 будет доказано следующее предложение.

Предложение 3.1. Если выполнено условие (25), то распределение $n^\alpha(\theta_n^* - \theta)$ при $\alpha = h/(h+1) \in (0, 1/4]$ сходится к некоторому устойчивому распределению.

В данном примере $n^\alpha(\theta_n^* - \theta) = O_p(1)$ при $\alpha \leq 1/4$. Следовательно, оценка θ_n^* является n^β -состоятельной при $\beta < \alpha \leq 1/4$, при

этом $E_{\theta} l'''(\theta, X_1) \neq 0$. В такой ситуации оценки Фишера (3) не подходят для одношагового поиска асимптотически эффективной оценки (см. замечание 3.1). Здесь можно использовать предлагаемые в данной работе одношаговые оценки $\theta_{n,k}^{**}$ (см. также замечание 3.5). Понятно, что индекс k при этом нужно выбирать в зависимости от скорости сближения θ_n^* и θ , определяемой значением h . Например, если $h \in (1/5, 1/3]$, то выполнено условие $n^{1/6}$ -состоятельности для $\theta_{n,1}^* = \theta_n^*$ (т.е. $k = 1$) и в качестве одношагового приближения нужно использовать оценку $\theta_{n,1}^{**}$, определенную в (6).

Отметим еще, что для плотности из (6) можно указать диапазон значений h , при которых скорость сходимости предварительной оценки и параметра позволяет использовать процедуру Фишера (3) для построения асимптотически эффективной оценки за один шаг. Так, просматривая доказательство предложения 3.1, нетрудно видеть, что $n^{\alpha}(\theta_n^* - \theta) = O_p(1)$ при $\alpha \in (1/4, 1/2]$ в случае, когда $h \in (1/3, 1]$. И этот диапазон h — область применения классических одношаговых оценок Фишера (3) (см. подробнее [10]).

З а м е ч а н и е 3.3. Отметим, что чем медленнее предварительная оценка (оценка $\theta_{n,k}^*$) сближается с параметром θ (т.е. чем больше значение k), тем более гладкой должна быть плотность распределения, чтобы имелась возможность за один шаг, с помощью $\theta_{n,k}^{**}$, находить асимптотически эффективную оценку. Этот факт наглядно демонстрируют условия (C_k) , (D_k) и определения (22), (20).

Аналогичная взаимосвязь между точностью предварительной оценки и гладкостью плотности распределения для выполнения тех или иных асимптотических свойств оценок Фишера (3) была отмечена в [10]: чем менее точна предварительная оценка, тем более гладкой должна быть плотность распределения для того, чтобы оценки Фишера являлись асимптотически эффективными.

З а м е ч а н и е 3.4. Подчеркнем, что все используемые нами ограничения, как и условия из работ [5] и [10], — локальные, т.е. ограничения накладываются на поведение производных логарифмической плотности лишь в некоторой окрестности $t \in \Theta$: $|t - \theta| \leq \delta$, $\delta > 0$ неизвестного параметра θ .

В частности, допускается, что вне этой окрестности функция правдоподобия может неограниченно возрастать, т.е. допускается возможность несуществования оценки максимального правдоподобия $\hat{\theta}_n$. Кроме того, допускается возможность, что $\hat{\theta}_n$ существует, но не состоятельна.

Таким образом, построенные в работе новые одношаговые оценки сохраняют свойство, присущее оценкам Фишера (3): оценки $\theta_{n,k}^{**}$, как и оценки (3), могут быть асимптотически эффективными даже в случае, когда оценка максимального правдоподобия не существует, либо не состоятельна.

Отмеченное, казалось бы, удивительное свойство обсуждаемых одношаговых оценок объясняется следующим. Оказывается, при некоторых локальных и очень слабых ограничениях (не гарантирующих состоятельность и даже существование оценки максимального правдоподобия) в окрестности θ существует и единственная точка локального максимума логарифмической функции правдоподобия $L_n(t)$ (обозначим эту точку через $\hat{\theta}_n(\theta)$), при этом с ростом n имеет место сходимость $\sqrt{n}(\hat{\theta}_n(\theta) - \theta) \Rightarrow \mathcal{N}(0, I^{-1}(\theta))$. Таким образом, асимптотическая эффективность одношаговых оценок в случае, когда оценки максимального правдоподобия не состоятельны, или даже не определены, объясняется тем, что эти оценки приближают точку локального максимума функции правдоподобия (ближайшую к θ точку $\hat{\theta}_n(\theta)$ локального максимума функции правдоподобия), обладающую некоторыми нужными свойствами.

Стоит отметить, что $\hat{\theta}_n(\theta)$ — это не статистика. В случае же, когда есть некоторая состоятельная оценка θ_n^* , в качестве асимптотически эффективной оценки параметра θ можно использовать суперпозицию $\hat{\theta}_n(\theta_n^*)$ — точку ближайшего к θ_n^* максимума, при этом $\mathbf{P}(\hat{\theta}_n(\theta) = \hat{\theta}_n(\theta_n^*)) \rightarrow 1$ с ростом n .

Несмотря на то что поиск локального экстремума $\hat{\theta}_n(\theta_n^*)$ — задача менее трудоемкая, чем поиск глобального экстремума (т.е. оценки $\hat{\theta}_n$), вычислить в явном виде $\hat{\theta}_n(\theta_n^*)$ может быть также затруднительно. В этой ситуации процедура одношагового оценивания — разумная альтернатива, позволяющая в явном виде определять асимптотически эффективные оценки. При этом выбор одношаговой оценки зависит от точности θ_n^* : если предварительная оценка $n^{1/4}$ -состоятельна, то можно использовать оценки Фишера (3), а в случае менее точной аппроксимации можно использовать построенные в данной работе одношаговые оценки.

Отметим еще, что в случае, когда оценка максимального правдоподобия $\hat{\theta}_n$ определена и состоятельна, то при широких условиях оценки $\hat{\theta}_n$, $\hat{\theta}_n(\theta_n^*)$ и величина $\hat{\theta}_n(\theta)$ совпадают с вероятностью, стремящейся к 1.

Упомянутые в этом замечании эффекты подробнее рассмотрены в автором в [9].

З а м е ч а н и е 3.5. Если предварительная оценка не достаточно точна для непосредственного использования одношаговой процедуры, то точность предварительной оценки иногда может быть улучшена многократным применением метода Ньютона.

Действительно, при некоторых дополнительных ограничениях для оценки Фишера θ_{n1}^{**} справедливо представление

$$\theta_{n1}^{**} - \theta = \frac{(\theta_n^* - \theta)L_n''(\theta_n^*) - L_n'(\theta_n^*)}{L_n''(\theta_n^*)} \sim \frac{\zeta_n}{\sqrt{n}} + (\theta_n^* - \theta)^2 \cdot C_n,$$

где

$$\zeta_n \Rightarrow \mathcal{N}(0, I^{-1}(\theta)) \quad \text{и} \quad C_n \xrightarrow{p} -\frac{\mathbf{E} l'''(\theta, X_1)}{2I(\theta)}.$$

Таким образом, если мы располагаем n^ε -состоятельной оценкой θ_n^* параметра θ и $\mathbf{E} |l'''(\theta, X_1)| < \infty$, то $\theta_{n1}^{**} = \theta_{n1}^{**}(\theta_n^*)$ будет $n^{2\varepsilon}$ -состоятельной оценкой (если $2\varepsilon < 1/2$). Далее, используя θ_{n1}^{**} в качестве предварительной оценки, мы получим $n^{4\varepsilon}$ -состоятельную оценку (если $4\varepsilon < 1/2$) и т.д. Нетрудно видеть, что продолжая этот процесс некоторое число раз, мы достигнем нужной точности (т.е. $\sqrt{n}$ -ограниченности), при этом количество итераций будет равно $[-\ln 2\varepsilon / \ln 2] + 1$.

Таким образом, располагая $n^{1/2(k+2)}$ -состоятельной предварительной оценкой, для отыскания асимптотически эффективной оценки можно, с одной стороны, использовать одношаговую статистику $\theta_{n,k}^{**}$, предложенную в настоящей работе. С другой стороны, применяя последовательно определенное количество раз преобразование Фишера, можно также найти асимптотически эффективную оценку за число шагов $\ln(k+2)/\ln 2 = \log_2(k+2)$.

4. Доказательства. В пп. 4.1 и 4.2 предполагаем фиксированным некоторое значение $k \geq 1$ и считаем выполненными предположения (A) и (B), особо это не оговаривая.

4.1. В этом пункте мы докажем теорему 2.1. Нам потребуется три вспомогательных утверждения.

Лемма 4.1. Пусть выполнено предположение (D_k). Тогда

$$\begin{aligned} \tau_\lambda(\delta) &:= \mathbf{E}_\theta \sup_{t \in \Theta: |t-\theta| \leq \delta} |\lambda_k(t, X_1) - \lambda_k(\theta, X_1)| \rightarrow 0 \quad \text{при } \delta \rightarrow 0, \\ \tau_g(\delta) &:= \mathbf{E}_\theta \sup_{t \in \Theta: |t-\theta| \leq \delta, t \neq \theta} \frac{|g_k^{(k+1)}(t, X_1) - g_k^{(k+1)}(\theta, X_1)|}{|t - \theta|} < \infty, \end{aligned}$$

где $\delta = \min\{\delta_1, \delta_2\}$ и функция $g_k(t, x)$ определена в (11).

Доказательство. Воспользуемся разложением

$$\lambda_k(t, x) = \sum_{m=0}^k \frac{(t-\theta)^m}{m!} \lambda_k^{(m)}(\theta, x) + \frac{(t-\theta)^{k+1}}{(k+1)!} \lambda_k^{(k+1)}(\tilde{t}, x),$$

где $\tilde{t}$ — некоторое промежуточное значение между t и θ . При $\delta < \delta_1$ имеем

$$\tau_\lambda(\delta) \leq \sum_{m=0}^k \frac{\delta^m}{m!} \mathbf{E}_\theta |\lambda_k^{(m)}(\theta, X_1)| + \frac{\delta^{k+1}}{(k+1)!} \mathbf{E}_\theta \sup_{t: |t-\theta| \leq \delta} |\lambda_k^{(k+1)}(\tilde{t}, x)|. \quad (26)$$

Эта оценка доказывает первое утверждение леммы, если еще учесть, что математические ожидания в правой части (26) конечны ввиду условия (D_k).

Второе утверждение леммы нетрудно извлечь из представления (16), если учесть следующие две оценки:

$$\begin{aligned}
& \mathbf{E}_\theta \sup_{t \in \Theta: |t - \theta| \leq \delta} \left| \lambda_k^{(k+1)}(t, X_1) - \lambda_k^{(k+1)}(\theta, X_1) \right| \\
& \leq \mathbf{E}_\theta \sup_{t \in \Theta: |t - \theta| \leq \delta_1} \left| \lambda_k^{(k+1)}(t, X_1) \right| + \mathbf{E}_\theta \left| \lambda_k^{(k+1)}(\theta, X_1) \right| < \infty, \\
& \mathbf{E}_\theta \sup_{t \in \Theta: |t - \theta| \leq \delta, t \neq \theta} \frac{|\lambda_k^{(k)}(t, X_1) - \lambda_k^{(k)}(\theta, X_1)|}{|t - \theta|} \\
& \leq \mathbf{E}_\theta \sup_{t \in \Theta: |t - \theta| \leq \delta_1} \left| \lambda_k^{(k+1)}(t, X_1) \right| < \infty.
\end{aligned}$$

Лемма 4.2. Пусть выполнены предположения $(\mathbf{C}_k)$ и $(\mathbf{D}_k)$. Тогда

$$G_{n,k} := \sum_{i=1}^n \frac{g_k^{(k+1)}(\tilde{\theta}_{n,k}^*, X_i) - g_k^{(k+1)}(\theta, X_i)}{n(\theta_{n,k}^* - \theta)} = O_p(1),$$

где $\tilde{\theta}_{n,k}^*$ — некоторая промежуточная точка между θ и $\theta_{n,k}^*$.

Доказательство. Для любого $K > 0$ и $\delta = \min\{\delta_1, \delta_2\}$ имеем

$$\begin{aligned}
& \mathbf{P}(|G_{n,k}| > K) \leq \mathbf{P}(|G_{n,k}| > K, |\theta_{n,k}^* - \theta| \leq \delta) + \mathbf{P}(|\theta_{n,k}^* - \theta| > \delta) \\
& \leq \mathbf{P}\left(\sup_{t: |t - \theta| \leq \delta, t \neq \theta} \sum_{i=1}^n \frac{|g_k^{(k+1)}(t, X_i) - g_k^{(k+1)}(\theta, X_i)|}{n|t - \theta|} > K, |\theta_{n,k}^* - \theta| \leq \delta\right) \\
& \quad + \mathbf{P}(|\theta_{n,k}^* - \theta| > \delta) \\
& \leq \frac{\tau_g(\delta)}{K} + \mathbf{P}(|\theta_{n,k}^* - \theta| > \delta). \tag{27}
\end{aligned}$$

При выводе (27) мы использовали формулу полной вероятности и неравенство Чебышёва с первым моментом. Следовательно, в силу леммы 4.1 и состоятельности оценки $\theta_{n,k}^*$, выполнено соотношение

$$\limsup_{n \rightarrow \infty} \mathbf{P}(|G_{n,k}| > K) \leq \frac{\tau_g(\delta)}{K} \rightarrow 0 \quad \text{при } K \rightarrow \infty,$$

что доказывает утверждение леммы.

Лемма 4.3. Пусть выполнены предположения $(\mathbf{C}_k)$ и $(\mathbf{D}_k)$. Тогда

$$\begin{aligned}
\Lambda_{n,k} &:= \frac{1}{n} \sum_{i=1}^n \lambda_k(\theta, X_i) \xrightarrow{p} I(\theta), \\
\bar{\Lambda}_{n,k} &:= \frac{1}{n} \sum_{i=1}^n (\lambda_k(\theta_{n,k}^*, X_i) - \lambda_k(\theta, X_i)) \xrightarrow{p} 0.
\end{aligned} \tag{28}$$

Доказательство. Первое утверждение леммы следует из закона больших чисел, поскольку функция $\lambda_k(\theta, x)$ такова, что $\mathbf{E}_\theta \lambda_k(\theta, X_1) = I(\theta)$.

Аналогично выводу (27), для любых $\varepsilon > 0$ и $\delta > 0$ имеем

$$\begin{aligned} \mathbf{P}(|\bar{\Lambda}_{n,k}| > \varepsilon) &\leq \mathbf{P}(|\bar{\Lambda}_{n,k}| > \varepsilon, |\theta_{n,k}^* - \theta| \leq \delta) + \mathbf{P}(|\theta_{n,k}^* - \theta| > \delta) \\ &\leq \mathbf{P}\left(\sup_{t: |t-\theta| \leq \delta} \frac{1}{n} \sum_{i=1}^n |\lambda_k(t, X_i) - \lambda_k(\theta, X_i)| > \varepsilon, |\theta_{n,k}^* - \theta| \leq \delta\right) \\ &\quad + \mathbf{P}(|\theta_{n,k}^* - \theta| > \delta) \\ &\leq \frac{\tau_\lambda(\delta)}{\varepsilon} + \mathbf{P}(|\theta_{n,k}^* - \theta| > \delta). \end{aligned} \quad (29)$$

Таким образом, $\limsup_{n \rightarrow \infty} \mathbf{P}(|\bar{\Lambda}_{n,k}| > \varepsilon) \leq \tau_\lambda(\delta)/\varepsilon$. А поскольку по лемме 4.1 $\tau_\lambda(\delta) \rightarrow 0$ при $\delta \rightarrow 0$, то в силу произвольности ε и δ второе утверждение леммы доказано.

Доказательство теоремы 2.1. Заметим прежде всего, что в условиях теоремы

$$\bar{G}_{nk} := \frac{1}{\sqrt{n}} \sum_{i=1}^n (g_k(\theta_{n,k}^*, X_i) - g_k(\theta, X_i)) \xrightarrow{P} 0. \quad (30)$$

Действительно, с учетом обозначения из леммы 4.2, имеем

$$\bar{G}_{nk} \equiv \sum_{m=1}^{k+1} \frac{(\theta_{n,k}^* - \theta)^m}{m!} \frac{1}{\sqrt{n}} \sum_{i=1}^n g_k^{(m)}(\theta, X_i) + \sqrt{n}(\theta_{n,k}^* - \theta)^{k+2} G_{n,k}. \quad (31)$$

Последнее слагаемое в правой части (31) сходится по вероятности к нулю в силу леммы 4.2 и условия $(\mathbf{C}_k)$. Далее, если $m = 1, \dots, k+1$, то $\mathbf{E}_\theta g_k^{(m)}(\theta, X_i) = 0$ по построению одношаговых оценок (см. п. 2.1), а $\mathbf{E}_\theta |g_k^{(m)}(\theta, X_i)|^2 < \infty$ в силу (16) и условия $(\mathbf{D}_k)$. Следовательно, в силу центральной предельной теоремы распределения $\sum_{i=1}^n g_k^{(m)}(\theta, X_i)/\sqrt{n}$ сходятся к нормальному, а потому все слагаемые суммы по m в правой части (31) сходятся по вероятности к нулю в силу состоятельности $\theta_{n,k}^*$.

Из (11) и (14) получаем, что

$$\sqrt{n}(\theta_{n,k}^{**} - \theta) = \frac{\sum g_k(\theta_{n,k}^*, X_i)/\sqrt{n}}{\sum \lambda_k(\theta_{n,k}^*, X_i)/n} = \frac{\sum l'(\theta, X_i)/\sqrt{n} + \bar{G}_{nk}}{\Lambda_{n,k} + \bar{\Lambda}_{n,k}}. \quad (32)$$

Сходимость $\sqrt{n}(\theta_{n,k}^{**} - \theta) \Rightarrow \mathcal{N}(0, I^{-1}(\theta))$ следует теперь из представления (32), леммы 4.3, первой сходимости в (14), а также (30).

4.2. Вывод теоремы 2.2 проведем в два этапа. Нам потребуются следующая лемма.

Лемма 4.4. Если выполнено условие $(\mathbf{E}_k)$ и функция $\lambda_k(t, x)$ является решением системы (23), то эта функция удовлетворяет и (10).

Доказательство. Уравнения (10) и (23) соответственно нам удобно переписать в следующем интегральном виде:

$$\int \lambda_k^{(m-1)} f = -\frac{1}{m} \int l^{(m+1)} f, \quad m = 1, \dots, k+1, \quad (33)$$

$$\int \lambda_k f^{(m-1)} = \frac{1}{m} \int l' f^{(m)}, \quad m = 1, \dots, k+1 \quad (34)$$

(условимся опускать аргументы у функций, а также переменную дифференцирования θ и переменную интегрирования x). Доказательство проведем методом индукции по m . Пусть при $m = 1$ функция λ_k удовлетворяет (34), т.е.

$$\int \lambda_k f = \int l' f'. \quad (35)$$

Дифференцируя тождество $\int l' f = 0$ по θ , имеем: $0 = \int l'' f + \int l' f'$. Следовательно, с учетом (35), $\int \lambda_k f = -\int l'' f$, т.е. выполнено (33) при $m = 1$.

Предположим теперь, что если λ_k удовлетворяет (34) при $m = 1, \dots, h-1$ для некоторого фиксированного $3 \leq h \leq k+1$, то λ_k удовлетворяет (33) при этих m . Покажем, что если, дополнительно, λ_k удовлетворяет (34) при $m = h$, то справедливо и (33) при $m = h$, т.е.

$$\int \lambda_k^{(h-1)} f = -\frac{1}{h} \int l^{(h+1)} f. \quad (36)$$

Дифференцируем равенство (34) $h-1$ раз при $m = 1$, $h-2$ раза при $m = 2$ и т.д. (т.е. при $m = 1, \dots, h$ равенство (34) дифференцируем $h-m$ раз). В итоге, полагая $C_0^0 = 1$, получаем

$$\sum_{i=1}^{h-j} C_{h-j}^i \int \lambda_k^{(i)} f^{(h-1-i)} = \frac{1}{j} \sum_{i=0}^{h-j} C_{h-j}^i \int l^{(i+1)} f^{(h-i)}, \quad j = 1, \dots, h. \quad (37)$$

При выводе (37) мы учли, что в силу предположения ($\mathbf{E}_k$) можно менять порядок дифференцирования и интегрирования в соответствующих интегралах, а также использовали следующее представление для n -й производной произведения двух функций:

$$(uv)^{(n)} = \sum_{i=0}^n C_n^i u^{(i)} v^{(n-i)} \quad \text{при} \quad C_n^i = \frac{n!}{i!(n-i)!}. \quad (38)$$

Положим

$$x_i := \int \lambda_k^{(i)} f^{(h-1-i)}, \quad \nu_i := \int l^{(i+1)} f^{(h-i)}, \quad b_{h-j} := \frac{1}{j} \sum_{i=0}^{h-j} C_{h-j}^i \nu_i \quad (39)$$

при $i = 0, \dots, h-1$ и $j = 1, \dots, h$. В этом случае (37) примет следующий вид:

$$x_{h-j} + C_{h-j}^{h-j-1} x_{h-j-1} + \dots + C_{h-j}^2 x_2 + C_{h-j}^1 x_1 + x_0 = b_{h-j}, \quad j = 1, \dots, h, \quad (40)$$

а равенство (36), которое нужно установить, можно переписать в виде

$$x_{h-1} = -\frac{\nu_h}{h}. \quad (41)$$

Нетрудно видеть, что соотношения (40) — это система линейных уравнений (с верхнетреугольной матрицей) относительно $x_0, \dots, x_{h-1}$. Используя обратный ход метода Гаусса и последовательно выражая $x_0 = b_0$, $x_1 = b_1 - b_0$, $x_2 = b_2 - 2b_1 + b_0$ и т.д., из (40) находим представление для x_{h-1} :

$$x_{h-1} = b_{h-1} - C_{h-1}^{h-2} b_{h-2} + C_{h-1}^{h-3} b_{h-3} - \dots \mp C_{h-1}^1 b_1 \pm b_0. \quad (42)$$

Чтобы теперь из (42) извлечь (41) и завершить доказательство леммы, нам потребуется несколько комбинаторных соотношений.

Во-первых, при любом $i = 1, \dots, h-1$

$$C_h^1 C_{h-1}^i - C_h^2 C_{h-2}^i + C_h^3 C_{h-3}^i - \dots \pm C_h^{h-i} C_i^i = C_h^0 C_h^i \quad (43)$$

(это соотношение содержится, например, в [11, с. 81]). Во-вторых, имеют место равенства

$$\sum_{i=0}^{h-1} C_h^i \nu_i = -\nu_h, \quad \frac{1}{j} C_{h-1}^{h-j} = \frac{1}{h} C_h^j. \quad (44)$$

Второе равенство в (44) нетрудно извлечь непосредственно из определения в (38), а чтобы получить первое равенство в (44), нужно h раз продифференцировать тождество $\int l' f = 0$ по θ , учесть, что в силу предположения (**Е**_k) можно менять порядок дифференцирования и интегрирования, а также использовать соотношение (38) и обозначения (39).

Наконец, подставляя значения b_{h-j} в соотношение (42) и используя тождества (43) и (44), получаем

$$\begin{aligned} x_{h-1} &= C_{h-1}^{h-1} \sum_{i=0}^{h-1} C_{h-1}^i \nu_i - \frac{1}{2} C_{h-1}^{h-2} \sum_{i=0}^{h-2} C_{h-2}^i \nu_i \\ &\quad + \frac{1}{3} C_{h-1}^{h-3} \sum_{i=0}^{h-3} C_{h-3}^i \nu_i - \dots \pm \frac{1}{h} \nu_0 \\ &= \frac{1}{h} \left(C_h^1 \sum_{i=0}^{h-1} C_{h-1}^i \nu_i - C_h^2 \sum_{i=0}^{h-2} C_{h-2}^i \nu_i + C_h^3 \sum_{i=0}^{h-3} C_{h-3}^i \nu_i - \dots \pm \nu_0 \right) \\ &= \frac{1}{h} \sum_{i=0}^{h-1} \left(\sum_{j=1}^{h-i} (-1)^{j+1} C_h^j C_{h-j}^i \right) \nu_i = \frac{1}{h} \sum_{i=0}^{h-1} C_h^0 C_h^i \nu_i = -\frac{1}{h} \nu_h. \end{aligned}$$

Таким образом, выполнено (41) и доказательство леммы завершено.

Лемма 4.5. Если выполнено условие $(\mathbf{E}_k)$, то функция $\lambda_k(\theta, x)$, определенная равенством (22), является решением системы (23).

Доказательство. В силу условий леммы и определения (20), нам заданы линейно независимые функции $e_{0,\theta}(\cdot), \dots, e_{k,\theta}(\cdot)$, которые будем называть базисом, и известны проекции искомой функции $\lambda_k(\theta, \cdot)$ на элементы этого базиса, определяемые соотношениями (23):

$$\langle \lambda_k(\theta, \cdot), e_{m-1,\theta}(\cdot) \rangle = \frac{1}{m} \langle e_{1,\theta}(\cdot), e_{m,\theta}(\cdot) \rangle \equiv a_{m-1,\theta}, \quad m = 1, \dots, k+1. \quad (45)$$

Хорошо известно, что в этом случае функция

$$\lambda_k(\theta, x) = c_{k0,\theta} e_{0,\theta}(x) + c_{k1,\theta} e_{1,\theta}(x) + \dots + c_{kk,\theta} e_{k,\theta}(x), \quad (46)$$

являющаяся линейной комбинацией элементов базиса со специально подобранными коэффициентами $c_{kj,\theta}$, $j = 0, \dots, k$, удовлетворяет системе (45). Чтобы определить эти коэффициенты, домножим скалярно равенство (46) на элементы базиса $e_{0,\theta}(\cdot), \dots, e_{k,\theta}(\cdot)$. Учитывая еще соотношения (45), получаем следующую линейную систему $(k+1)$ уравнения относительно $(k+1)$ неизвестной $c_{kj,\theta}$, $j = 0, \dots, k$:

$$c_{k0,\theta} \langle e_{0,\theta}(\cdot), e_{j,\theta}(\cdot) \rangle + \dots + c_{kk,\theta} \langle e_{k,\theta}(\cdot), e_{j,\theta}(\cdot) \rangle = a_{j,\theta}, \quad j = 0, \dots, k. \quad (47)$$

При этом матрица с элементами $\langle e_{i,\theta}(\cdot), e_{j,\theta}(\cdot) \rangle$, $i, j = 0, \dots, k$, определяющая эту систему линейных уравнений (матрица Грама, построенная по элементам базиса), положительно определена ввиду линейной независимости этих элементов.

В силу равенств $l'(\theta, x) = f'_\theta(x)/f_\theta(x)$, $l''(\theta, x) = f''_\theta(x)/f_\theta(x) - (f'_\theta(x)/f_\theta(x))^2$ и условия $(\mathbf{B})$ получаем, что $\int f'_\theta(x) \mu(dx) = 0$ и $\int f''_\theta(x) \mu(dx) = 0$. А потому, с учетом обозначений (19) и (20),

$$\langle e_{1,\theta}(\cdot), e_{1,\theta}(\cdot) \rangle \equiv a_{0,\theta} = I(\theta), \quad \langle e_{0,\theta}(\cdot), e_{0,\theta}(\cdot) \rangle \equiv \int f_\theta(x) \mu(dx) = 1,$$

$$\langle e_{j,\theta}(\cdot), e_{0,\theta}(\cdot) \rangle = \langle e_{0,\theta}(\cdot), e_{j,\theta}(\cdot) \rangle = \int f_\theta^{(j)}(x) \mu(dx) = 0 \quad \text{при } j \neq 0.$$

Следовательно, у матрицы Грама, определяющей систему уравнений (47), все элементы первой строки и первого столбца, за исключением диагонального — нулевые. А потому $c_{k0,\theta} = I(\theta)$ для любых k и функция $\lambda_k(\theta, x)$, определенная в (21) и (22), является решением (23). Лемма доказана.

Доказательство теоремы 2.2. Это утверждение немедленно следует из лемм 4.4 и 4.5.

4.3. **Доказательство предложения 3.1.** Заметим, что $\mathbf{E}_\theta X_1 < \infty$ при $h > 0$ и $\mathbf{E}_\theta X_1^2 = \infty$ при $h < 1$. Согласно [12, с. 643]

для сходимости соответствующих распределений к устойчивому закону нужно, чтобы выполнялось соотношение

$$\frac{n\mu(a_n)}{a_n^2} \rightarrow \text{const}, \quad \text{где} \quad \mu(x) = \int_{-x}^x y^2 f_\theta(y) dy \quad \text{и} \quad a_n = n^{1-\alpha}.$$

Учитывая (25), имеем

$$\mu(x) = (h+1) \int_0^x \frac{y^2/\theta}{(1+y/\theta)^{h+2}} dy = C_0 \int_1^{x+1} \frac{(y-1)^2}{y^{h+2}} dy \sim C_1 x^{-h+1}.$$

Следовательно,

$$\frac{n\mu(a_n)}{a_n^2} \sim n^{(1-\alpha)(1-h)+1-2(1-\alpha)} \sim \text{const},$$

если только $-(1-\alpha)(1+h)+1=0$. Из последнего соотношения находим, что $\alpha = h/(h+1)$ и $\alpha \leq 1/4$ при $h \leq 1/3$.

СПИСОК ЛИТЕРАТУРЫ

1. Fisher R. A. Theory of statistical estimation. — Proc. Cambridge Philos. Soc., 1925, v. 22, p. 700–725.
2. Боровков А. А. Математическая статистика. Новосибирск: Изд-во ИМ СО РАН, 1997, 771 с.
3. Закс Ш. Теория статистических выводов. М.: Мир, 1975, 776 с.
4. Леман Э. Теория точечного оценивания. М.: Наука, 1991, 444 с.
5. Le Cam L. On the asymptotic theory of estimation and testing hypotheses. — Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability. V. 1: Contributions to the Theory of Statistics. Berkeley: Univ. of California Press, 1956, p. 129–156.
6. Verrill S. Rate of convergence of k -step Newton estimators to efficient likelihood estimators. — Statist. Probab. Lett., 2007, v. 77, № 12, p. 1371–1376.
7. Janssen P., Jurečková J., Veraverbeke N. Rate of convergence of one- and two-step M -estimators with applications to maximum likelihood and Pitman estimators. — Ann. Statist., 1985, v. 13, № 3, p. 1222–1229.
8. Jurečková J., Sen P. K. Effect of the initial estimator on the asymptotic behavior of one-step M -estimator. — Ann. Inst. Statist. Math., 1990, v. 42, № 2, p. 345–357.
9. Linke Yu. Yu. Asymptotic properties of one-step M -estimators based on nonidentically distributed observations and applications to some regression problems. arXiv: 1503.03393v5.
10. Линке Ю. Ю., Саханенко А. И. Об условиях асимптотической нормальности одношаговых оценок Фишера для однопараметрических семейств распределений. — Сиб. электрон. матем. изв., 2014, т. 11, с. 464–475.
11. Феллер В. Введение в теорию вероятностей и ее приложения, т. 1. М.: Мир, 1984, 527 с.
12. Феллер В. Введение в теорию вероятностей и ее приложения, т. 2. М.: Мир, 1984, 751 с.

Поступила в редакцию
3.III.2014