S©MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru

Том 11, стр. 464–475 (2014) УДК 519.233.22
MSC 62F12

ОБ УСЛОВИЯХ АСИМПТОТИЧЕСКОЙ НОРМАЛЬНОСТИ ОДНОШАГОВЫХ ОЦЕНОК ФИШЕРА ДЛЯ ОДНОПАРАМЕТРИЧЕСКИХ СЕМЕЙСТВ РАСПРЕДЕЛЕНИЙ

Ю.Ю. ЛИНКЕ, А.И.САХАНЕНКО

Abstract. We consider asymptotic behavior of one-step statistical estimators introduced by R. Fisher as approximations for consistent maximum likelihood estimators. Some sufficient conditions are found for these one-step estimators to be asymptotically normal even in the cases when either the maximum likelihood estimators may not exist or exist but be inconsistent. Investigated are connections between the smoothness conditions for the density of the sample distribution and the rate of proximity of the preliminary estimator and the parameter which are needed for fulfillment of the properties under considerations.

Keywords: one-step estimators, asymptotical normality, maximum likelihood estimator, Newton's method, preliminary estimator, proximity of estimation.

1. Введение

1.1. Пусть $X_1, \ldots, X_n$ — независимые наблюдения произвольной природы, общее распределение которых имеет плотность $f(\theta, x)$ относительно некоторой меры. Рассматривается задача оценивания неизвестного числового параметра $\theta \in \Theta \subset \mathbb{R}$. В широких условиях некоторым эталоном точности при оценивании этого неизвестного θ принято считать ОМП — оценку максимального правдоподобия $\widehat{\theta}_n$. Однако известны примеры распределений, когда отыскать ОМП затруднительно (например, в случае существования большого числа локальных максимумов у логарифмической функции правдоподобия), но можно вычислить (например, по методу моментов) некоторую достаточно хорошую

Linke, Yu.Yu. and Sakhanenko A.I., On conditions for asymptotic normality of Fisher's one-step estimators in one-parameter families of distributions.

Работа поддержана РФФИ (грант 14-01-00220).

Поступила 14 марта 2014 г., опубликована 16 июня 2014 г.

(скажем, состоятельную) оценку θ_n^* для θ. Для таких ситуаций Р. Фишером в [1] была предложена одношаговая процедура приближенного вычисления ОМП, основанная на методе Ньютона. Начиная с работ Р.Фишера, одну из следующих двух статистик

$$\theta_{n,L}^{**} = \theta_n^* - \frac{L_n'(\theta_n^*)}{L_n''(\theta_n^*)} \qquad \text{и} \qquad \theta_{n,I}^{**} = \theta_n^* + \frac{L_n'(\theta_n^*)}{nI(\theta_n^*)} \tag{1}$$

обычно используют в качестве приближений для ОМП. Здесь θ_n^* — некоторая начальная оценка параметра θ, $I(t)$ — информация Фишера, соответствующая плотности $f(t,x)$, а $L_n(t)$ — логарифмическая функция правдоподобия, построенная по выборке $X_1,\dots,X_n$, т.е.

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

Естественным образом возникает вопрос об условиях асимптотической нормальности и асимптотической эффективности оценок из (1), т.е. вопрос об условиях, достаточных для справедливости следующих сходимостей по распределению:

$$\sqrt{nI(\theta)}\big(\theta_{n,L}^{**} - \theta\big) \Rightarrow \mathcal{N}(0,1) \quad \text{при} \quad n\to\infty, \tag{3}$$

$$\sqrt{nI(\theta)}\big(\theta_{n,I}^{**} - \theta\big) \Rightarrow \mathcal{N}(0,1) \quad \text{при} \quad n\to\infty. \tag{4}$$

Эта задача рассматривалась, например, в монографиях [2], [3]. При этом в большинстве известных авторам работ при доказательстве (3) и (4) (как и при доказательстве асимптотической эквивалентности $\widehat{\theta}_n$ и оценок из (1)) на предварительную оценку θ_n^* налагается очень жесткое ограничение:

$$\sqrt{n}(\theta_n^* - \theta) = O_p(1) \quad \text{при} \quad n\to\infty, \tag{5}$$

которое мы будем называть $\sqrt{n}$-ограниченностью оценки θ_n^*. Исключением является лишь монография [2; теорема 5.5.4], где при выводе (3) предполагается $n^{1/4}$-состоятельность оценки θ_n^*, т.е. следующая сходимость:

$$n^{1/4}(\theta_n^* - \theta) \xrightarrow{p} 0 \quad \text{при} \quad n\to\infty. \tag{6}$$

1.2. Для удобства дальнейших ссылок введем два предположения. Подчеркнем, что далее штрихи обозначают дифференцирование по параметру θ или по переменной t, которая часто заменяет этот параметр в функциях из (2).

$(R_{1,\theta})$. Наблюдаются первые n элементов из последовательности независимых одинаково распределенных случайных величин X_1, $X_2,\dots$ со значениями в произвольном измеримом пространстве $(\mathcal{X},\mathcal{A})$ с плотностью $f(\theta,x)$ относительно некоторой меры μ на $\mathcal{A}$, где неизвестный параметр θ может принимать значения из некоторого открытого множества $\Theta\subset\mathbb{R}$. Кроме того, для μ-почти всех $x\in\mathcal{X}$ функция $l(t,x)=\ln f(t,x)$ дважды дифференцируема по t при всех $t\in\Theta$, причем

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

$(R_{2,\theta})$. Для μ-почти всех $x\in\mathcal{X}$ функция $l(t,x)$ дважды непрерывно дифференцируема по t при всех $t\in\Theta$, причем при некотором $\delta_\theta>0$

$$\mathbf{E}_\theta H_2(\theta,X_1) < \infty, \quad \text{где} \quad H_2(\theta,x) := \sup\{|l''(t,x)| : t\in[\theta-\delta_\theta,\theta+\delta_\theta]\cap\Theta\}. \tag{8}$$

Подчеркнем, что всюду в работе θ — это неизвестный параметр, а потому при практических применениях утверждений работы мы должны проверять выполнение их условий при всех возможных значениях этого неизвестного параметра. Но чтобы не загромождать утверждения очевидными оговорками, мы формулируем их при некотором фиксированном значении этого параметра θ.

Теорема A. *Пусть выполнены предположения* $(R_{1,\theta})$ *и* $(R_{2,\theta})$ *и верно условие* (5). *Тогда имеет место сходимость* (3). *Если же, дополнительно, функция* $I(t)$ *непрерывна в точке* $t = \theta$, *то также справедлива и сходимость* (4).

Отметим, что в известной авторам литературе в формулировках аналогичных утверждений всегда присутствует еще целый ряд условий (см., например, [2] – [5], а также ссылки в [6]). Во многом это связано с тем, что оценки (1) рассматриваются как приближения для ОМП $\widehat{\theta}_n$. И в первую очередь вместо (3) часто доказывается следующая сходимость:

$$\sqrt{n}\big(\theta^{**}_{n,L} - \widehat{\theta}_n\big) \xrightarrow{p} 0, \tag{9}$$

что заставляет вводить целый ряд ограничений, нужных для существования ОМП и для ее асимптотической нормальности. Тем не менее, как будет показано в пункте 2.2 (следствие 3), теорема A справедлива и без дополнительных условий.

Отметим еще, что в монографии [3] аналог теоремы A содержит следующее восходящее к Г.Крамеру и более жесткое, чем $(R_{2,\theta})$, условие.

$(R_{3,\theta})$. Для μ-почти всех $x \in \mathcal{X}$ функция $l(t,x)$ трижды дифференцируема по t при всех $t \in \Theta$ и

$$\mathbf{E}_\theta H_3(\theta, X_1) < \infty, \quad \text{где} \quad H_3(\theta,x) := \sup\{|l'''(t,x)| : t \in [\theta - \delta_\theta, \theta + \delta_\theta] \cap \Theta\} \tag{10}$$

для некоторого $\delta_\theta > 0$.

Но если выполнено это более жесткое условие на гладкость плотности, то при получении сходимости (3) можно ослабить ограничение (5) на точность сближения предварительной оценки θ^* и неизвестного параметра θ.

Теорема B. *Пусть справедливы предположения* $(R_{1,\theta})$, $(R_{3,\theta})$ *и выполнено условие* (6). *Тогда имеет место сходимость* (3).

В замечании 3 мы сравним условия теоремы B с предположениями, использовавшимися при выводе аналогичной теоремы 5.5.4 из [2].

1.3. Асимптотическое поведение рассматриваемых одношаговых приближений Фишера (1) существенно зависит как от точности предварительной оценки θ^*_n, так и от гладкости функции $l(t,x)$. При этом эти две характеристики взаимосвязаны в том смысле, что чем хуже оценка θ^*_n приближает неизвестный параметр θ, тем более гладкой должна быть указанная функция $l(t,x)$, если мы хотим, чтобы имела место сходимость (3) (или (4)). Так в теореме A на гладкость функции $l(t,x)$ налагается минимальное ограничение: требуется, по сути, лишь непрерывность ее второй производной $l''(t,x)$. A в теореме B другая крайность: число производных у функции $l(t,x)$ возрастает до трех, но при этом ослабляются условия на точность оценки θ^*_n с $\sqrt{n}$-ограниченности (5) до $n^{1/4}$-состоятельности (6).

Первой целью настоящей статьи является получение достаточных условий для сходимостей (3) и (4) в более широком спектре возможных ограничений на гладкость функции $l(t,x)$. Мы будем предполагать, что вторые производные $l''(t,x)$ удовлетворяют условию Гельдера с некоторым показателем α и исследуем (см. теоремы 1-2 и следствия 1-2) вопрос о том, какие в этом случае придется наложить требования на точность оценок θ_n^*. В качестве двух крайних частных случаев (см. следствия 3 и 4) мы получим утверждения теорем A и B, причем при чуть более слабых предположениях.

В доказательствах мы используем восходящую к Р. Фишеру и Г. Крамеру идею приближения функции правдоподобия отрезком ряда Тейлора. Второй целью данной работы (см. теорему 3 и следствие 5) является доказательство того факта, что этим способом не возможно ослабить условие (6) на точность оценки θ_n^* без введения дополнительных ограничений на класс плотностей.

2. Основные результаты

2.1. В этом пункте мы приведем наиболее общие достаточные условия асимптотической нормальности оценок $\theta_{n,L}^{**}$ и $\theta_{n,I}^{**}$. При $\delta > 0$ и $\alpha \geq 0$ положим

$$h_{\alpha,\theta}(\delta, x) := \sup\left\{\frac{|l''(t,x) - l''(\theta,x)|}{|t-\theta|^\alpha} : t \in [\theta-\delta, \theta+\delta] \cap \Theta,\ t \neq \theta\right\}, \tag{11}$$

$$\overline{L_n''}(\theta) := L_n''(\theta)/n + I(\theta), \qquad \beta(\alpha) = 1/(2+2\alpha). \tag{12}$$

Отметим, что всюду далее $0 \leq \alpha \leq 1$ и $1/4 \leq \beta(\alpha) \leq 1/2$.

Теорема 1. *Пусть выполнено предположение* $(R_{1,\theta})$ *и, кроме того,*

$$\mathbf{E}_\theta h_{\alpha,\theta}(\delta, X_1) \to 0 \quad \text{при} \quad \delta \to 0, \quad \text{где} \quad 0 \leq \alpha < 1, \tag{13}$$

$$n^{\beta(\alpha)}(\theta_n^* - \theta) = O_p(1) \quad \text{при} \quad n \to \infty. \tag{14}$$

Тогда при $n \to \infty$

$$\Delta_{n,L} := \sqrt{n}\big(\theta_{n,L}^{**} - \theta\big)I(\theta) - L_n'(\theta)/\sqrt{n} \xrightarrow{p} 0, \tag{15}$$

$$\begin{aligned}\Delta_{n,I}^* := {}& \sqrt{n}\big(\theta_{n,I}^{**} - \theta\big)I(\theta^*) - L_n'(\theta)/\sqrt{n} \\ & -\sqrt{n}(\theta_n^* - \theta)\big(I(\theta^*) - I(\theta) + \overline{L_n''}(\theta)\big) \xrightarrow{p} 0.\end{aligned} \tag{16}$$

В частности, при этих условиях из (15) *вытекает сходимость* (3).

Следствие 1. *Пусть верны все условия теоремы* 1 *и дополнительно*

$$I(t) - I(\theta) = o(|t-\theta|^\alpha) \quad \text{при} \quad t \to \theta, \tag{17}$$

$$n^{\alpha\beta(\alpha)}\overline{L_n''}(\theta) \xrightarrow{p} 0 \quad \text{при} \quad n \to \infty. \tag{18}$$

Тогда имеет место также и сходимость (4).

Теорема 2. *Пусть справедливо предположение* $(R_{1,\theta})$ *и, кроме того,*

$$\mathbf{E}_\theta h_{\alpha,\theta}(\delta_\theta, X_1) < \infty \quad \text{при некотором} \quad \delta_\theta > 0, \quad \text{где} \quad 0 < \alpha \leq 1, \tag{19}$$

$$n^{\beta(\alpha)}(\theta_n^* - \theta) \xrightarrow{p} 0 \quad \text{при} \quad n \to \infty. \tag{20}$$

Тогда имеют место сходимости (15), (16) *и* (3).

Следствие 2. *Пусть выполнены все условия теоремы* 2 *и*

$$I(t) - I(\theta) = O(|t-\theta|^\alpha) \quad \text{при} \quad t \to \theta, \tag{21}$$

$$n^{\alpha\beta(\alpha)}\overline{L_n''}(\theta) = O_p(1) \quad \text{при} \quad n \to \infty. \tag{22}$$

Тогда справедлива сходимость (4).

З а м е ч а н и е 1. Как будет показано в пункте 3.4, для справедливости условия (22) достаточно выполнения следующего простого предположения:

$$\mathbf{E}_\theta|l''(\theta, X_1)|^{\gamma(\alpha)} < \infty \quad \text{при} \quad \gamma(\alpha) = \frac{2+2\alpha}{2+\alpha}. \tag{23}$$

2.2. Рассмотрим теперь вопрос о справедливости теорем A и B из введения. Прежде всего заметим, что при $\alpha = 0$ утверждение теоремы 1 и следствия 1 можно переписать в следующем виде.

Следствие 3. *Пусть выполнено предположение* $(R_{1,\theta})$, *условие* (5) *и*

$$\mathbf{E}_\theta \sup\{|l''(t,x) - l''(\theta,x)| : t \in [\theta-\delta, \theta+\delta] \cap \Theta\} \to 0 \quad \textit{при} \quad \delta \to 0. \tag{24}$$

В этом случае имеют место сходимости (15), (16) *и* (3). *Если же, дополнительно, функция* $I(t)$ *непрерывна в точке* $t = \theta$, *то справедливо также* (4).

З а м е ч а н и е 2. Если верно предположение $(R_{2,\theta})$, то функция $l''(t,x)$ непрерывна по t и условие (24) выполнено в силу теоремы Лебега о мажорируемой сходимости под знаком интеграла. Значит, теорема A является частным случаем следствия 3.

Из теоремы 2 при $\alpha = 1$ и $\beta(\alpha) = 1/4$ вытекает

Следствие 4. *Пусть верно предположение* $(R_{1,\theta})$ *и*

$$\mathbf{E}_\theta h_{1,\theta}(\delta_\theta, X_1) < \infty \quad \textit{при некотором} \quad \delta_\theta > 0. \tag{25}$$

В этом случае для справедливости сходимостей (15) *и* (3) *достаточно выполнения условия* (6).

Поскольку условие (25) слабее, чем предположение $(R_{3,\theta})$, то теорема B является очевидным частным случаем следствия 4. Однако, если выполнено предположение $(R_{3,\theta})$, то утверждение теоремы B можно усилить следующим образом.

Теорема 3. *Пусть верны предположения* $(R_{1,\theta})$ *и* $(R_{3,\theta})$, *а оценка* θ_n^* — *состоятельна. Тогда при* $n \to \infty$ *справедливо следующее соотношение:*

$$\Delta_{n,L} = \sqrt{n}(\theta_n^* - \theta)^2\big(-\mathbf{E}_\theta l'''(\theta, X_1)/2 + o_p(1)\big) + o_p(1). \tag{26}$$

Следствие 5. *Пусть выполнены все условия теоремы* 3 *и*

$$\mathbf{E}_\theta l'''(\theta, X_1) \neq 0. \tag{27}$$

В этом случае условие (6) *является необходимым и достаточным для того, чтобы сходимость* (15) *имела место.*

Таким образом, метод доказательства, используемый в настоящей работе (как и аналогичные методы из, например, работ [2]–[4]) требует от предварительной оценки θ_n^* как минимум ее $n^{1/4}$-состоятельность, т.е. сходимость (6).

2.3. В этом пункте мы приведем несколько примеров и замечаний. Как отмечено во введении, при использовании одношаговых оценок Фишера обычно в качестве предварительной оценки рассматривают $\sqrt{n}$-ограниченную оценку. Приведем пример, когда оценка θ_n^* медленнее сближается с параметром θ.

Пример. Пусть общее распределение случайных величин $X_1, X_2, \ldots$ сосредоточено на положительной полуоси и имеет следующую плотность:

$$f(\theta, x) = \frac{(h+1)}{\theta(1+x/\theta)^{2+h}} \qquad \text{при} \quad h > 0 \quad \text{и} \quad x > 0. \tag{28}$$

В этом случае вычислить оценку максимального правдоподобия затруднительно, но оценку метода моментов по первому моменту найти легко:

$$\theta_n^* = h\sum_{i=1}^n X_i/n \qquad \text{и} \qquad \mathbf{E}_\theta\theta_n^* = h\mathbf{E}_\theta X_1 = \theta.$$

Из (28) также нетрудно извлечь, что $\forall r < 1+h$ $\mathbf{E}_\theta X_1^r < \infty$ и $\mathbf{E}_\theta X_1^{1+h} = \infty$. Значит $\mathbf{D}_\theta X_1 = \infty$ при $h < 1$ и в этом случае оценка $\theta_n^* = h\sum_{i=1}^n X_i/n$ не является $\sqrt{n}$-ограниченной.

С другой стороны, из неравенства Бара–Эссена (см, например, [7], стр. 79) при $1 \le r < 1+h < 2$ имеем:

$$n^{r-1}\mathbf{E}_\theta|\theta_n^* - \theta|^r = h^r n^{-1}\mathbf{E}_\theta\Big|\sum_{i=1}^n (X_i - \mathbf{E}_\theta X_i)\Big|^r \le 2h^r\mathbf{E}_\theta|X_1 - \mathbf{E}_\theta X_1|^r < \infty.$$

Следовательно,

$$\theta_n^* - \theta = O_p(n^{-(1-1/r)}) = o_p(n^{-1/4}) \quad \text{при} \quad 4/3 < r < 1+h < 2. \tag{29}$$

Таким образом, при $1/3 < h < 1$ выборка с плотностью из (28) обладает теми свойствами, что предварительная оценка не является $\sqrt{n}$-ограниченной, но удовлетворяет условию (6). Кроме того, в этом случае условия (14) и (20) выполнены при всех $\beta(\alpha) \in \big(1/4, h/(1+h)\big)$ в силу (29).

З а м е ч а н и е 3. В монографии [2, теорема 5.5.4] утверждается, что для сходимости (3) в случае, когда предварительная оценка $n^{1/4}$-состоятельна, достаточно выполнения условий теоремы 5.5.3 из [2]. При этом в доказательстве теоремы 5.5.4. (см. формулу (5.5.19) в [2]) используется соотношение, которое в наших обозначениях можно переписать следующим образом:

$$L_n''(\theta_n^*)/n = L_n''(\theta)/n + O_p(\theta_n^* - \theta). \tag{30}$$

Далее, условия упомянутой теоремы 5.5.3 из [2] отличаются от предположений нашего следствия 4 заменой условия (25) на следующее более слабое ограничение:

(31)

$$\sup_{t:\ |t-\theta|<\delta}\big|l''(t,X_1) - l''(\theta,X_1)\big| \le H(\theta,X_1)o_p(1) \quad \text{при} \quad \delta\to 0, \quad \mathbf{E}_\theta H(\theta,X_1) < \infty.$$

Однако перед теоремой 5.5.3 в [2] утверждается, что из ее условий вытекает лишь следующее более слабое, чем (30), соотношение $L_n''(\theta_n^*)/n = -I(\theta) + o_p(1)$ (см. формулу (5.5.11) в [2]), которое можно переписать в виде:

$$L_n''(\theta_n^*)/n = L_n''(\theta)/n + o_p(1). \tag{32}$$

Такое противоречие между формулами (30) и (32) наводит на мысль, что вместо условия (31) в формулировке теоремы 5.5.4 в [2] должно стоять более жесткое предположение. Например, наше условие (25) из следствия 4 достаточно для получения соотношения (30) вместо (32).

З а м е ч а н и е 4. Поскольку для μ-почти всех $x \in \mathcal{X}$ функция $l(t,x) = \ln f(t,x)$ дважды дифференцируема по t при всех $t \in \Theta$ в силу предположения $(R_{1,\theta})$, то для всех таких x она определена, конечна и $f(t,x) > 0$. Но это означает, что

$$\mu(\mathcal{X}_0) = 0, \quad \text{где} \quad \mathcal{X}_0 := \{x \in \mathcal{X} : f(t,x) = 0 \text{ хотя бы при одном } t \in \Theta\}.$$

Значит, множество $\mathcal{X}$ может отличаться от носителя распределения $\mathbf{P}_t$ лишь на подмножество множества $\mathcal{X}_0$ меры нуль.

Таким образом, предположение $(R_{1,\theta})$ включает в себя, в неявном виде, условие о совпадении носителей всех неизвестных распределений $\mathbf{P}_t$, $t \in \Theta$. Впрочем, аналогичные рассуждения показывают, что во всех работах [2]-[6] эти носители также совпадают, даже если это не указано в явном виде, как в [6].

Замечание 5. Подчеркнем, что все используемые нами ограничения — локальные, т.е. каждый раз ограничения накладываются лишь на поведение плотностей в некоторой окрестности $[\theta-\delta_\theta, \theta+\delta_\theta]\cap\Theta$ неизвестного параметра θ. В частности, мы допускаем, что вне этой окрестности функция правдоподобия может неограниченно возрастать и что могут существовать распределения $\mathbf{P}_t$, которые совпадают с $\mathbf{P}_\theta$ при $|t-\theta| > \delta_\theta$. Тем самым мы допускаем как возможность несуществования оценки максимального правдоподобия $\widehat{\theta}_n$ (когда, например, функция правдоподобия стремится к бесконечности с ростом параметра), так и возможность, что она существует, но не состоятельна (например, если распределения $\mathbf{P}_t$ периодически повторяются, то существует бесконечное число ОМП, из которых мы не сможем выбрать состоятельную).

Замечание 6. Как уже отмечалось во введении, всюду в работе θ — это неизвестный параметр. При практических применениях утверждений работы мы должны проверять выполнение их условий при всех значениях $\theta \in \Theta$ (т.е. так же, как это всегда делается в математической статистике (см., например, [4])).

3. Доказательства

3.1. Далее в работе, не оговаривая этого больше, мы считаем, что все пределы берутся при $n \to \infty$ и предполагаем, что справедливо условие $(R_{1,\theta})$, а оценка θ_n^* — состоятельна, поскольку эти два ограничения выполнены во всех доказываемых утверждениях. В этом случае обе изучаемые оценки $\theta_{n,L}^{**}$ и $\theta_{n,I}^{**}$ определены с вероятностями, стремящимися к 1, а их определения (1) можно переписать в следующем виде:

(33)
$$\theta_{n,L}^{**} - \theta = \frac{L_n'(\theta_n^*) - (\theta_n^* - \theta)L_n''(\theta_n^*)}{-L_n''(\theta_n^*)} \quad \text{и} \quad \theta_{n,I}^{**} - \theta = \frac{L_n'(\theta_n^*) + nI(\theta_n^*)(\theta_n^* - \theta)}{nI(\theta_n^*)}.$$

Введем обозначения:

$$\delta_n^* := \theta_n^* - \theta \quad \text{и} \quad \rho_n(\delta) := \big(L_n''(\theta+\delta) - L_n''(\theta)\big)/n. \tag{34}$$

Из (15), (16), (33) и (34) нетрудно извлечь равенства:

$$\Delta_{n,I}^* = \sqrt{n}\overline{\rho}_n(\delta_n^*) \quad \text{при} \quad \overline{\rho}_n(\delta) := \big(L_n'(\theta+\delta) - L_n'(\theta) - \delta L_n''(\theta)\big)/n, \tag{35}$$

$$\Delta_{n,L} = \frac{L_n'(\theta)/\sqrt{n} + \sqrt{n}\overline{\rho}_n(\delta_n^*) - \sqrt{n}\delta_n^*\rho_n(\delta_n^*)}{1 - \overline{L_n''}(\theta)/I(\theta) - \rho_n(\delta_n^*)/I(\theta)} - \frac{L_n'(\theta)}{\sqrt{n}}, \tag{36}$$

где величина $\overline{L_n''}(\theta)$ была определена в (12).

Представления (35) и (36) будут играть ключевую роль в доказательствах. Нам также потребуются две сходимости

$$L_n(\theta)/\sqrt{n} \Rightarrow \mathcal{N}(0, I(\theta)) \quad \text{и} \quad \overline{L_n''}(\theta) \xrightarrow{p} 0, \tag{37}$$

вытекающие из предположения $(R_{1,\theta})$, центральной предельной теоремы и закона больших чисел. А для оценки погрешностей нам потребуются следующие утверждения.

Лемма 1. *Если выполнено условие* (13), *то*

$$\lambda_n := \sup_{v\in[0,1]} |\rho_n(v\delta_n^*)| = |\delta_n^*|^\alpha o_p(1) \tag{38}$$

и справедливо следующее равенство:

$$\Delta_{n,L} = \rho_n^*(\delta_n^*)\big(1 + o_p(1)\big) + o_p(1) \quad \text{\textit{при}} \quad \rho_n^*(\delta) = \sqrt{n}\,\overline{\rho}_n(\delta) - \sqrt{n}\,\delta\rho_n(\delta). \tag{39}$$

Если же верно условие (19), *то* $\lambda_n = |\delta_n^*|^\alpha O_p(1)$ *и равенство* (39) *также верно.*

Лемма 2. *Если верно условие* $(R_{3,\theta})$, *то*

$$\overline{\lambda}_n := \sup_{v\in[0,1]} |\rho_n(v\delta_n^*) - v\delta_n^* l| = |\delta_n^*|\, o_p(1) \quad \text{\textit{при}} \quad l := \mathbf{E}_\theta l'''(\theta, X_1). \tag{40}$$

Эти основные леммы будут доказаны в пунктах 3.2 и 3.3. Кроме того, нам понадобятся следующие равенства

$$\sqrt{n}\,|\delta_n^*|^{1+\alpha} = \big(n^{\beta(\alpha)}|\delta_n^*|\big)^{1+\alpha} \quad \text{и}$$
$$\overline{\rho}_n(\delta) = \delta\int_0^1 L_n''(\theta + v\delta)dv - \delta L_n''(\theta) = \delta\int_0^1 \rho_n(v\delta)dv. \tag{41}$$

Д о к а з а т е л ь с т в о теоремы 1. Из (38) и (41) вытекает, что в условиях теоремы

$$\begin{aligned}\sqrt{n}\,\big|\overline{\rho}_n(\delta_n^*)\big| + \sqrt{n}\,\big|\delta_n^*\rho(\delta_n^*)\big| &\le 2\sqrt{n}|\delta_n^*|\lambda_n = \\ &= \big(n^{(\beta(\alpha)}|\delta_n^*|\big)^{1+\alpha} o_p(1) = \big(O_p(1)\big)^{1+\alpha} o_p(1) \underset{p}{\to} 0.\end{aligned} \tag{42}$$

Из этого соотношения и представления (35) вытекает сходимость (16). А сходимость (15) следует из (39) и (42). □

Для доказательства теоремы 2 достаточно повторить вывод предыдущего утверждения, поменяв местами в формуле (42) символы $o_p(1)$ и $O_p(1)$.

Д о к а з а т е л ь с т в о следствия 1. Ввиду (41) в условиях следствия

$$\begin{aligned}\sqrt{n}\,\delta_n^*\big(I(\theta^*) - I(\theta)\big) &= \sqrt{n}\,|\delta_n^*|^{1+\alpha} o_p(1) = \\ &= \big(n^{\beta(\alpha)}|\delta_n^*|\big)^{1+\alpha} o_p(1) = \big(O_p(1)\big)^{1+\alpha} o_p(1) \underset{p}{\to} 0,\end{aligned} \tag{43}$$

$$\sqrt{n}\,\delta_n^*\overline{L_n''}(\theta) = n^{\beta(\alpha)}\delta_n^* \cdot n^{\alpha\beta(\alpha)}\overline{L_n''}(\theta) = O_p(1)\cdot o_p(1) \underset{p}{\to} 0. \tag{44}$$

Подставляя эти соотношения в (16) мы получим, что

$$\sqrt{n}\,(\theta_{n,I}^{**} - \theta)I(\theta_n^*) - L_n'(\theta)/\sqrt{n} \underset{p}{\to} 0. \tag{45}$$

Но оценка θ_n^* — состоятельна, а функция $I(t)$ непрерывна при $t = \theta$. Значит, $I(\theta_n^*) \underset{p}{\to} I(\theta)$ и из (45), с учетом (37), вытекает требуемая сходимость (4). □

Для доказательства следствия 2 нужно повторить вывод предыдущего утверждения, поменяв в формулах (43) и (44) местами символы $O_p(1)$ и $o_p(1)$.

Д о к а з а т е л ь с т в о теоремы 3. Из (39) и (41) имеем:

$$\rho_n^*(\delta_n^*)+\sqrt{n}(\delta_n^*)^2l/2=\sqrt{n}\,\delta_n^*\int\limits_0^1\bigl[\rho_n(v\delta_n^*)-v\delta_n^*l\bigr]dv+\sqrt{n}\,\delta_n^*\bigl[\delta_n^*l-\rho_n(\delta_n^*)\bigr].$$

Но отсюда и из (40) находим:

$$\bigl|\rho_n^*(\delta_n^*)+\sqrt{n}(\delta_n^*)^2l/2\bigr|\le 2\sqrt{n}\lambda_n|\delta_n^*|=\sqrt{n}|\delta_n^*|^2o_p(1). \tag{46}$$

Заметим теперь, что условие $(R_{3,\theta})$ теоремы 3 жестче, чем условие (19) при $\alpha=1$. Следовательно, мы можем воспользоваться утверждением леммы 1. Подставляя (46) в (39), получаем:

$$\Delta_{n,L}=\sqrt{n}(\delta_n^*)^2\bigl(-l/2+o_p(1)\bigr)\bigl(1+o_p(1)\bigr)+o_p(1).$$

Но из полученного соотношения вытекает требуемое утверждение (26) теоремы 3. $\square$

3.2. Д о к а з а т е л ь с т в о леммы 1. Прежде всего заметим, что справедливо неравенство

$$\lambda_n\le|\delta_n^*|^\alpha\overline{h}_n(\delta_n^*)\quad\text{при}\quad\overline{h}_n(\delta):=n^{-1}\sum_{i=1}^n h_{\alpha,\theta}(\delta,X_i). \tag{47}$$

Действительно, из определения (11) для любого $v\in[0,1]$, имеем:

$$\bigl|l''(\theta+v\delta_n^*,x)-l''(\theta,x)\le|v\delta_n^*|^\alpha h_{\alpha,\theta}\bigl(|v\delta_n^*|,x\bigr)\le|\delta_n^*|^\alpha h_{\alpha,\theta}\bigl(|\delta_n^*|,x\bigr).$$

Подставляя это неравенство в определения (2) и (34) находим, что

$$\forall v\in[0,1]\qquad\bigl|\rho_n(v\delta_n^*)\bigr|\le|\delta_n^*|^\alpha\overline{h}_n(\delta_n^*).$$

Но отсюда и из (38) вытекает (47).

Далее, для любых неслучайных $\delta>0$ и $x>0$

$$\mathbf{P}_\theta\bigl(\overline{h}_n(\delta_n^*)>x,\,|\delta_n^*|\le\delta\bigr)\le\mathbf{P}_\theta\bigl(\overline{h}_n(\delta)>x\bigr)\le\mathbf{E}_\theta\overline{h}_n(\delta)/x=\mathbf{E}_\theta h_{\alpha,\theta}(\delta,X_1)/x,$$

а потому

$$\mathbf{P}_\theta\bigl(\overline{h}_n(\delta_n^*)>x\bigr)\le\mathbf{P}_\theta\bigl(|\delta_n^*|>\delta\bigr)+\mathbf{E}_\theta h_{\alpha,\theta}(\delta,X_1)/x \tag{48}$$

и, следовательно, с учетом состоятельности оценки θ_n^*,

$$\limsup_{n\to\infty}\mathbf{P}_\theta\bigl(\overline{h}_n(\delta_n^*)>x\bigr)\le\mathbf{E}_\theta h_{\alpha,\theta}(\delta,X_1)/x\qquad\forall x>0\;\forall\delta>0. \tag{49}$$

Если выполнено условие (13), то из (49) при любом $\varepsilon>0$ находим:

$$\limsup_{n\to\infty}\mathbf{P}_\theta\bigl(\overline{h}_n(\delta_n^*)>\varepsilon\bigr)\le\mathbf{E}_\theta h_{\alpha,\theta}(\delta,X_1)/\varepsilon\to0\quad\text{при}\quad\delta\to0.$$

Значит, $\overline{h}_n(\delta_n^*)\underset{p}{\to}0$. Но из этого факта и неравенства (47) вытекает соотношение (38).

Если же верно условие (19), то из (49) имеем:

$$\limsup_{n\to\infty}\mathbf{P}_\theta\bigl(\overline{h}_n(\delta_n^*)>x\bigr)\le\mathbf{E}_\theta h_{\alpha,\theta}(\delta,X_1)/x\to0\quad\text{при}\quad x\to\infty.$$

Значит, $\overline{h}_n(\delta_n^*)=O_p(1)$ и из (47) следует утверждение $\lambda_n=|\delta_n^*|^\alpha O_p(1)$.

Докажем, наконец, (39). Поскольку оценка θ_n^* — состоятельна, то из (38) при $\alpha\ge0$ и из соотношения $\lambda_n=|\delta_n^*|^\alpha O_p(1)$ при $\alpha>0$ вытекает, что

$$\delta_n^*\underset{p}{\to}0\qquad\text{и}\qquad\bigl|\rho_n(\delta_n^*)\bigr|\le\lambda_n\underset{p}{\to}0.$$

Но из этих фактов и (37) имеем:

$$\lambda_n^* := \overline{L_n''}(\theta)/I(\theta) + \rho_n(\delta_n^*)/I(\theta) \underset{p}{\to} 0, \qquad \lambda_n^* L_n'(\theta)/\sqrt{n} \underset{p}{\to} 0. \tag{50}$$

Подставляя (50) в (36), находим:

$$\Delta_{n,L} = \frac{\lambda_n^* L_n'(\theta)/\sqrt{n} + \rho_n^*(\delta_n^*)}{1-\lambda_n^*} = \frac{o_p(1) + \rho_n^*(\delta_n^*)}{1-o_p(1)}.$$

Из последнего соотношения нетрудно извлечь (39). □

3.3. При выводе леммы 2 нам понадобится функция

$$\tau(\delta,x) := \big(l''(\theta+\delta,x) - l''(\theta,x)\big)/\delta - l'''(\theta,x) \quad \text{при} \quad \delta \ne 0, \quad \tau(0,x) := 0 \tag{51}$$

и обозначения

$$\overline{H}_n(\delta) := n^{-1}\sum_{i=1}^n H(\delta, X_i), \qquad \overline{L_n'''}(\theta) := L_n'''(\theta)/n - l. \tag{52}$$

Д о к а з а т е л ь с т в о леммы 2. Заметим прежде всего, что

$$\lim_{\delta\to 0} \mathbf{E}_\theta H(\delta, X_1) = 0 \quad \text{при} \quad H(\delta,x) := \sup\left\{|\tau(t,x)| : |t| \le \delta, \theta + t \in \Theta\right\}. \tag{53}$$

Действительно, при любых фиксированных x функция $\tau(\delta,x)$ непрерывна при $\delta = 0$. Следовательно, $H(\delta,x) \to 0$ при $\delta \to 0$. Кроме того,

$$\tau(t,x) = \int_0^1 l'''(\theta+tv,x)dv - l'''(\theta,x) \quad \text{и} \quad \big|\tau(t,x)\big| \le 2H_3(\theta,x)$$

при $|t| \le \delta_\theta$ в силу условия (10). Значит,

$$\sup_{|\delta|\le\delta_\theta} H(\delta,x) \le 2H_3(\theta,x) \quad \text{и} \quad \mathbf{E}_\theta H_3(\theta,X_1) < \infty.$$

Таким образом, соотношение (53) вытекает из теоремы Лебега о мажорируемой сходимости под знаком интеграла.

Сравнивая определения (34) и (51), имеем

$$\rho_n(v\delta_n^*) - v\delta_n^* l = n^{-1} v\delta_n^* \sum_{i=1}^n \tau(v\delta_n^*, X_i) + v\delta_n^*\overline{L_n'''}(\theta).$$

Из этого равенства и обозначений (52), (53) нетрудно извлечь, что

$$\overline{\lambda}_n \le |\delta_n^*|\big(\overline{H}_n(\delta_n^*) + |\overline{L_n'''}(\theta)|\big). \tag{54}$$

Далее, при всех неслучайных $\varepsilon > 0$ и $\delta > 0$

$$\begin{aligned} \mathbf{P}_\theta\big(\overline{H}_n(\delta_n^*) > \varepsilon, \, |\delta_n^*| \le \delta\big) \le \mathbf{P}_\theta\big(\overline{H}_n(\delta) > \varepsilon\big) \le \mathbf{E}_\theta\overline{H}_n(\delta)/\varepsilon = \\ = \mathbf{E}_\theta H(\delta, X_1)/\varepsilon \to 0 \quad \text{при} \quad \delta \to 0. \end{aligned} \tag{55}$$

Подчеркнем, что сходимость к нулю в (55) следует из соотношения (53). Таким образом,

$$\mathbf{P}_\theta\big(\overline{H}_n(\delta_n^*) > \varepsilon\big) \le \mathbf{P}_\theta\big(|\delta_n^*| > \delta\big) + \mathbf{E}_\theta H(\delta, X_1)/\varepsilon. \tag{56}$$

Учитывая состоятельность θ_n^* и переходя в (56) к пределу сначала при $n \to \infty$, а затем при $\delta \to 0$, при всех $\varepsilon > 0$ получаем

$$\limsup_{n\to\infty} \mathbf{P}_\theta\big(\overline{H}_n(\delta_n^*) > \varepsilon\big) \le \mathbf{E}_\theta H(\delta, X_1)/\varepsilon \to 0 \quad \text{при} \quad \delta \to 0.$$

Тем самым мы доказали, что $\overline{H}_n(\delta_n^*) \underset{p}{\to} 0$. Наконец, из определения (52) и из закона больших чисел вытекает, что $\overline{L'''_n}(\theta) \underset{p}{\to} 0$. Утверждение (40) леммы немедленно следует из оценки (54) и двух последних сходимостей. □

3.4. Нам потребуется еще одна вспомогательная

Лемма 3. *Если выполнены условия замечания* 1, *то* $1 < \gamma(\alpha) \le 2$ *и*

$$\mathbf{E}_\theta \big| n^{\alpha\beta(\alpha)} \overline{L''_n}(\theta) \big|^{\gamma(\alpha)} \le 2\mathbf{E}_\theta |\nu_1|^{\gamma(\alpha)} < \infty \quad \textit{при} \quad \nu_i = l''(\theta, X_i) + I(\theta). \tag{57}$$

Д о к а з а т е л ь с т в о. В силу выбора чисел $\beta(\alpha)$ и $\gamma(\alpha)$ в (12) и (23), соотношение (57) можно переписать следующим образом

$$n^{-1}\mathbf{E}_\theta \Big| \sum_{i=1}^n \nu_i \Big|^{\gamma(\alpha)} \le n^{-1} 2 \sum_{i=1}^n \mathbf{E}_\theta |\nu_i|^{\gamma(\alpha)} = 2\mathbf{E}_\theta |\nu_1|^{\gamma(\alpha)}. \tag{58}$$

Но соотношение (58) очевидно верно, так как оно совпадает с известным неравенством Бара–Эссена (см, например, [7], стр. 79), которое справедливо, так как величины $\nu_1, \ldots, \nu_n$ — независимы, одинаково распределены и имеют нулевые средние. □

Для вывода утверждения из замечания 1 нам осталось заметить, что при $\zeta_n = n^{\alpha\beta(\alpha)} \overline{L''_n}(\theta)$

$$\sup_n \mathbf{P}_\theta \big(|\zeta_n| > x \big) \le \sup_n \mathbf{E}_\theta |\zeta_n|^{\gamma(\alpha)} / x^{\gamma(\alpha)} \le 2\mathbf{E}_\theta |\nu_1|^{\gamma(\alpha)} / x^{\gamma(\alpha)} \to 0 \quad \text{при} \quad x \to \infty$$

в силу (57). Значит, $\zeta_n = O_p(1)$.

Список литературы

[1] R.A. Fisher, *Theory of statistical estimation*, Proc. Camb. Phil.Soc., **22** (1925), 700–725. JFM 51.0385.01

[2] Закс Ш. *Теория статистических выводов*, М.:Мир, 1975. MR0420924

[3] Леман Э. *Теория точечного оценивания*, М.: Наука, 1991. MR1143059

[4] Боровков А.А. *Математическая статистика*, Новосибирск: Изд-во ИМ СО РАН, 1997. MR1845991

[5] S. Verrill, *Rate of convergence of k-step Newton estimators to efficient likelihood estimators,* Statistical and Probability Letters, **77** (2007), 1371–1376. MR2392808

[6] P. Janssen, J. Jureckova, N. Veraverbeke, *Rate of convergence of one- and two-step M-estimators with applications to maksimum likelihood and Pitman estimators*, Ann. Stat., 13, **3** (1985), 1222–1229. MR0803768

[7] Петров В.В. *Суммы независимых случайных величин*, М.: Наука, 1972. MR0322927

Юлиана Юрьевна Линке
Институт математики им. С. Л. Соболева СО РАН,
пр. академика Коптюга, 4,
630090, Новосибирск, Россия;
Новосибирский государственный университет,
ул. Пирогова, 2,
630090, Новосибирск, Россия
E-mail address: linke@math.nsc.ru

Александр Иванович Саханенко
Институт математики им. С. Л. Соболева СО РАН,
пр. академика Коптюга, 4,
630090, Новосибирск, Россия;
Новосибирский государственный университет,
ул. Пирогова, 2,
630090, Новосибирск, Россия
E-mail address: aisakh@mail.ru