Доклады Академии наук СССР
1971. Том 198, № 3

УДК 519.25 *МАТЕМАТИКА*

И. А. ИБРАГИМОВ, Р. З. ХАСЬМИНСКИЙ

О ПРЕДЕЛЬНОМ ПОВЕДЕНИИ БАЙЕСОВСКИХ ОЦЕНОК И ОЦЕНОК МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

(Представлено академиком Ю. В. Линником 20 XI 1970)

1. Пусть $X_1, X_2, \ldots, X_n$ — повторная выборка из генеральной совокупности с распределением P_θ, зависящим от неизвестного параметра $\theta \in \Theta$. Допустим, что при всех θ P_θ суть распределения в k-мерном эвклидовом пространстве $\mathbf{R}^k$, а параметрическое множество Θ есть открытый интервал (может быть, и бесконечный) действительной прямой. Предположим, что все меры P_θ абсолютно непрерывны относительно некоторой меры ν в $\mathbf{R}^k$ и положим $f(x, \theta) = dP_\theta / d\nu$. Будем предполагать в дальнейшем, что $f(x, \theta)$ измерима по отношению к $d\nu \times d\theta$ и что для любых двух точек $\theta, \theta' \in \Theta$, $\theta \neq \theta'$

$$\int |f(x, \theta) - f(x, \theta')| \, \nu(dx) > 0.$$

Задавшись неотрицательной измеримой функцией $\pi(y)$, $y \in \mathbf{R}^1$, рассмотрим оценку $\tilde{\theta}_n$ параметра θ, определенную равенством

$$\tilde{\theta}_n = \int_\Theta y p_n(y) \, dy, \tag{1}$$

где

$$p_n(y) = \pi(y) \prod_{j=1}^n f(X_j, y) \Big/ \left(\int_\Theta \prod_{j=1}^n f(X_j, z) \pi(z) \, dz \right). \tag{2}$$

В ([1]) был предложен новый метод исследования асимптотического поведения оценки $\tilde{\theta}_n$, основанный на асимптотическом анализе нормированного отношения правдоподобия, которое мы трактуем как случайную функцию θ (сходный метод для изучения оценок максимального правдоподобия использован в ([2])). В настоящей статье уточняются некоторые результаты ([1]) для случая, когда семейство P_θ имеет конечное информационное количество Фишера. Кроме оценок (1), мы будем рассматривать также оценку максимального правдоподобия $\hat{\theta}_n$, определяя ее как один из корней уравнения (например, наибольший)

$$\prod_1^n f(X_j, \hat{\theta}_n) = \max_{\theta \in \Theta} \prod_1^n f(X_j, \theta),$$

и байесовские оценки относительно функции риска $|\theta|^\alpha$. Напомним, что оценка t_n называется байесовской относительно функции риска $|\theta|^\alpha$ и «априорной плотности» $\pi(\theta)$, если

$$\int_\Theta |y - t_n|^\alpha p_n(y) \, dy = \min_{\theta \in \Theta} \int |y - \theta|^\alpha p_n(y) \, dy. \tag{3}$$

В частности $\tilde{\theta}_n$ — апостериорное среднее, есть байесовская оценка относительно $|\theta|^2$, апостериорная медиана есть байесовская оценка относительно $|\theta|$.

2. Перечислим ограничения, которые будут налагаться на семейство P_θ и функции $\pi(y)$. При этом условимся обозначать через θ_0 истинное значение параметра и писать $P\{\cdot\}$, $E\{\cdot\}$ вместо $P_{\theta_0}\{\cdot\}$, $E_{\theta_0}\{\cdot\}$.

Условия группы I.

I_1. При любом фиксированном x функция $f(x,\theta)$ определена и абсолютно непрерывна на замыкании Θ^c множества Θ.

I_2. При всех $\theta\in\Theta$ информационное количество Фишера

$$I(\theta)=\int\frac{|f'_\theta(x,\theta)|^2}{f(x,\theta)}\nu(dx)<\infty.$$

(При этом считается, что подынтегральное выражение равно нулю там, где $f(x,\theta)=0$.)

I_3. При всех $\delta>0$, $\theta\in\Theta$

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int\limits_{\theta-\varepsilon}^{\theta+\varepsilon}dv\int\limits_{\Delta}\frac{|f'_v(x,v)|^2}{f(x,v)}\nu(dx)=0,$$

$$\Delta=\left\{x:\left|\ln\frac{f(x,\theta+\varepsilon)}{f(x,\theta)}\right|>\delta\right\}.$$

При этом, конечно, считается

$$\left\{x:\left|\ln\frac{f(x,\theta+\varepsilon)}{f(x,\theta)}\right|=\infty\right\}\subset\left\{x:\left|\ln\frac{f(x,\theta+\varepsilon)}{f(x,\theta)}\right|>\delta\right\}.$$

I_4. Функция $I(\theta)$ непрерывна на Θ^c.

I_5. Существует число $p>0$ такое, что

$$\sup_{\theta\in\Theta}(1+|\theta|)^{-p}I(\theta)<\infty.$$

II. Для некоторого $\varepsilon>0$

$$\sup_{\theta\in\Theta}|\theta-\theta_0|^\varepsilon\int\sqrt{f(x,\theta)f(x,\theta_0)}\,\nu(dx)<\infty.$$

III. Функция $\pi(\theta)$ непрерывна в окрестности точки θ_0, $\pi(\theta_0)\neq 0$, и для некоторого $p\geqslant 0$

$$\sup_{\theta\in\Theta}(1+|\theta|^{-p}\pi(\theta)<\infty.$$

Наиболее сложное из этих условий — условие I_3. Оно выполняется, например, в случаях:

1) если для какого-нибудь $\delta>0$ интегралы

$$I_{2+\delta}(\theta)=\int\frac{|f'_\theta(x,\theta)|^{2+\delta}}{(f(x,\theta))^{1+\delta}}\nu(dx)$$

ограничены на компактах;

2) если $f(x,\theta)=f(x-\theta)$, ν — мера Лебега и $\int\frac{|f'(x)|^2}{f(x)}dx<\infty$ (случай оценки параметра сдвига).

3. Пусть $\Theta=(\alpha,\beta)$. Определим случайные процессы $Z_n(\theta)$ равенствами

$$Z_n(\theta)=\begin{cases}\prod\limits_1^n\dfrac{f(X_j,\theta_0+\theta/\sqrt{n}}{f(X_j,\theta_0)} & \text{для } \theta\in[(\alpha-\theta_0)\sqrt{n},(\beta-\theta_0)\sqrt{n}],\\ 0 & \text{для } \theta\,\overline{\in}\,[(\alpha-\theta_0)\sqrt{n}-1,(\beta-\theta_0)\sqrt{n}+1].\end{cases}$$

Z_n линеен в оставшейся части $\mathbf{R}^1$. Обозначим $C_0(-\infty,\infty)=C_0$ пространство функций $\varphi(t)$, непрерывных на $\mathbf{R}^1$ и таких, что $\lim\limits_{t\to\pm\infty}\varphi(t)=0$, снабженное равномерной метрикой. Положим, наконец, $Z(\theta)=$ $=\exp\{\sqrt{I_0}\theta\xi-I_0\theta^2/2\}$, где $I_0=I(\theta_0)$, а ξ нормальная случайная величина с параметрами 0, 1.

Теорема 1. *Если выполняются условия* $\mathrm{I}_1-\mathrm{I}_4$, *то при* $n\to\infty$ *конечномерные распределения процессов* $Z_n(\theta)$ *сходятся к конечномерным распределениям процесса* $Z(\theta)$. *Если выполняются условия* I, II, *распределения в* C_0, *порожденные процессами* Z_n, *слабо сходятся к распределению, порожденному в* C_0 *процессом* Z.

Все приводимые дальше результаты основаны на теореме 1. Чтобы продемонстрировать ее возможности, покажем, например, как из этой теоремы выводится следующий результат.

Если выполняются условия I, II, *разность* $\sqrt{n}(\hat\theta_n-\theta_0)$ *асимптотически нормальна с параметрами* $0, I_0^{-1}$.

Для доказательства, задавшись числом x, рассмотрим функционалы на C_0:

$$f_x(\varphi)=\max_{y\leqslant x}|\varphi(y)|,\quad F_x(\varphi)=\sup_{y>x}|\varphi(y)|,\quad \varphi\in C_0.$$

Эти функционалы непрерывны в C_0 и потому

$$P\{\sqrt{n}(\hat\theta_n-\theta_0)\geqslant x\}=P\{F_x(Z_n)-f_x(Z_n)\geqslant 0\}\underset{n\to\infty}{\longrightarrow}$$
$$\to P\{F_x(Z)-f_x(Z)\geqslant 0\}=P\{\xi/\sqrt{I_0}\geqslant x\}.$$

4. В качестве следствия теоремы 1 можно получить ряд результатов относительно $\tilde\theta_n$ и $p_n(\theta)$. Эти результаты выводятся путем формального перехода к пределу в (1), (2) (см. ([1])) и дальнейшего оправдания такого перехода.

Теорема 2. *Если выполняются условия* $\mathrm{I}_1-\mathrm{I}_4$, II, III, *то распределение разности* $\sqrt{n}(\tilde\theta_n-\theta_0)$ *асимптотически нормально с параметрами* $0, I_0^{-1}$. *Если выполняются условия* I—III, *то для любого числа* $p>0$

$$\lim_{n\to\infty}n^{1/2p}E|\tilde\theta_n-\theta_0|^p=I_0^{-1/2p}E|\xi|^p=\left(\frac{2}{I_0}\right)^{1/2p}\Gamma\left(\frac{p+1}{2}\right)\Big/\sqrt{\pi}.$$

Положим

$$\Delta_n(\theta)=\frac{1}{\sqrt{n}}p_n\left(\theta_0+\frac{\theta}{\sqrt{n}}\right)-\sqrt{\frac{I_0}{2\pi}}\exp\left\{-\frac{I_0}{2}(\theta-\sqrt{n}(\tilde\theta_n-\theta_0))^2\right\}.$$

Теорема 3. *Если выполнены условия* I—III, *то справедливы утверждения:* 1) *случайные величины* $\max_\theta|\Delta_n(\theta)|$, $\int|\theta|^p|\Delta_n(\theta)|^q d\theta$, $p,\ q>0$ *сходятся к нулю по вероятности;* 2) *для любого* $N>0$ *найдется постоянная* C_N *такая, что*

$$P\{\max_{|\theta|>A}|\Delta_n(\theta)|>A^{-N}\}\leqslant C_NA^{-N}.$$

5. Теоремы 2, 3, в свою очередь, позволяют исследовать асимптотическое поведение оценок t_n, $\hat\theta_n$.

Теорема 4. *Пусть выполнены условия* I—III.

Тогда распределения разностей $\sqrt{n}(\hat\theta_n-\theta_0)$, $\sqrt{n}(t_n-\theta_0)$ *асимптотически нормальны с параметрами* $0, I_0^{-1}$ *и при всех* $p\geqslant 0$

$$\lim_{n\to\infty}n^{1/2p}E|\hat\theta_n-\theta_0|^p=\lim_{n\to\infty}n^{1/2p}|t_n-\theta_0|^p=I_0^{-1/2p}E|\xi|^p.$$

Существует ряд работ, посвященных асимптотическому анализу распределений байесовских оценок (см., например, ([3])), однако все они используют традиционные ограничения, связанные с непрерывностью $\ln f(x,\theta)$ и существованием $f_{\theta\theta}''$. Нам известна лишь одна работа ([4]), где изучаются моменты оценок максимального правдоподобия.

6. Лишь для простоты и краткости мы ограничили изложение случаем одномерного параметра θ. Наши результаты остаются справедливыми и для

$\Theta \subset R^m$. В этом случае роль информационного количества играет информационная матрица и условия I должны выполняться для всех ее элементов. Кроме того надлежит потребовать, чтобы эта матрица была невырожденной при всех θ. Предельный процесс $Z(\theta)$ теперь имеет вид

$$Z(\theta) = \exp\{(I_0^{1/2}\xi, \theta) - {}^1/_2 (I_0\theta, \theta)\}.$$

Здесь

$$I_0 = \left\| E \frac{\partial \ln f(X, \theta)}{\partial \theta_i} \frac{\partial \ln f(X, \theta)}{\partial \theta_j} \right\| -$$

информационная матрица Фишера, ξ — m-мерный вектор, компоненты которого суть независимые нормально распределенные случайные величины с параметрами 0, 1.

Отметим также, что условия п.п. 1, 2 отнюдь не являются максимальными. Так, например, в качестве множества значений случайных величин X_j можно взять произвольное измеримое пространство с σ-конечной мерой ν, относительно θ достаточно предположить, что оно открыто, и т. д.

Авторы признательны участникам семинара акад. Ю. В. Линника по теории последовательного оценивания за полезное обсуждение.

Ленинградский государственный университет
им. А. А. Жданова

Институт проблем передачи информации
Академии наук СССР
Москва

Поступило
20 XI 1970

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] И. А. Ибрагимов, Р. З. Хасьминский, ДАН, 194, № 2 (1970). [2] L. Le Cam, Ann. Math. Stat., 41, № 3, 802 (1970). [3] M. T. Chao, Ann. Math. Stat., 41, № 3, 601 (1970). [4] Yu. V. Linnik, N. M. Mitrofanova, Contrib. to Stat., Presented to Prof. Mahalonobis, 1964, p. 229.