Доклады Академии наук СССР
1970. Том 194, № 2

УДК 519.25 *МАТЕМАТИКА*

И. А. ИБРАГИМОВ, Р. З. ХАСЬМИНСКИЙ

ОБ АСИМПТОТИЧЕСКОМ ПОВЕДЕНИИ ОБОБЩЕННЫХ БАЙЕСОВСКИХ ОЦЕНОК

(Представлено академиком А. Н. Колмогоровым 11 III 1970)

1. Пусть $x_1, \ldots, x_n$ — выборка с независимыми элементами из генеральной совокупности с распределением P_θ, зависящим от неизвестного параметра $\theta \in \Theta$. Допустим, что P_θ при всех θ суть распределения в R_n, а параметрическое множество $\Theta \subset R_m$. Предположим также, что все меры R_θ абсолютно непрерывны относительно некоторой меры ν и положим $f(x, \theta) = dP_\theta / d\nu$.

Нашей целью будет исследование асимптотического поведения конечномерных распределений случайного поля

$$p_n(\theta) = \prod_1^n f(x_i, \theta)\, \pi(\theta) \Big/ \int \prod_1^n f(x_i, \theta)\, \pi(\theta)\, \nu(d\theta) \tag{1}$$

и оценок $\bar{\theta}_n$ параметра θ, имеющих вид

$$\tilde{\theta}_n = \int \theta p_n(\theta)\, \nu(d\theta) \tag{2}$$

для измеримых в R_m функций $\pi(\theta)$. Если $\pi(\theta)$ — плотность распределения вероятностей в R_m относительно меры ν, то поле $p_n(\theta)$ и оценку $\tilde{\theta}_n$ естественно трактовать как байесовскую апостериорную плотность и соответственно апостериорное среднее случайной величины, имеющей априорное распределение с плотностью $\pi(\theta)$. Мы сохраним название апостериорной плотности и байесовской оценки за правыми частями (1) и (2), хотя в наших рассмотрениях θ не есть случайная величина. Будем обозначать θ_0 истинное (неизвестное наблюдателю) значение параметра θ.

Назовем функцию $\varphi(n)$ правильным нормирующим множителем (п.н.м.) для (1), (2), если конечномерные распределения случайного поля $\eta_n(\theta) = \varphi^{-1}(n) p_n(\theta_0 + \theta / \varphi(n))$ и распределение случайной величины $\zeta_n = \varphi(n)(\tilde{\theta}_n - \theta_0)$ при $n \to \infty$ сходятся к собственным предельным.

В настоящей заметке мы занимаемся отысканием п.н.м. и соответствующих предельных распределений при тех или иных ограничениях, наложенных на семейство P_θ. Чтобы сделать изложение менее громоздким, мы ограничимся случаем, когда случайные величины x_i и параметр θ одномерны, а $\nu(dy) = dy$ — мера Лебега *.

2. Пусть $f(x, \theta)$ — измеримая функция (x, θ), а «априорная плотность» $\pi(\theta)$ непрерывна в точке θ_0, ограничена на компактах и

$$\iint |\theta| f(x, \theta) f(x, \theta_0) |\pi(\theta)|\, dx\, d\theta < \infty.$$

Относительно семейства P_θ ниже будут считаться выполненными какие-либо из следующих условий

I. 1) $\lim\limits_{|\theta| \to \infty} \int f(x, \theta) f(x, \theta_0)\, dx = 0$; 2) $\lim\limits_{\varepsilon \to 0} \int f^{1-\varepsilon}(x, \theta_0)\, dx = 1$.

---

* Постановкой некоторых из рассмотренных здесь задач мы обязаны Ю. В. Линнику.

II. Каковы бы ни были вещественные числа $\theta_1, \ldots, \theta_s$, найдутся такие интервалы $[0, T_1], \ldots, [0, T_s]$, что для всех $t_j \in [0, T_j]$ при $\xi \downarrow 0$

$$\int \prod_{j=1}^{s} \left(\frac{f(x, \theta_0 + \xi\theta_j)}{f(x, \theta_0)} \right)^{t_j} f(x, \theta_0)\, dx = 1 + \xi^\alpha a(t_1, \ldots, t_s;\ \theta_1, \ldots, \theta_s) + o(\xi^\alpha), \tag{3}$$

где число α не зависит от $\theta_1, \ldots, \theta_s, t_1, \ldots, t_s$, а функция a отлична от тождественного нуля.

III. Для всех θ_1, θ_2 с $|\theta|_1, |\theta_2| \leqslant H < \infty$ при $\xi \to 0$

$$\left| \int \ln \frac{f(x, \theta_0 + \xi\theta_1)}{f(x, \theta_0 + \xi\theta_2)} f(x, \theta_0 + \xi\theta_2)\, dx \right| \leqslant |\xi|^\alpha C_1(|\theta_2 - \theta_1|),$$

$$\left| \int \ln^2 \frac{f(x, \theta_0 + \xi\theta_1)}{f(x, \theta_0 + \xi\theta_2)} f(x, \theta_0 + \xi\theta_2)\, dx \right| \leqslant |\xi|^\alpha C_2(|\theta_2 - \theta_1|);$$

здесь (зависящие от H) функции $C_i(h)$, $i = 1, 2$, стремятся к нулю при $h \to 0$ и в обоих интегралах функцию $f(x, \theta_0 + \xi\theta_i)$ надо заменить единицей там, где она равна нулю.

Обозначим через $Y(\theta)$ (сепарабельный) случайный процесс, конечномерные распределения которого определяются равенствами

$$M_r \exp\{\Sigma t_j Y(\theta_j)\} = \exp\{a(t_1, \ldots, t_s, \theta_1, \ldots, \theta_s)\},$$

причем здесь и ниже $M(\cdot) = M_{\theta_0}(\cdot)$, $P(\cdot) = P_{\theta_0}(\cdot)$.

Теорема 1. *Пусть выполнены предположения* I—III. *Тогда п.н.м.* $\varphi(n) = n^{1/\alpha}$, *конечномерные распределения процесса* $\eta_n(\theta) = \varphi^{-1}(n) p_n(\theta_0 + \theta / \varphi(n))$ $(\theta \in (-\infty, +\infty))$ *сходятся к соответствующим распределениям процесса* $\eta_0(\theta) = (\int e^{Y(\theta)} d\theta)^{-1} e^{Y(\theta)}$, *а нормированные разности* $\zeta_n = \varphi(n)(\bar\theta_n - \theta_0)$ *сходятся по распределению к случайной величине* $\int \theta\eta_0(\theta)\, d\theta$.

3. Если, например, левую часть (3) можно продифференцировать дважды под знаком интеграла, мы найдем, что $\alpha = 2$,

$$\eta_0(\theta) = \frac{1}{\sqrt{2\pi I}} \exp\left\{ -\left(\theta - \frac{\xi}{\sqrt{I}}\right)^2 / 2I \right\}; \quad Y(\theta) = \sqrt{I}\,\theta\xi - \frac{\theta^2}{2} I, \tag{4}$$

где ξ — стандартная нормальная случайная величина, а I — информационное количество Фишера. Поэтому нормированная апостериорная плотность асимптотически гауссова, а разность $\sqrt{n}(\bar\theta_n - \theta_0)$ также асимптотически нормальна с параметрами 0, $I^{-1/2}$ (см. также п. 5). В сходных предположениях можно также доказать, что $\bar\theta_n$ близка к оценке наибольшего правдоподобия $\hat\theta_n$ в том смысле, что $\sqrt{n}(\bar\theta_n - \hat\theta_n) \to 0$ по вероятности или с вероятностью 1, в зависимости от ограничений на P_θ. Мы не будем подробнее останавливаться на этих результатах, потому что близкие теоремы имеются в литературе (см., например, [1], где можно найти также восходящую к С. Н. Бернштейну и Р. Мизесу историю вопроса).

4. Обратимся к разрывным плотностям $f(x, \theta)$. Будем считать, что функция $f(x, \theta)$ ограничена, непрерывно дифференцируема по θ на каждом из интервалов $(x_k(\theta), x_{k+1}(\theta))$, а в точках $x_k(\theta)$, $k = 1, \ldots, r$, имеет разрывы первого рода. Предположим, что функции $x_k(\theta)$ непрерывно дифференцируемы и обозначим

$$p(k) = \lim_{x \downarrow x_k(\theta_0)} f(x, \theta_0), \quad q(k) = \lim_{x \uparrow x_k(\theta_0)} f(x, \theta_0).$$

Рассмотрим две независимые пуассоновские случайные меры, сосредоточенные на множестве $R = \{1, 2, \ldots, r\}$. Нам удобно будет обозначать эти меры общим символом $\nu(\theta, A)$, $A \in R$. Это не вызовет недоразумений, так как для одной из этих мер параметр $\theta \in (-\infty, 0]$, для второй $\theta \in [0, \infty)$, $\nu(0, A) = 0$. Распределения мер $\nu(\theta, A)$ определяются равенствами: $M\nu(\theta, k) = p(k) x_k'(\theta_0)\theta$, если $\theta x_k'(\theta_0) > 0$, и $M\nu(\theta, k) = -q(k) x_k'(\theta_0)\theta$, если $\theta x_k'(\theta_0) < 0$. Обозначим R^* подмножество точек k из R, для которых $p(k) q(k) = 0$, и положим

$$\tau_+ = \inf_{\theta > 0}\{\theta\colon\ \nu(\theta, R^*) > 0\}, \quad \tau_- = \sup_{\theta < 0}\{\theta\colon \nu(\theta, R^*) > 0\}$$

(если R^* пусто, будем считать $\tau_+ = \infty$, $\tau_- = -\infty$).

Теорема 2. *Пусть, помимо предположений п. 4, выполнены предположения* 1 *п.* 2 *и* $|f_0'(x, \theta)| \leqslant q(x)$, *где* $\int q(x)\,dx < \infty$. *Тогда справедливо заключение теоремы* 1 *с п.н.м.* $\varphi(n) = n$ *и предельным процессом*

$$Y(\theta) = \theta \sum_{k=1}^{r} (p(k) - q(k))\, x_k'(\theta_0) + \int_R \operatorname{sign}(x_z'(\theta_0)\,\theta) \ln \frac{q(z)}{p(z)}\, \nu(\theta, dz)$$

$$\text{для } \theta \in (\tau_-, \tau_+),$$

$$Y(\theta) = -\infty \qquad \text{для } \theta \,\overline{\in}\, [\tau_-, \tau_+).$$

При этом предел по распределению для $n(\bar\theta - \theta_0)$ *можно записать в виде*

$$\left(\int_{\tau_-}^{\tau_+} \exp Y(\theta)\, d\theta\right)^{-1} \int_{\tau_-}^{\tau_+} \theta \exp Y(\theta)\, d\theta.$$

Условиям этой теоремы удовлетворяют, например, те рассмотренные в [2] случаи, когда: 1) $f(x, \theta) = \beta$ для $0 \leqslant x \leqslant \theta$, $f(x, \theta) = \gamma$ для $\theta < x \leqslant 1$ и $f(x, \theta) \equiv 0$ для $x \,\overline{\in}\, [0, 1]$; 2) $f(x, \theta)$ имеет одну точку разрыва.

5. В этом пункте мы предположим, что $f(x, \theta) = f(x - \theta)$, так что оценке подлежит параметр сдвига. Положим также $\pi(\theta) \equiv 1$. Оценка (2) есть теперь оценка Питмена для параметра сдвига (см. [3]); можно без потери общности положить $\theta_0 = 0$. Пусть, кроме того,

$$\int |x|\, f(x)\, dx < \infty. \tag{5}$$

Теорема 3. *Пусть выполнено условие* (5), *функция* $f(x)$ *абсолютно непрерывна, а информационное количество Фишера*

$$I = \int \frac{|f'(x)|^2}{f(x)}\, dx$$

конечно. Тогда конечномерные распределения процесса $\eta_n(\theta) = n^{-1/2} p_n(\theta_0 + \theta / \sqrt{n})$ *сходятся к соответствующим распределениям процесса* $\eta_0(\theta)$, *определяемого формулой* (4), *а разность* $\sqrt{n}\,(\bar\theta_n - \theta_0)$ *асимптотически нормальна с параметрами* $0, I^{-1/2}$.

Теорема 4. *Пусть выполняются условия теоремы* 3 *и* $M|x_i|^p < \infty$ *для какого-нибудь* $p > 1$. *Тогда*

$$\lim_{n\to\infty} n^{p/2} M\, |\tilde\theta_n - \theta_0|^p = I^{-p/2} \pi^{-1/2} 2^{(p+1)/2}\, \Gamma((p+1)/2). \tag{6}$$

Налагая дальнейшие ограничения на моменты случайных величин $l^{(k)}(x_i)$, где $l(x) = \ln f(x)$, можно получить дальнейшие члены асимптотического разложения левой части (6). Например, пусть выполнены условия теоремы 3 и, кроме того, конечны моменты

$$M\, |l^{(k)}(x_i)|^s,\ k = 1, \ldots, 5,\ s \leqslant 20; \quad M\,(\max_{|\theta| \leqslant \varepsilon} |l^{(5)}(x_i - \theta)|)^s,\ s \leqslant 20.$$

Тогда

$$M\,(\tilde\theta_n - \theta_0)^2 = \frac{1}{In} + \frac{C_1}{n^2} + o(n^{-2}).$$

Аналогично, если $\tilde\theta_n$ есть общая оценка (2), но с $\pi(\theta) \equiv 1$, $\nu(dy) \equiv dy$, и конечны моменты $M\,|l_\theta^{(k)}(x_1, \theta_0)|^s$, $k = 1, \ldots, 5$; $s \leqslant 20$, $M \max\limits_{|\theta| \leqslant \varepsilon} |l_\theta^{(5)}(x_0, \theta_0 - \theta)|^s$, $s \leqslant 20$, где $l(x, \theta) = \ln f(x, \theta)$, и если $\max M\tilde\theta_n^2 < \infty$, то

$$M(\tilde\theta_n - \theta_0) = \frac{c(\theta_0)}{n} + o\left(\frac{1}{n}\right); \quad c(\theta_0) = \frac{1}{I^2}[Ml_\theta'''(x_1, \theta_0) + Ml'(x_1, \theta_0)\, l''(x_1, \theta_0)],$$

$$M(\tilde\theta_n - \theta_0)^2 = \frac{1}{In} + \frac{c_1(\theta_0)}{n^2} + o\left(\frac{1}{n^2}\right), \quad I = M\,|l'(x_1, \theta_0)|^2.$$

6. Чтобы выяснить, что происходит при нарушении условий регулярности, рассмотрим следующий класс плотностей $f(x, \theta) = f(x - \theta)$. Функция $f(x)$ дважды непрерывно дифференцируема в каждом из интервалов (x_k, x_{k+1}), $k = 0, \ldots, r$, $x_0 = -\infty$, $x_{r+1} = \infty$, а при $x \to x_k$, $k = 1, \ldots, r$, ее производная $f'(x)$ удовлетворяет условиям:

$$f'(x) \sim a_k^- \alpha_k^- |x - x_k|^{\alpha_k^- - 1}, \quad x \to x_k - 0; \quad f'(x) \sim a_k^+ \alpha_k^+ |x - x_k|^{\alpha_k^+ - 1}, \quad x \to x_k + 0,$$

где $-1 \not< \alpha_k^\pm < 1$, $\alpha_k^\pm \neq 0$. Нам будет удобнее включить сюда также и случай $\alpha_k^\pm = 0$, считая, что равенство $\alpha_k^+ = 0$ $(\alpha_k^- = 0)$ означает, что $f(x_k + 0) = a_k^+$ $(f(x_k - 0) = a_k^-)$.

Введем обозначения

$$\gamma = \min_k(\alpha_k^+, \alpha_k^-), \quad A^\pm = \{k : \alpha_k^\pm = \gamma\}, \quad C = \{k : f(x_k) = 0\},$$

$$\gamma_1 = \min\{2 \min_{k \overline{\in} C} \alpha_k^\pm, \min_{k \in C} \alpha_k^\pm\}, \quad A_1^\pm = \{k : k \overline{\in} C,\ 2\alpha_k^\pm = \gamma_1\},$$

$$A_2^\pm = \{k : k \in C,\ \alpha_k^\pm = \gamma_1\}.$$

Справедливы следующие утверждения:

1) *Если* $\gamma \leqslant 0$, *то п.н.м.* $\varphi(n) = n^{1/(1+\gamma)}$; *при этом единственными параметрами, определяющими предельное распределение для* $\eta_n(\theta)$ *и* ζ_n, *являются числа* γ *и* $a_k^\pm$ *для* $k \in A^\pm$.

2) *Если* $\gamma > 0$, *то п.н.м.* $\varphi(n) = n^{1/(1+\gamma_1)}$, *а единственными определяющими предельные распределения для* $\eta_n(\theta)$, ζ_n *параметрами служат числа* γ_1 *и* $a_k^\pm$ *для* $k \in A_1^\pm \cup A_2^\pm$.

Эти утверждения позволяют упорядочить особенности функции $f(x)$ по их «информативности», если ограничиться классом особенностей степенного вида. Их можно распространить и на случай особенностей вида $|x - x_k|^\alpha l(x - x_k)$, где $l(x)$ — медленно меняющаяся в смысле Карамата функция (уже в (3) множитель ξ^α можно заменить на $\xi^\alpha l(\xi)$). Интересно отметить, что медленно меняющиеся функции в п.н.м. могут появиться и в «крайних» точках степенных особенностей функции $f(x)$. Например, для функции $f(x)$, имеющей лишь особенности вида $a_k|x - x_k|$ и $c_k + a_k|x - x_k|^{1/2}$, $c_k > 0$, п.н.м. $\varphi(n) = \sqrt{n \ln n}$.

Мы выражаем глубокую благодарность А. Н. Колмогорову и Ю. В. Линнику за внимание к работе и полезное обсуждение.

Ленинградский государственный университет
им. А. А. Жданова

Институт проблем передачи информации
Академии наук СССР
Москва

Поступило
1 III 1970

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] L. Le Cam, Univ. Calif. Public. Stat., 1, 277 (1953); Сборн. пер. Математика, 4, 2, 69 (1960). [2] H. Chernoff, H. Rubin, Proc. III Berkeley Symp., 1, Univ. Calif. Press, 1956, p. 19. [3] А. М. Каган, Тр. Матем. инст. им. В. А. Стеклова АН СССР, **104**, 19 (1968).