Секция «Теория вероятностей и математическая статистика»

Асимптотика апостериорного распределения

Заикин Артем Александрович

Аспирант

Казанский (Приволжский) федеральный университет, Казань, Россия

E-mail: Kaskrin@gmail.com

Пусть наблюдается выборка $\mathbf{X} = (X_1, ..., X_n)^T$ с наблюдениями из $\mathfrak{X} \subseteq \mathbb{R}$. Пусть далее распределение $\mathbf{P}_\theta$ наблюдения происходит из класса распределений, индексированного одномерным параметром $\theta \in \Theta \subseteq \mathbb{R}$. Кроме того, у распределения с.в. X_1 существует плотность $p(x|\theta) = d\mathbf{P}_\theta/d\nu$. Обозначим $p_n(\mathbf{X}|\theta) = \prod_{i=1}^n p(X_i|\theta)$. Пусть θ есть реализация случайной величины ϑ из распределения $\mathbf{G}$ с плотностью $g(\theta)$ по лебеговской мере. Обозначим символом $\mathbf{P}$ совместную меру распределения ϑ и $\mathbf{X}$. Тогда $\mathbf{P}_\theta$ будет условным распеделением $\mathbf{X}$ при значении $\vartheta = \theta$.

Известный результат Бернштейна-фон Мизеса гласит о том, что при довольно слабых условиях условное распределение параметра ϑ при значении выборки $\mathbf{X} = x^{(n)}$ (апостериорное распределение) асимптотически нормально, причем предел не зависит от априорного распределения $\mathbf{G}$. Иногда точности этого утверждения может быть не достаточно, и поэтому возникает необходимость в асимптотических разложениях.

Ставится задача найти асимптотическое разложение апостериорного распределения по степеням $n^{-1/2}$ в некоторой точке θ_0, а именно, для заданного k найти такие величины $H_m\left(A, x^{(n)}\right)$, что для любого $\varepsilon > 0$ существует $C > 0$ такое, что

$$\sup_n \mathbf{P}_{\theta_0}\left\{\left|\mathbf{P}\left(\sqrt{n}\vartheta \in A|\mathbf{X}\right) - \sum_{m=0}^{k} n^{-m/2} H_m\left(A, \mathbf{X}\right)\right| > Cn^{-(k+1)/2}\right\} < \varepsilon,$$

$$\sup_n \mathbf{P}_{\theta_0}\left\{\left|H_m\left(A, \mathbf{X}\right)\right| > Cn^{-m/2}\right\} < \varepsilon,\ 0 \le m \le k.$$

К настоящему моменту было получено несколько результатов, касающихся асимптотических разложений в похожих условиях. Однако все они касались оценок параметра (так, что $A = A(\mathbf{X})$). Стоит отметить работы Джонсона, Гусева и Венга, проведенные в этом направлении.

Пусть $\Phi(x)$ – функия распределения стандартного нормального закона. $\mathbf{E}_\theta$ – математическое ожидание, отвечающее $\mathbf{P}_\theta$. Кроме того,

$$l_i(\theta) = \ln p(x_i|\theta), \quad \Delta_m(\theta) = \sum_{i=1}^n l_i^{(m)}(\theta), \quad \pi_m(\theta) = \frac{d^m}{d\theta^m}\ln g(\theta),$$

$$I(\theta) = \mathbf{E}_\theta\left(l_1^{(1)}(\theta)\right)^2, \quad \Phi_\Delta(x) = \Phi\left(\left(x - \frac{\Delta_1(\theta_0)}{\sqrt{n}I(\theta_0)}\right)\sqrt{I(\theta_0)}\right).$$

Перечислим условия, налагаемые на плотности p и g, а также особые условия в фиксированной точке $\theta_0 \in \mathring{\Theta}$:

1) Для $\theta \neq \lambda$ выполняется $\int |p(x|\theta) - p(x|\lambda)|\,\nu(dx) > 0$.
2) Существует $\delta_1 > 0$ такое, что $\sup_\Theta |\theta - \theta_0|^{\delta_1} \int \sqrt{p(x|\theta)p(x|\theta_0)}\nu(dx) < \infty$.
3) Функция $p(x|\theta)$ определена для каждого $x \in \mathfrak{X}$, непрерывна на $\overline{\Theta}$, имеет $k+3$ ($k \ge 1$) непрерывных производных на Θ.
4) $\mathbf{E}_{\theta_0}\left(l_1^{(2)}\right)^2 < \infty$.
5) При $i = 3, ..., k+3$ $\mathbf{E}_{\theta_0}\left|l_1^{(i)}\right| < \infty$.

6) Существует $\delta_2 \geq 0$ такое, что $\sup_{\theta\in\Theta}(1+|\theta|^{\delta_2})^{-1}I(\theta) < \infty$.
7) $I(\theta) > 0$ для всех $\theta \in \Theta$.
8) g имеет $k+1$ непрерывные производные на Θ.

Введем дополнительные обозначения. $I_1(m,v)$ – множество индексов $(i_1,...,i_m)$, которые удовлетворяют

$$\sum_{j=1}^{m} i_j = v,\ i_j > 0,\ j = \overline{1,m}.$$

Пусть далее

$$S_n(z) = \sqrt{n}\,(\Theta - \theta_0) \cap (-\infty, z); \quad F_m(z) = \int_{S_n(z)} h^m d\Phi_\Delta(h).$$

Основной результат доклада: при выполнении условий 1)-8) существует $n_0 \in \mathbb{N}$, что для всех $n > n_0$ справедливо разложение апостериорного распределения $\sqrt{n}\vartheta$ при значениях выборки $\mathbf{X} = x^{(n)}$:

$$\mathbf{P}\left(\sqrt{n}(\vartheta-\theta_0) < z|\mathbf{X} = x^{(n)}\right) = V_0^{-1}\left(\Phi_\Delta(z) + \sum_{m=1}^{k} n^{-m/2}H_m(z)\right) + n^{-(k+1)/2}\xi,$$

где

$$H_m(z) = f_m(z) + \sum_{j=1}^{m-1} V_j f_{m-j}(z), \quad m = \overline{1,k},$$

$$V_0 = \Phi_\Delta(\infty),$$

$$V_m = \sum_{j=1}^{m}(-1)^j V_0^{-j} \sum_{I_1(j,m)} \prod_{l=1}^{j} f_{i_l}(\infty), \quad m = \overline{1,k},$$

$$f_m(z) = \sum_{j=1}^{m} \frac{1}{j!} \sum_{I_1(j,m)} \prod_{l=1}^{j} q_{i_l}(z), \quad m = \overline{1,k},$$

$$q_1(z) = F_1(z)\pi_1(\theta_0) + F_2(z)\frac{\Delta_2(\theta_0) + nI(\theta_0)}{2\sqrt{n}} + F_3(z)\frac{\Delta_3(\theta_0)}{6n},$$

$$q_m(z) = F_m(z)\frac{\pi_m(\theta_0)}{m!} + F_{m+2}(z)\frac{\Delta_{m+2}(\theta_0)}{n(m+2)!}, \quad m = \overline{2,k},$$

причем $H_m(z)$ для всех $z \in \mathbb{R}$ и остаток ξ – равномерно по n ограниченные случайные величины.

Источники и литература

1) R.A. Johnson, Asymptotic expansions associated with posterior distributions. // The Annals of Mathematical Statistics, Vol. 41, No. 3, 1970 – 851-864.
2) С.И.Гусев, Асимптотические разложения, связанные с некоторыми статистическими оценками в гладком случае. I. Разложения случайных величин. // Теория вероятностей и ее применения, Том XX, 3. – М.: Наука, 1975. – 488-514.
3) R.C.Weng, C.-H.Hsu, A Study of Expansions of Posterior Distributions // Communications in Statistics—Theory and Methods, 42, 2013 – 346–364. 2010 – 741-764.