0.1 Асимптотическая нормальность

Пусть заданы объекты из некоторого распределения:

$$\mathbf{X}^n = \{X_i\}_{i=1}^n,$$

где n число объектов.

Пусть задано некоторое открытое подмножество $\mathbf{\Theta} \in \mathbb{R}^d$. Подмножество $\mathbf{\Theta}$ задает множество статистических моделей $\mathcal{P}^n = \{P_\theta^n | \theta \in \mathbf{\Theta}\}$. Пусть для каждого n существует мера P_0^n которая доминирует все меры из множества $\mathcal{P}^n$. Пусть также все меры задаются своей плотностью p_θ^n.

Определение 1. *Рассмотрим некоторую внутреннюю точку $\theta^* \in \mathbf{\Theta}$ и последовательность $\delta_n \to 0$. Пусть существует вектор $\Delta_{\theta^*}^n$ и невырожденная матрица $\boldsymbol{V}_{\theta^*}$, такие, что последовательность $\{\Delta_{\theta^*}^n\}$ ограничена по вероятностной мере, а также для любого компакта $K \subset \mathbb{R}^d$ выполняется:*

$$\sup_{h \in K} \Big| \log \frac{p^n_{\theta^* + \delta_n h}}{p^n_{\theta^*}}\big(\boldsymbol{X}^n\big) - h^\mathsf{T} \boldsymbol{V}_{\theta^*} \Delta^n_{\theta^*} - \frac{1}{2} h^\mathsf{T} \boldsymbol{V}_{\theta^*} h \Big| \xrightarrow{P_0^n} 0.$$

Тогда модель $\mathcal{P}^n$ удовлетворяет условия локальной асимптотической нормальности в точке θ^* (local asymptotic normality).

Априорное распределение заданное на множестве $\mathbf{\Theta}$ обозначим Π, а его плотность π. Предположим, что π положительно в некоторой окрестности точки θ^*.

Апостериорное распределение построенное на основе множестве объектов $\mathbf{X}^n$ обозначим $\Pi_n\big(A|\mathbf{X}^n\big)$, где A некоторое борелевское множество. Будем обозначать случайную величину из апостериорного распределения как ϑ.

0.2 Теорема Бернштейна фон Мизеса

Теорема 1. *Пусть для некоторой точки θ^* выполено условия локальной асимптотической нормальности (Опр.0.1). Пусть задано априорное распределение Π. Пусть для некоторой последовательности чисел $M_n \to \infty$ выполняется следующее условие:*

$$P_0^n \Pi_n\big(||\vartheta - \theta^*|| > \delta_n M_n | \boldsymbol{X}^n\big) \to 0. \tag{0.1}$$

Тогда последовательность апостериорных распределений сходится к последовательности нормальных:

$$\sup_B \Big| \Pi_n\big(\frac{\vartheta - \theta^*}{\delta_n} \in B | \boldsymbol{X}^n\big) - N_{\Delta^n_{\theta^*}, V^{-1}_{\theta^*}}\big(B\big) \Big| \xrightarrow{P_0^n} 0.$$

Доказательство. Апостериорное распределение для величины $H = \frac{\vartheta - \theta^*}{\delta_n}$ полученное для выборки $\mathbf{X}^n$ обозначим Π_n. Также обозначим $N_{\Delta^n_{\theta^*}, V^{-1}_{\theta^*}}$ как Φ_n. Рассмотрим некоторый компакт $K \subset \mathbb{R}^d$. Рассмотрим условное апостериорное распределение:

$$\Pi_n^K\big(B|\mathbf{X}^n\big) = \Pi_n\big(B \cap K|\mathbf{X}^n\big) / \Pi_n\big(K|\mathbf{X}^n\big),$$
$$\Phi_n^K\big(B\big) = \Phi_n\big(B \cap K\big)/\Phi_n\big(K\big).$$

Рассмотрим некоторый компакт $K \subset \mathbb{R}^d$. Для любой окрестности $U\big(\theta^*\big) \subset \mathbf{\Theta}$ существует некоторый номер n, такой, что $\theta^* + K\delta_n \subset U\big(\theta^*\big)$.

Рассмотрим функцию $f_n : K \times K \to \mathbb{R}$:

$$f_n\big(g, h\big) = \left(1 - \frac{\phi_n\big(h\big) s_n\big(g\big) \pi_n\big(g\big)}{\phi_n\big(g\big) s_n\big(h\big) \pi_n\big(h\big)}\right)_+,$$

где ϕ_n, π_n — распределение Φ_n и Π_n соответственно, s_n является отношением правдоподобия:

$$s_n(h) = \frac{p^n_{\theta^*+h\delta_n}}{p^n_{\theta^*}}.$$

Рассмотрим две произвольные последовательности $\{h_n\}, \{g_n\} \subset K$:

$$\begin{aligned}
&\log\frac{\phi_n(h_n)s_n(g_n)\pi_n(g_n)}{\phi_n(g_n)s_n(h_n)\pi_n(h_n)} = \\
=&(g_n-h_n)^{\mathsf{T}}\mathbf{V}_{\theta^*}\Delta^n_{\theta^*} + \frac{1}{2}h^{\mathsf{T}}\mathbf{V}_{\theta^*}h_n - \frac{1}{2}g_n^{\mathsf{T}}\mathbf{V}_{\theta^*}g_n + o(1) - \\
=&-\frac{1}{2}(h_n-\Delta^n_{\theta^*})^{\mathsf{T}}\mathbf{V}_{\theta^*}(h_n-\Delta^n_{\theta^*}) + \frac{1}{2}(g_n-\Delta^n_{\theta^*})^{\mathsf{T}}\mathbf{V}_{\theta^*}(g_n-\Delta^n_{\theta^*}) = o(1),
\end{aligned} \tag{0.2}$$

где первое слагаемое получено используя локальную асимптотическую нормальность (Опр.0.1), а второе с плотности нормального распределения. Тогда из (0.2) получаем, что:

$$\sup_{g,h\in K} f_n(g,h) \xrightarrow{P_0}_{n\to\infty} 0. \tag{0.3}$$

Обозначим за Ξ_n событие, что $\Pi_n(K) > 0$. Рассмотрим некоторое $\eta > 0$, которое задает следующее множество:

$$\Omega_n = \{\sup_{g,h\in K} f_n(g,h) \le \eta\}_*, \tag{0.4}$$

где $*$ обозначает измеримое покрытие множества. Из (0.3) и (0.4) получаем следующее неравенство:

$$P_0^n||\Pi_n^K - \Phi_n^K||\mathbb{I}_{\Xi_n} \le P_0^n||\Pi_n^K - \Phi_n^K||\mathbb{I}_{\Xi_n\cap\Omega_n} + 2P_0^n||\Pi_n^K - \Phi_n^K||\mathbb{I}_{\Xi_n\setminus\Omega_n}, \tag{0.5}$$

где $\mathbb{I}_{\Xi_n}$ — индикаторная функция, $||\cdot||$ является вариационной нормой (total-variational norm). Второе слагаемое равняется нулю в силу (0.3). Используя свойство данной нормы первое слагаемое принимает следующий вид:

$$\begin{aligned}
\frac{1}{2}P_0^n||\Pi_n^K - \Phi_n^K||\mathbb{I}_{\Xi_n\cap\Omega_n} &= P_0^n\int_K\left(1-\frac{d\Phi_n^K}{d\Pi_n^K}\right)_+ d\Pi_n^K\mathbb{I}_{\Xi_n\cap\Omega_n} = \\
&= P_0^n\int_K\left(1-\int_K\frac{s_n(g)\pi_n(g)\phi_n^K(h)}{s_n(h)\pi_n(h)\phi_n^K(g)}d\Phi_n^K(g)\right)_+ d\Pi_n^K\mathbb{I}_{\Xi_n\cap\Omega_n}.
\end{aligned}$$

Используя неравенство Йенсена, а также (0.3) получаем следующее:

$$\frac{1}{2}P_0^n||\Pi_n^K - \Phi_n^K||\mathbb{I}_{\Xi_n\cap\Omega_n} \le P_0^n\int\left(1-\frac{s_n(g)\pi_n(g)\phi_n^K(h)}{s_n(h)\pi_n(h)\phi_n^K(g)}\right)_+ d\Phi_n^K(g)d\Pi_n^K\mathbb{I}_{\Xi_n\cap\Omega_n} \le \eta.$$

Подставляя в (0.5) получаем, что для любого компакта $K \subset \mathbb{R}^d$ выполняется, что $P_0^n||\Pi_n^K - \Phi_n^K||\mathbb{I}_{\Xi_n} \to 0$.

Рассмотрим последовательность шаров $\{K_m\}$ с центом в нуле с радиусом M_m, причем $M_m \to \infty$.

Рассмотрим множество $\{\Xi_n | \Xi_n = \{\Pi_n(K_n) > 0\}\}$, по условию теоремы (0.1) получим, что $P_0^n(\Xi_n) \to 0$. Также получаем, что $P_0^n||\Pi_n^{K_n} - \Phi_n^{K_n}|| \to 0$.

Теперь рассмотрим $P_0^n||\Pi_n - \Phi_n||$:

$$\begin{aligned}
P_0^n||\Pi_n - \Phi_n|| &\le P_0^n||\Pi_n - \Pi_n^{K_n}|| + P_0^n||\Phi_n - \Phi_n^{K_n}|| \\
&\le 2\left(\Pi(\mathbb{R}^d\setminus K_n)\right) + 2\left(\Phi(\mathbb{R}^d\setminus K_n)\right) \to 0,
\end{aligned} \tag{0.6}$$

так как увеличивая радиус компакта в бесконечность мы покроем все множество $\mathbb{R}^d$. Выражение (0.6) заканчивает доказательство данной теоремы. □

Список литературы

[1] *Kleijn, B. J. K., and van der Vaart, A. W.* (2012). The Bernstein-Von-Mises theorem under misspecification. Electronic Journal of Statistics, 6, 354-381. https://doi.org/10.1214/12-EJS675