# 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 $p_\theta^n$.

**Определение 1.1.** *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:*

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

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

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

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 $\vartheta$.

# 2 Теорема Бернштейна фон Мизеса

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

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

*Тогда последовательность апостериорных распределений сходится к последовательности нормальных:*

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

*Доказательство.* Апостериорное распределение для величины $H = \frac{\vartheta - \theta^*}{\delta_n}$ полученное для выборки $\mathbf{X}^n$ обозначим $\Pi_n$. Также обозначим $N_{\Delta_{\theta^*}^n, V_{\theta^*}^{-1}}$ как $\Phi_n$. 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 $U(\theta^*) \subset \mathbf{\Theta}$ существует некоторый номер $n$, такой, что $\theta^* + K\delta_n \subset U(\theta^*)$.

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

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

где $\phi_n, \pi_n$ — распределение $\Phi_n$ и $\Pi_n$ соответственно, $s_n$ является отношением правдоподобия:

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

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

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

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

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

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

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

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

$$P^n_0||\Pi^K_n - \Phi^K_n||\mathbb{I}_{\Xi_n} \leq P^n_0||\Pi^K_n - \Phi^K_n||\mathbb{I}_{\Xi_n\cap\Omega_n} + 2P^n_0||\Pi^K_n - \Phi^K_n||\mathbb{I}_{\Xi_n\setminus\Omega_n},$$

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

$$\begin{aligned}
\frac{1}{2}P^n_0||\Pi^K_n - \Phi^K_n||\mathbb{I}_{\Xi_n\cap\Omega_n} &= P^n_0\int_K \left(1 - \frac{d\Phi^K_n}{d\Pi^K_n}\right)_+ d\Pi^K_n \mathbb{I}_{\Xi_n\cap\Omega_n} = \\
&= P^n_0\int_K \left(1 - \int_K \frac{s_n(g)\pi_n(g)\phi^K_n(h)}{s_n(h)\pi_n(h)\phi^K_n(g)} d\Phi^K_n(g)\right)_+ d\Pi^K_n \mathbb{I}_{\Xi_n\cap\Omega_n}.
\end{aligned}$$

Используя неравенство Йенсена, а также (**??**) получаем следующее:

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

Подставляя в (**??**) 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 $\big\{\Xi_n|\Xi_n = \{\Pi_n\big(K_n\big) > 0\}\big\}$, по условию теоремы (**??**) получим, что $P_0^n\big(\Xi_n\big) \to 0$. Также получаем, что $P_0^n||\Pi_n^{K_n} - \Phi_n^{K_n}|| \to 0$.

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

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

так как увеличивая радиус компакта в бесконечность мы покроем все множество $\mathbb{R}^d$. Выражение (**??**) 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`