14.02.20. "Fundamentals of Stein's method", Nathan Ross

На семинаре мы обсудили основные определения и утверждения, связанные с методом Штейна для нормальной аппроксимации.

Определение 1 *Метрикой (расстоянием) между двумя вероятностными распределениями μ и ν называется*

$$d_{\mathcal{H}}(\mu, \nu) = \sup_{h \in \mathcal{H}} \left| \int h(x) d\mu(x) - \int h(x) d\nu(x) \right|,$$

где $\mathcal{H}$ – некоторое семейство тестовых функций.

В частности, мы будем работать со следующими метриками:

1. Если $\mathcal{H} = \{\mathbb{I}[\cdot \leqslant x] : x \in \mathbb{R}\}$, то метрика называется Колмогоровской, и обозначается d_{K}.
2. Если $\mathcal{H} = \{h : \mathbb{R} \to \mathbb{R} : |h(x) - h(y)| \leqslant |x - y|\}$ (липшицевы фукнции с константой Липшица $L = 1$), то метрика называется метрикой Вассерштейна и обозначается d_{W}.
3. Если $\mathcal{H} = \{\mathbb{I}[A \in \mathbb{R}] : A \in \mathfrak{B}(\mathbb{R})\}$, то метрика называется метрикой полной вариации, и обозначается d_{TV}.

Они связаны следующими соотношениями:

Предложение 1 *1. Для случайных величин W и Z, $d_{\mathrm{K}}(W, Z) \leqslant d_{\mathrm{TV}}(W, Z)$.*

2. Если с.в. Z имеет плотность, ограниченную константой C, то для любой с.в. W,

$$d_{\mathrm{K}}(W, Z) \leqslant \sqrt{2C d_{\mathrm{W}}(W, Z)}.$$

Далее мы ввели оператор

$$\mathcal{A}f(x) = f'(x) - xf(x),$$

называющийся характеристическим для стандартного нормального распределения. Название обусловлено тем, что

$$W \sim \mathcal{N}(0, 1) \Leftrightarrow \mathbb{E}\mathcal{A}f(W) = 0 \text{ для любой абс. непр. } f \text{ с } \mathbb{E}|f'(Z)| < \infty.$$

Оказалось так же, что если за f_h для любой $h \in \mathcal{H}$ обозначить решение уравнения

$$f_h'(w) - w f_h(w) = h(w) - \Phi(h),$$

где $\Phi(h) = \mathbb{E}h(Z)$ для $Z \sim \mathcal{N}(0,1)$, то для произвольной с.в. W и $Z \sim \mathcal{N}(0,1)$, расстояние между ними можно выразить в терминах функции f_h:

$$d_{\mathcal{H}}(X,Y) = \sup_{h\in\mathcal{H}} |\mathbb{E}h(X) - \mathbb{E}h(Y)| = \sup_{h\in\mathcal{H}} |\mathbb{E}\left[f_h'(W) - W f_h(W)\right]|.$$

Далее идея метода заключается в том, чтобы подобрать хорошую оценку для этого мат. ожидания, используя свойства решения f_h и информацию о с.в. W. Имеются оценки на f_h:

Лемма 1 *1. Если h ограничена, то*

$$\|f_h\| \leqslant \sqrt{\frac{\pi}{2}}\|h(\cdot) - \Phi(h)\|, \quad \text{и} \quad \|f_h'\| \leqslant 2\|h(\cdot) - \Phi(h)\|$$

2. Если h абсолютно непрерывна, то

$$\|f_h\| \leqslant 2\|h'\|, \quad \|f_h'\| \leqslant \sqrt{\frac{2}{\pi}}\|h'\|, \quad \text{и} \quad \|f_h''\| \leqslant 2\|h'\|.$$

Отсюда, если W – с.в. и $Z \sim \mathcal{N}(0,1)$, то

$$d_{\mathrm{W}}(W,Z) \leqslant \sup_{f\in\mathcal{F}} |\mathbb{E}\left[f'(W) - W f(W)\right]|,$$

где $\mathcal{F} = \left\{ f : \|f\|, \|f''\| \leqslant 2, \|f'\| \leqslant \sqrt{2/\pi} \right\}$.

В конце, используя полученные утверждения, мы получили оценку для суммы независимых случаных величин.

Теорема 1 *Пусть $X_1, \ldots, X_n$ – независимые случайные величины с $\mathbb{E}|X_i|^4 < \infty, \mathbb{E}X_i = 0$ и $\mathbb{E}X_i^2 = 1$. Пусть $Z \sim \mathcal{N}(0,1)$. Обозначим $W = \left(\sum_{i=1}^n X_i\right)/\sqrt{n}$. Тогда имеет место оценка:*

$$d_{\mathrm{W}}(W,Z) \leqslant \frac{1}{n^{3/2}} \sum_{i=1}^n \mathbb{E}|X_i|^3 + \frac{\sqrt{2}}{\sqrt{\pi}n} \sqrt{\sum_{i=1}^n \mathbb{E}\left[X_i^4\right]}.$$

Автор замечает, что требование наличия четвертых моментов можно ослабить, но структура приводимого доказательства в дальнейшем будет повторяться, поэтому формулирует и доказывает именно его.