Секция «Теория вероятностей и математическая статистика»

Асимптотически d-оптимальные оценки

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 $\mathbf{P}$ совместное распределение ϑ и $\mathbf{X}$. Тогда $\mathbf{P}_\theta$ будет условным распределением $\mathbf{X}$ при значении $\vartheta = \theta$.

Пусть задана некоторая функция потерь на $L : \Theta \times \Theta \mapsto \mathbb{R}$, такая, что $L(\theta, \theta) = 0$. Для произвольной оценки T_n можно определить d-риск как величину $\mathcal{R}(T_n) = \mathbf{E}\{L(T_n, \vartheta)|T_n\}$ (в данной записи предполагается, что соответствующая регулярная вероятность существует). Оценка θ_n^* называется оценкой с равномерно минимальным d-риском (ОРМdР), если $\mathbf{P}(\mathcal{R}(\theta_n^*) \leq \mathcal{R}(T_n)) = 1$ для любой другой оценки T_n. Общая теория и подход к построению подобных оценок указаны в статье (Simushkin, 1983).

К сожалению, работа с ОРМdР сопряжена со значительными трудностями. В частности вопрос о широких условиях существования оценки остается открытым. Поэтому имеет смысл работать с оценками, которые близки к ОРМdР в плане значений d-риска. Пусть задана некоторая функция $\phi : \mathbb{N} \mapsto \mathbb{R}$, такая, что $\phi(n) \to 0, n \to \infty$. Определим оценку с $\phi(n)$-асимптотически равномерно минимальным d-риском в классе оценок $\mathcal{K}$ как оценку θ_n^*, которая удовлетворяет $\mathbf{P}(\mathcal{R}(\theta_n^*) \leq \mathcal{R}(T_n) + \phi(n)) \to 1, \forall T_n \in \mathcal{K}$. Функцию $\phi(n)$ можно понимать как уровень допустимого отклонения функции d-риска асимптотически оптимальной оценки от истинно оптимального уровня.

В докладе рассматривается степенная функция потерь $L(d, \theta) = |d - \theta|^k, k > 1$. Основным результатом является следующая

Теорема 1. При выполнении определенных условий регулярности оценка максимального правдоподобия $\widehat{\theta}_n$ является оценкой с $n^{-k/2}$-асимптотически равномерно минимальным d-риском среди всех оценок T_n, которые удовлетворяют

$$\forall \varepsilon > 0 \ \exists b > 0 : \sup_{\theta \in \Theta} \sup_n \mathbf{P}_\theta \left(\sqrt{n}|T_n - \theta| > b \right) < \varepsilon.$$

Ограничение на класс $\sqrt{n}$-состоятельности оценок довольно естественно, однако не обязательно. Подобное утверждение справедливо и для класса всех оценок параметра θ (в том числе и несостоятельных в привычном понимании этого слова), однако скорость сходимости меньше.

Теорема 2. При выполнении определенных условий регулярности существует последовательность $\varepsilon_n \to 0$ такая, что оценка максимального правдоподобия $\widehat{\theta}_n$ является оценкой с ε_n-асимптотически равномерно минимальным d-риском среди всех оценок T_n.

Источники и литература

1) Simushkin S.V., Volodin I.N. Statistical inference with a minimal d-risk. // Lect. Notes Math., No. 1021, 1983~— 629–636.