



Сравнение оценок, существование наилучшей оценки

Пусть X_1, \dots, X_n — выборка объема n из параметрического семейства распределений F_θ , где $\theta \in \Theta$.

Говорят, что оценка θ^* лучше оценки θ_1 в смысле среднеквадратического подхода, если для любого $\theta \in \Theta$

$$E_\theta (\theta_1^* - \theta)^2 \leq E_\theta (\theta_1 - \theta)^2,$$

и хотя бы при одном θ это неравенство строгое.

Существует ли среди всех оценок наилучшая в смысле среднеквадратического подхода? Скептик сразу ответит «нет». Покажем, что он прав. Предположим, что мы имеем дело с невырожденной задачей: ни для какой статистики θ^* невозможно тождество: $\theta^* = \theta$ при любых $\theta \in \Theta$.

В классе всех возможных оценок наилучшей в смысле среднеквадратического подхода оценки не существует.

Доказательство

Доказательство теоремы 4. Пусть, напротив, θ^* — наилучшая, то есть для любой другой оценки θ_1 , при любом $\theta \in \Theta$ выполнено

$$E_\theta (\theta^* - \theta)^2 \leq E_\theta (\theta_1 - \theta)^2.$$

Пусть θ_1 — произвольная точка Θ . Рассмотрим статистику $\theta_1^* \equiv \theta_1$. Тогда

$$E_\theta (\theta^* - \theta)^2 \leq E_\theta (\theta_1 - \theta)^2 \text{ при любом } \theta \in \Theta.$$

В частности, при $\theta = \theta_1$ получим $E_{\theta_1} (\theta^* - \theta_1)^2 \leq E_{\theta_1} (\theta_1 - \theta_1)^2 = 0$. Поэтому $E_{\theta_1} (\theta^* - \theta_1)^2 = 0$. Но, поскольку θ_1 произвольно, при любом $\theta \in \Theta$ выполняется $E_\theta (\theta^* - \theta)^2 = 0$. А это возможно только если $\theta^* \equiv \theta$ (оценка в точности отгадывает неизвестный параметр), т.е. для вырожденной с точки зрения математической статистики задачи. \square

Вырожденными являются, например, следующие задачи:

- * для выборки из I_θ , $\theta \in \mathbb{R}$, выполнено тождество $X_1 \equiv \theta$;
- * для выборки из $U_{\theta, \theta+1}$, $\theta \in \mathbb{Z}$, выполнено тождество $[X_1] \equiv \theta$.

Упражнение. Объяснить словесно доказательство теоремы 4.

Если в классе всех оценок наилучшей не существует, то, возможно, следует разбить класс всех оценок на отдельные подклассы и в каждом искать наилучшую.

Обычно рассматривают оценки, имеющие одинаковое **смещение**

$$b(\theta) = E_\theta \theta^* - \theta.$$

Обозначим через $K_b = K_{b(\theta)}$ класс оценок, имеющих смещение, равное заданной функции $b(\theta)$:

$$K_b = \{\theta^* : E_\theta \theta^* = \theta + b(\theta)\}, \quad K_0 = \{\theta^* : E_\theta \theta^* = \theta\}.$$

Здесь K_0 — класс несмещенных оценок.

Определение 9.

Оценка $\theta^* \in K_b$ является эффективной оценкой в классе K_b , если она лучше (не хуже) всех других оценок класса K_b в смысле среднеквадратического подхода. То есть для любой $\theta_1^* \in K_b$, для любого $\theta \in \Theta$

$$E_\theta (\theta^* - \theta)^2 \leq E_\theta (\theta_1^* - \theta)^2.$$

https://tvims.nsu.ru/chernova/ms/ms_2006.pdf (глава 3)

<https://portal.edu.asu.ru/pluginfile.php/18909/course/section/10118/thstat3.pdf?ysclid=l4rgpq2dx334407587>

