

Мартынюк И.А.

Национальный авиационный университет, Киев

ОПТИМИЗАЦИЯ ПАРАМЕТРА СГЛАЖИВАНИЯ ПРИ ЯДЕРНОМ ОЦЕНИВАНИИ ПЛОТНОСТИ ВЕРОЯТНОСТИ

Для построения гладкой оценки плотности распределения по выборке

$$\eta_1, \dots, \eta_N$$

используются ядерные оценки:

$$f(x) = \frac{1}{N \cdot h} \cdot \sum_{n=1}^N K\left(\frac{x - \eta_n}{h}\right),$$

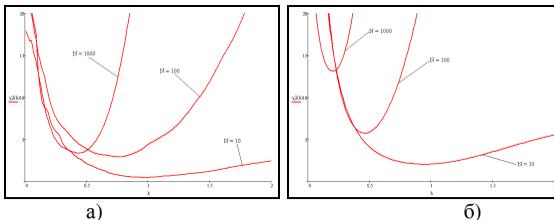
где $K(x)$ – функция-ядро, обладающая свойством $\int_{-\infty}^{\infty} K(x) dx = 1$, N – объем выборки, h – параметр сглаживания или параметр масштаба, η_n – n -й отсчет из выборки.

Широкое применение при построении ядерных оценок находят прямоугольное и гауссовское ядро:

$$K(x) = \begin{cases} 0.5 & \text{if } -1 \leq x \leq 1, \\ 0 & \text{otherwise;} \end{cases} \quad K(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{x^2}{2}\right).$$

Качество восстановленной плотности зависит от выбора величины параметра h . Таким образом, возникает оптимизационная задача выбора параметра сглаживания. В работе проведено моделирование выборок различных объемов из нормального распределения ($\mu=0, \sigma=1$) и исследована зависимость качества оценки плотности распределения от параметра масштаба ядра по критерию согласия хи-квадрат.

Результаты моделирования показывают, что в этой зависимости наблюдается минимум, который и определяет оптимальное значение параметра h .



Графики зависимости качества оценки от параметра сглаживания при разных объемах выборки: а) – прямоугольное ядро; б) – гауссовское ядро

Научный руководитель – И.Г.Прокопенко, д.т.н., проф.