Секция «Теория вероятностей и математическая статистика»

Оценка плотности в условиях мультипликативного шума

Научный руководитель – Шкляев Александр Викторович

Потапова Евдокия Сергеевна
Студент (специалист)
Московский государственный университет имени М.В.Ломоносова,
Механико-математический факультет, Кафедра математической статистики и случайных процессов, Москва, Россия
E-mail: evdokiapotapova@yandex.ru

Рассматривается математическая модель задачи, возникающей при проведении химических исследований. Наблюдаемыми объектами в опытах выступают движущиеся частицы примесей в жидкости, имеющие размеры порядка 10^{-8} метра. Жидкость просвечивают лучом лазера, попадающие в луч частицы фиксируются на камеру. В каждый момент времени на диафрагме появляется изображение, для каждой частицы совокупность таких изображений составляет проекцию движения частицы на площадь камеры. Размер изображений напрямую не связан с размерами частиц.
Наблюдая за движением изображений на диафрагме, мы можем оценить векторы перемещений частиц в плоскости камеры. Для каждой отдельно взятой частицы эти перемещения образуют двумерное броуновское движение с нулевым сносом и дисперсией $\sigma^2 = c/d$, где c — некоторая константа, d — размер частицы.
Задача, которая ставится на основе данных, полученных с камеры, состоит в том, чтобы оценить линейные размеры d частиц примесей.
В итоге задача приводится к следующей вероятностной постановке: известна выборка $Z_1, \dots, Z_n$ - н.о.р. случайные величины, причем: $Z_i = X_i Y_i$, где $Y_i \sim \chi^2_k$. Цель исследования состоит в том, чтобы оценить плотность X_i.
Решение данной задачи опирается на *метод Фурье* оценки плотности распределения ([1],[2]). В этом методе строится оценка плотности вида:

$$\hat{f}(x) = \sum_{\nu=0}^{K} \hat{f}_\nu \psi_\nu(x),$$

где $X_1, \dots, X_n$ - н.о.р. случайные величины с плотностью f с носителем на отрезке $[-\pi, \pi]$, ψ_ν - базис, состоящий из ортогональных функций, K - параметр,

$$\begin{cases} \hat{f}_0 = \frac{1}{2\pi n} \sum_{k=1}^{n} \psi_0(X_k), \\ \hat{f}_\nu = \frac{1}{\pi n} \sum_{k=1}^{n} \psi_\nu(X_k), \qquad \nu = 1, 2, \dots \ , \end{cases} \tag{1}$$

Предлагаемый способ отделения шума заключается в подборе некой функции, которая при применении к зашумленным данным сводит их к незашумленным: $g_\nu(t)$: $Eg_\nu(Z) = E\psi(\nu X), \nu = 0, 1, 2, \dots$. В результате такой операции коэффициенты ряда Фурье записывается в виде:

$$\begin{cases} \hat{f}_0 = \frac{1}{2\pi n} \sum_{k=1}^{n} g_0(Z_k), \\ \hat{f}_\nu = \frac{1}{\pi n} \sum_{k=1}^{n} g_\nu(Z_k), \qquad \nu = 1, 2, \dots \ . \end{cases} \tag{2}$$

После получения оценки плотности возникает ряд задач, среди которых приведение плотности к виду вероятностной ([3]), а также правильный выбор параметра K ([4]).
Описанный метод реализован и исследован на модельных данных.

Источники и литература

1) Silverman, B. W., Density estimation for statistics and data analysis, 1952
2) Wasserman, L., All of nonparametric statistics, 2006
3) Ingrid K.Glad, Nils Lid Hjort, Nikolai G.Ushakov, Correction of density estimators that are not densities, Scandinavian Journal of Statistics, Vol. 30, No. 2 (Jun., 2003), pp. 415-427
4) Jeffrey D. Hart, On the choice of a truncation point in Fourier series density estimation, Journal of Statistical Computation and Simulation, Vol. 21, 1985, pp. 95-116