

Ядерные оценки плотности. Непараметрическая регрессия the equivalence

Оценка Розенבלата-Парзена

$$\hat{f}(t) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{t - x_i}{h}\right)$$

При этом Функция K
неотрицательная
симметричная
 $\int_{-\infty}^{\infty} K(t) dt$

То есть Функция K - плотность распределения.

$$\int_{-\infty}^{\infty} (\hat{f}(t) - f(t))^2 dt$$

$$\int_{-\infty}^{\infty} |\hat{f}(t) - f(t)| dt$$

Ядро Епанечникова

$$K(t) = \frac{3}{4\sqrt{5}} \left(1 - \frac{z^2}{5}\right)$$

Ширина окна - ?
Постоянная или переменная, зависящая от данных

Функция K
неотрицательная ?????
симметричная ????? Краевые эффекты
 L^1 норма равна единице

Кластеризация

расстояние до k-го ближайшего соседа `spacing`
Проклятие размерности

Непараметрическая (ядерная) регрессия

$$\hat{f}(t) = \frac{\sum_{i=1}^n y_i K\left(\frac{t - x_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{t - x_i}{h}\right)}$$

Сходится по вероятности к $E(Y|X)$