KERNEL REGRESSION

$\binom{Y_1}{X_1}, \dots, \binom{Y_n}{X_n}$ i.i.d

I. „GLOBAL" POLYNOMIALS:

$m(x) = \mathbb{E}[Y_i | X_i = x] = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^p$

II. Regressogram

- breaks: $a_0 < a_1 < a_2 \dots < a_L$

$\hat{m}_n(x) = \text{For } x \in (a_l, a_{l+1}]$ $\quad \hat{m}_n(x) = \dfrac{\sum_{i: X_i \in (a_l, a_{l+1}]} Y_i}{\#\{i : X_i \in (a_l, a_{l+1}]\}}$

estim. of $m(x)$ $\quad$ estim of $m(x)$

III. Nadaraya - Watson's estimator

- ~~is~~ „locally constant" fitting
- for a kernel K: $\hat{m}_n^{(NW)}(x) = \sum_{i=1}^{n} w_{ni}(x) Y_i$,

where $w_{ni}(x) = \dfrac{K\left(\frac{X_i - x}{h_n}\right)}{\sum_{j=1}^{n} K\left(\frac{X_j - x}{h_n}\right)}$, K - kernel, h_n - bandwidth

If $K(x) = \frac{1}{2} \mathbb{1}\{|x| < 1\}$

$\rightsquigarrow \hat{m}_n^{(NW)}(x) = \dfrac{\sum_{i: X_i \in (x - h_n, x + h_n)} Y_i}{\#\{i : X_i \in (x - h_n, x + h_n)\}}$ $\quad$ ← MOVING AVERAGE

IV. Local - linear estimator

- $\hat{m}_n^{(LL)}(x) = \hat{\beta}_0(x)$, where

$$(\hat{\beta}_0(x), \hat{\beta}_1(x)) = \underset{b_0, b_1}{\arg\min} \sum_{i=1}^{n} (Y_i - b_0 - b_1(X_i - x))^2 K\left(\frac{X_i - x}{h_n}\right)$$

WEIGHTED REGRESSION

If $K(x) = \frac{1}{2} \mathbb{1}\{|x| < 1\}$

$$(\hat{\beta}_0(x), \hat{\beta}_1(x)) = \underset{b_0, b_1}{\arg\min} \sum_{i: X_i \in (x - h_n, x + h_n)} (Y_i - b_0 - b_1(X_i - x))^2$$

VI CHOICE OF h_n

VI -

$$h_n^{(CV)} = \underset{h_n > 0}{\operatorname{argmin}}\ CV(h_n),$$

where $CV(h_n) = \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \hat{m}_{h_n}^{(-i)}(X_i)\right]^2$,

ESTIM OF $m(x)$ based on $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \dots, \begin{pmatrix} X_{i-1} \\ Y_{i-1} \end{pmatrix}, \begin{pmatrix} X_{i+1} \\ Y_{i+1} \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$

VI (B) - NEAREST-NEIGHBOUR BW CHOICE

- Let $K(x) > 0 \Leftrightarrow x \in (-1,1)$
- then choose h_n so that $\{\#\{i : |X_i - x| \le h_n\} \ge k$

$\rightsquigarrow h_n^{(NN)}(x)$ ... the k-th smallest value from $\{|X_1 - x|, \dots, |X_n - x|\}$

CHOICE OF k - transform X_i to ranks R_i, KERNEL

+ CV on $\begin{pmatrix} X_1 \\ R_1 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ R_n \end{pmatrix}$, 2

VII LOWESS FUNCTION

- „robustified" kernel regression — TRICUBE KERNEL
- 0 step: $\hat{m}^{(0)}(x)$ is calculated as $\hat{m}_{h_n}^{(LL)}(x)$ with $k = [nf]$

FOR k step $= 1, \dots, ITER$

$$r_i^{(k)} = Y_i - \hat{m}^{(k-1)}(x),$$

$$\delta_i^{(k)} = B\left(\frac{r_i^{(k)}}{\text{med}\{|r_1^{(k)}|, \dots, |r_n^{(k)}|\}}\right), \quad i = 1, \dots, n$$

$$\hat{m}^{(k)}(x) = \hat{\beta}_0^{(k)}(x), \text{ where}$$

$$\left(\hat{\beta}_0(x), \hat{\beta}_1(x)\right) = \underset{(b_0, b_1)}{\operatorname{argmin}} \sum_{i=1}^{n}\left(Y_i - b_0 - b_1(X_i - x)\right)^2 \cdot K\left(\frac{X_i - x}{h_n(x)}\right)\delta_i^{(k)}$$

VII. CONDITIONAL VARIANCE ESTIMATION $T^2(x) = \mathrm{var}(Y|X)$

1. STEP: ESTIM $m(x) = E[Y|X=x]$
AND CALCULATE RESIDUALS $\hat{\varepsilon}_i = Y_i - \hat{m}(X_i)$

2. STEP, ESTIMATE $\mathrm{var}(Y|X)$ BY FITTING KERNEL
REGRESSION TO $\begin{pmatrix} \hat{\varepsilon}_1^2 \\ X_1 \end{pmatrix}, \dots \begin{pmatrix} \hat{\varepsilon}_n^2 \\ X_n \end{pmatrix}$