Homework 7

Liu Huihang
SA18017026
QQ: 184050473
MAIL: huihang@mail.ustc.edu.cn

Problem 1 (Slide 7 page 6) *Consider Bart Simpson density as following*

$$f(x) = \frac{1}{2}\phi(x;0,1) + \frac{1}{10}\sum_{j=0}^{4}\phi(x;(j/2)-1,1/10)$$

where $\phi(x;\mu,\sigma)$ *denotes a Normal density with mean* μ *and standard deviation* σ*.*

Generate 1000 *sample from Bart Simpson distribution, estimate its density by histogram and compare different interval segmentation method. Estimate density use naive density estimator with image and compare the effect of different bandwidth h.*

Problem 2 (Slide 7 page 46) *Prove the asymptotic normality of KDE* $\hat{f}_h(x)$ *if* f'' *exists and* $h = cn^{-1/5}$.

Solution

We have

$$E[\hat{f}_h(x) - f(x)] = \frac{1}{2}h^2 f''(x)\kappa_{21} + o(h^2)$$

$$Var[\hat{f}_h(x)] = \frac{1}{nh}f(x)\kappa_{02} + o(\frac{1}{nh})$$

$$\frac{\hat{f}_h(x) - E[\hat{f}_h(x)]}{\sqrt{Var[\hat{f}_h(x)]}} \rightsquigarrow N(0,1)$$

Thus, $h = cn^{-1/5}$

$$\frac{\hat{f}_h(x) - f(x) - \frac{1}{2}h^2 f''(x)\kappa_{21}}{\sqrt{\frac{1}{nh}f(x)\kappa_{02}}} \rightsquigarrow N(0,1)$$

$$\frac{\hat{f}_h(x) - f(x)}{\sqrt{\frac{n^{-4/5}}{c}f(x)\kappa_{02}}} - \frac{\frac{1}{2}c^2 n^{-2/5} f''(x)\kappa_{21}}{\sqrt{\frac{n^{-4/5}}{c}f(x)\kappa_{02}}} = n^{2/5}\frac{\hat{f}_h(x) - f(x)}{\sqrt{\frac{1}{c}f(x)\kappa_{02}}} - \frac{\frac{1}{2}c^2 f''(x)\kappa_{21}}{\sqrt{\frac{1}{c}f(x)\kappa_{02}}}$$

this leads to

$$n^{2/5}\left(\hat{f}_h(x) - f(x)\right) \rightsquigarrow N\left(\frac{1}{2}c^2 f''(x)\kappa_{21}, \frac{1}{c}f(x)\kappa_{02}\right)$$

$\blacksquare$

Problem 3 *Generate 1000 samples from distribution* $0.3N(0,1)+0.7N(1,0.32)$, *draw one figure with different estimated density functions with different bandwidth* h *where* h *is given by* kedd.

Problem 4 (Slide 7 page 44) *Suppose now that we have estimated an unknown density* f *using some kernel* K_A *and bandwidth* h_A, *what bandwidth* h_B *should we use in the estimation with kernel* K_B *when we want to get approximately the same degree of smoothness as we had in the case of* K_A *and* h_A?

Prove that the answer is given by the following formula:

$$h_B = h_A \frac{\delta_0^B}{\delta_0^A}.$$

Solution

Assume that we choose h to minimize the AMSE,

$$\text{AMSE} = \frac{1}{4}h^4(f''(x))^2\kappa_{21}^2 + \frac{1}{nh}f(x)\kappa_{02},$$

so this leads to

$$h(x) = \left(\frac{f(x)\kappa_{02}}{(f''(x))^2\,\kappa_{21}^2}\right)^{1/5} n^{-1/5} = \delta_0\left(\frac{f(x)}{(f''(x))^2}\right)^{1/5} n^{-1/5}$$

Thus, to get the same degree of smoothness as we had in the case of K_A and h_A, we should choose

$$h_B = h_A \frac{\delta_0^B}{\delta_0^A}.$$

When we choose h who doesnot minimize the AMSE, the relation between h and δ_0 would hold approximately. ■

Homework 8

Problem 1 *Let* $X_1, \ldots, X_n$, *i.i.d from density* $f(x)$, *denote* r *derivative of density function by* $\hat{f}_h^{(r)}$, *please find the asymptotic distribution of* $\hat{f}_h^{(r)}$.

Solution

We have

$$E\left[f_h^{(r)}(x)\right] = \frac{1}{h}\int K\left(\frac{x-z}{h}\right) f^{(r)}(z)dz$$
$$= f^{(r)}(x) + \frac{1}{2} f^{(r+2)}(x)\kappa_{21}h^2 + o\left(h^2\right),$$

and

$$\operatorname{Var}\left[f_h^{(r)}(x)\right] = \frac{f(x)}{nh^{2r+1}}\int\left[K^{(r)}(u)\right]^2 du + o\left(\frac{1}{nh^{2r+1}}\right).$$

Choose $h(x) = cn^{-1/(2r+5)}$. Thus,

$$\frac{f_h^{(r)}(x) - E\left[f_h^{(r)}(x)\right]}{\sqrt{\operatorname{Var}\left[f_h^{(r)}(x)\right]}} \rightsquigarrow N(0,1),$$

Because

$$\begin{aligned}
&\frac{f_h^{(r)}(x) - f^{(r)}(x) - \frac{1}{2}f^{(r+2)}(x)\kappa_{21}h^2}{\sqrt{\frac{f(x)}{nh^{2r+1}}\int\left[K^{(r)}(u)\right]^2 du}} \\
=&\frac{f_h^{(r)}(x) - f^{(r)}(x) - \frac{1}{2}f^{(r+2)}(x)\kappa_{21}c^2 n^{-2/(2r+5)}}{\sqrt{n^{-4/(2r+5)}c^{-2r-1}f(x)\int\left[K^{(r)}(u)\right]^2 du}} \\
=&n^{2/(2r+5)}\frac{f_h^{(r)}(x) - f^{(r)}(x)}{\sqrt{c^{-2r-1}f(x)\int\left[K^{(r)}(u)\right]^2 du}} - \frac{\frac{1}{2}c^2 f^{(r+2)}(x)\kappa_{21}}{\sqrt{c^{-2r-1}f(x)\int\left[K^{(r)}(u)\right]^2 du}}
\end{aligned}$$

so we obtain

$$n^{2/(2r+5)}\left(f_h^{(r)}(x) - f^{(r)}(x)\right) \rightsquigarrow N\left(\frac{1}{2}c^2 f^{(r+2)}(x)\kappa_{21}, c^{-2r-1}f(x)\int\left[K^{(r)}(u)\right]^2 du\right).$$

■

Problem 2 *Generate 1000 samples from distribution* $0.3N(0,1) + 0.7N(1, 0.32)$*, draw one figure with different estimated derivative of density functions with different bandwidth* h *where* h *is given by* kedd*.*

Problem 3 (Adaptive KDE) *Write program to draw the figure of 21 pages of courseware. (Referring algorithm steps in reading materials)*

Problem 4 (Kernel Density Estimation with Boundary Correction) *Reade accessory materials. Write program to realize the result in 39 pages of slide. (Reading material, Figure 3)*