Oulun yliopisto
Matemaattisten tieteiden laitos
Funktioiden estimointi
9. harjoitus, viikko 12, 2014

1. Tarkastellaan tiheysfunktiota

$$f(x) = \frac{3}{4} f(x; -\frac{3}{2}, 1) + \frac{1}{4} f(x; \frac{3}{2}, (\frac{1}{3})^2),$$

jossa $f(\cdot; \mu, \sigma^2)$ on normaalijakauman $N(\mu, \sigma^2)$ tiheysfunktio.

Generoi kokoa $n = 100$ oleva otos tiheysfunktiosta f, ja laske tiheysfunktion Breiman-Meisel-Purcel estimaattori. (Valitaan esim. lähinaapurien lukumäärä $k = 10$ ja $h = 1$.)

2. Tässä tehtävässä estimoimme tiheysfunktiota f, jolla tiedämme olevan reunan pisteessä 0.

(a) Olkoon $\tilde{f}_n$ reunaheijastusmenetelmällä saatu estimaatti silotusparametrilla h. Osoita, että $\tilde{f}_n'(0+) = 0$.

(b) Generoi kokoa $n = 100$ oleva otos eksponenttijakaumasta $\mathrm{Exp}(1)$ ja laske reunaheijastusmenetelmällä saatava estimaatti sekä tavallinen ydinestimaatti samalla silotusparametrilla h. (Vihje, jos U on tasajakautunut välillä $[0, 1]$, ja määritellään $X = -\log(U)$, niin $X \sim \mathrm{Exp}(1)$.)

3. Olkoon d-ulotteisella satunnaisvektorilla X tasainen jakauma yksikkökuutiossa $B = [0, 1]^d$.

(a) Tarkastellaan kuutiota

$$B_h(x_0) = [x_{01} - h, x_{01} + h] \times \cdots \times [x_{0d} - h, x_{0d} + h], \; x_0 = [x_{01}, \ldots, x_{0d}]^T$$

joka sisältyy aidosti yksikkökuutioon. Todista, että

$$\mathbb{P}(X \in B_h(x_0)) \to 0, \qquad \text{kun } d \to \infty.$$

(b) Olkoon $\varepsilon > 0$ ja olkoon $A(\varepsilon)$ niiden pisteiden joukko, joiden etäisyys yksikkökuution B reunasta on pienempi kuin ε. Todista (edelliseen kohtaan nojautuen), että

$$\mathbb{P}(X \in A(\varepsilon)) \to 1, \qquad \text{kun } d \to \infty.$$

4. Kun d-ulotteisessa avaruudessa on ydinestimaattorin silotusparametrimatriisille H_n valittu muoto $H_n = h_n^2 I_d$, (jossa $h_n > 0$), niin osoita, että luentojen s. 112 kaava

$$\mathrm{AMISE}[\hat{f}_n(\cdot; H_n)] = \frac{1}{4}\mu_2(K)^2 R(\mathrm{tr}[H_n G_f]) + \frac{R(K)}{n|H_n^{1/2}|}$$

sievenee s. 111 annettuun muotoon

$$\mathrm{AMISE}[\hat{f}_n(\cdot; H_n)] = \frac{1}{4} h_n^4 \mu_2(K)^2 R(\nabla^2 f) + \frac{R(K)}{n h_n^d}.$$

5. Olkoon f normaalijakauman $N(\mu, \sigma^2 I_d)$ ja φ normaalijakauman $N(0, I_d)$ tiheysfunktio.

(a) Osoita (muuttujanvaihdolla), että $R(\nabla^2 f) = \sigma^{-d-4} R(\nabla^2 \varphi)$.

(b) Osoita, että $\nabla^2 \varphi(x) = (x^T x - d)\varphi(x)$.

(c) Osoita, että $R(\nabla^2 \varphi) = 2^{-d} \pi^{-d/2} d(d+2)/4$.

(d) Kun ytimenä käytetään Gaussin ydintä (ts. $K = \varphi$) ja estimoidaan tiheysfunktiota f, niin osoita, että AMISE-optimaalisen silotusparametrin kaava

$$h_n^* = \left[\frac{d R(K)}{\mu_2(K)^2 R(\nabla^2 f)} \right]^{1/(d+4)} n^{-1/(d+4)}$$

sievenee muotoon

$$h_n^* = \left[\frac{4}{d+2} \right]^{1/(d+4)} \sigma n^{-1/(d+4)}.$$