Clase 17

jueves, 25 de octubre de 2018 9:06

Veamos ahora las propiedades del estimador basado en núcleos en el caso general

Teorema 88: Si $f_n(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)$ entonces

$f_n(x)$ es asintóticamente insesgado en x punto de continuidad de f si se cumple:

a) $\int K(t)\,dt = 1$

b) $|x K(x)| \longrightarrow 0$ si $|x| \to \infty$

c) K acotada y simétrica

d) $h \to 0$

Dem: $E(f_n(x)) = \frac{1}{h} E\left(K\left(\frac{x-X_1}{h}\right)\right) = \frac{1}{h}\int K\left(\frac{x-t}{h}\right) f(t)\,dt$

Dado que x es un punto de continuidad de f y se cumplen las condiciones del lema de Bochner se tiene que

$$\frac{1}{h}\int K\left(\frac{x-t}{h}\right) f(t)\,dt \longrightarrow f(x) \quad \text{si } h \to 0$$

entonces $E(f_n(x)) \to f(x)$ o sea el estimador es asintóticamente insesgado en x de continuidad de f

Teorema 89: Consideremos el estimador $f_n(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)$

Supongamos que se cumplen las condiciones del teorema 88 y además que $nh \to \infty$.

Entonces $\hat{f}_n(x)$ es un estimador consistente de $f(x)$

Dem: Calculemos la varianza de $\hat{f}_n(x)$

$$\operatorname{Var}(\hat{f}_n(x)) = \frac{1}{nh^2} \operatorname{Var}\left(K\left(\frac{x-X_1}{h}\right)\right) \leq \frac{1}{nh^2} E\left(K^2\left(\frac{x-X_1}{h}\right)\right) =$$

$$= \frac{1}{nh} \frac{1}{h} \int K^2\left(\frac{x-t}{h}\right) f(t)\,dt$$

Consideremos la función $K^2(x)$

Como $|x K(x)| \to 0$ si $|x| \to \infty$ entonces $|x^2 K^2(x)| \to 0$ si $|x| \to \infty$

Además $|x K^2(x)| \leq x^2 K^2(x)$ si $|x| > 1$

Luego $|x K^2(x)| \to 0$ si $|x| \to \infty$

Por otro lado $\int |K^2(t)|\,dt \leq \|K\|_\infty \int K(t)\,dt < \infty$

pues $\int K(t)\,dt < \infty$ y K es acotada.

Claramente $K^2(x)$ es simétrica ya que K lo es.

Por lo tanto podemos aplicar el lema de Bochner obteniendo que:

$$\frac{1}{h} \int K^2\left(\frac{x-t}{h}\right) f(t)\,dt \to f(x) \int K^2(t)\,dt$$

$$\frac{1}{h}\int K^2\left(\frac{x-t}{h}\right) f(t)\,dt \to f(x)\int K^2(t)\,dt$$

Entonces $\mathrm{Var}(f_n(x)) \to \dfrac{f(x)}{nh}\displaystyle\int K^2(t)\,dt \to 0$ si $nh\to\infty$

Por el teorema 88 $E(f_n(x)) \xrightarrow[h\to 0]{} f(x)$ y

$$\mathrm{Var}(f_n(x)) \xrightarrow[nh\to\infty]{} 0$$

A partir de esto $f_n(x) \xrightarrow{P} f(x)$ o sea

$f_n(x)$ es un estimador consistente de $f(x)$

Obs Lo que probamos es la convergencia en L^2 es decir

$$E\left[(f_n(x)-f(x))^2\right] \to 0$$

Definición 90: Decimos que $f_n \to f$ uniformemente casi seguramente si

$$\lim_{n\to\infty} \|f_n(x)-f(x)\|_\infty = 0 \quad \text{c.s}$$

o sea $P\left(\omega\in\Omega \,/\, \lim\limits_{n\to\infty}\|f_n(x)-f(x)\|_\infty \neq 0\right) = 0$

Definición 91: $Z_n \to Z$ completamente ($Z_n \xrightarrow{c} Z$) si y solo si

$$\sum_{n=1}^{\infty} P(|Z_n - Z| > \varepsilon) < \infty$$

Teorema 92: El estimador f_n converge uniformemente casi seguramente si

a) K es continua por la derecha

b) K es de variación acotada

c) $\lim_{|x|\to\infty} K(x) = 0$

d) f es uniformemente continua

e) $\sum_{n=1}^{\infty} e^{-\gamma n h^2} < \infty$ para todo $\gamma > 0$

Dem: Se omite

Obs: Hay resultados para la convergencia casi segura que piden $\frac{nh}{\log\log n} \to \infty$ y para la convergencia completa que piden $\frac{nh}{\log n} \to \infty$

Kernels ($K(u)$)

1) Uniforme $\quad K(u) = \frac{1}{2} I(|u| \leq 1)$

2) Triangular $\quad K(u) = (1-|u|) I(|u| \leq 1)$

3) Epanechnikov $\quad K(u) = \frac{3}{4}(1-u^2) I(|u| \leq 1)$

4) Quartic $\quad K(u) = \frac{15}{16}(1-u^2) I(|u| \leq 1)$

5) Triweight $\quad K(u) = \frac{35}{32}(1-u^2)^3 I(|u| \leq 1)$

6) Gaussiano $\quad K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$

7) Coseno $\qquad K(u) = \frac{\pi}{4} \cos\left(\frac{\pi}{2} u\right) I(|u| \leq 1)$

Ejercicio Probar que $\int f_n(t)\,dt = 1$

Cálculo de la Varianza y el Sesgo

$$Var(\hat{f}_n(x)) = \frac{1}{nh^2} Var\left(K\left(\frac{x-X_1}{h}\right)\right) = \frac{1}{nh^2}\left[E\left(K^2\left(\frac{x-X_1}{h}\right)\right) - E^2\left(K\left(\frac{x-X_1}{h}\right)\right)\right]$$

$$= \frac{1}{nh^2}\left[\int K^2\left(\frac{x-t}{h}\right) f(t)\,dt - \left(\int K\left(\frac{x-t}{h}\right) f(t)\,dt\right)^2\right]$$

Como $\frac{1}{h}\int K\left(\frac{x-t}{h}\right) f(t)\,dt \to f(x)$ si $h \to 0$

entonces

$$\frac{1}{h}\int K\left(\frac{x-t}{h}\right) f(t)\,dt = f(x) + O(h)$$

Vimos también que

$$\frac{1}{h}\int K^2\left(\frac{x-t}{h}\right) f(t)\,dt \longrightarrow f(x) \underbrace{\int K^2(t)\,dt}_{\|K\|_2^2}$$

o sea $\quad \frac{1}{h}\int K^2\left(\frac{x-t}{h}\right) f(t)\,dt = \|K\|_2^2\left(f(x)+o(h)\right)$

Entonces

$$Var\left(f_n(x)\right) = \frac{1}{n}\left[\frac{1}{h}\|K\|_2^2\left(f(x)+o(h)\right) - \left(f(x)+o(h)\right)^2\right]$$

$$= \frac{1}{nh}\|K\|_2^2 f(x) + \frac{1}{nh}\|K\|_2^2\, o(h) - \frac{\left(f(x)+o(h)\right)^2}{n}$$

Pero $\quad \frac{1}{nh}\|K\|_2^2\, o(h) = o(1/nh)$ pues $\quad \frac{\|K\|_2^2\, o(h)/nh}{1/nh} \to 0$

y $\quad \frac{\left(f(x)+o(h)\right)^2}{n} = o(1/nh)$ pues $\quad \frac{\left(f(x)+o(h)\right)^2/n}{1/nh} \to 0$

luego

$$Var\left(f_n(x)\right) = \frac{1}{nh}\|K\|_2^2 f(x) + o(1/nh) \quad , \; nh \to \infty$$

o sea el orden de $Var(f_n(x))$ es $\frac{1}{nh}$

Analicemos el sesgo

$$B(f_n(x)) = E\left(f_n(x)\right) - f(x) = \frac{1}{h}\int K\left(\frac{x-t}{h}\right) f(t)\,dt - f(x) =$$

$$= \int K(u)\, f(x+uh)\,du - f(x)\int K(u)\,du =$$

$$= \int K(u) \left[f(x-uh) - f(x) \right] du =$$

Si suponemos adicionalmente que

- $\exists \, f''$ y $\sup |f''| < M$
- $\int K(u)\, u \, du = 0$ (Si K es simétrica se cumple)
- $\int u^2 K(u) \, du = \mu_2(K) < \infty$

$$= \int K(u) \left[f'(x) h u + f''(\xi) \frac{h^2 u^2}{2} \right] du \doteq \frac{h^2}{2} \int u^2 K^2(u) f''(\xi) \, du$$

Como $\xi \in (x, x+uh)$ entonces $f''(\xi) = f''(x) + o(h)$

$$\doteq \frac{h^2}{2} \mu_2(K) f(x) + o(h^2) \quad , \quad h \to 0$$

pues $\frac{h^2}{2} o(h) \mu_2(K) = o(h^2) \quad h \to 0$

Teorema 93: Sea $f_n(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$

Si x es un punto de continuidad de f y se cumple

a) $\int K(t)\, dt = 1$

b) K es simétrico y acotado

c) $|x K(x)| \longrightarrow 0$ si $|x| \to \infty$

d) $\int u^2 K^2(u)\,du = \mu_2(K) < \infty$

e) f es dos veces diferenciable y $\sup|f''(x)| < M$

Entonces

$$ECM(f_n(x)) = \frac{1}{nh} f(x) \|K\|_2^2 + \frac{h^4}{4}\left(f''(x)\mu_2(K)\right)^2 + o\left(\frac{1}{nh}\right) + o(h^4)$$

si $h \to 0$ y $nh \to \infty$

Dem: Es inmediato a partir de que

$$ECM(f_n(x)) = Var(f_n(x)) + B^2(f_n(x))$$

Teorema 94: Si se cumplen las hipótesis del teorema 93 la ventana que minimiza el error cuadrático medio es:

$$h_0 = \left(\frac{f(x)\|K\|_2^2}{\left(f''(x)\,\mu_2(K)\right)^2 n}\right)^{1/5}$$

Dem: $h_0 = \underset{h}{\arg\min}\, ECM(f_n(x))$

$$\frac{\partial ECM(f_n(x))}{\partial h} = -\frac{f(x)\|K\|_2^2}{\ldots} + h^3\left(f''(x)\mu_2(K)\right)^2$$

$$\frac{\partial ECM(f_h(x))}{\partial h} = -\frac{f(x)\,\|K\|_2^2}{n h^2} + h^3\left(f''(x)\,\mu_2(K)\right)^2 = 0$$

o sea $$h_0 = \left[\frac{f(x)\,\|K\|_2^2}{\left(f''(x)\,\mu_2(K)\right)^2 n}\right]^{1/5}$$

Con este h_0 el error cuadrático medio es

$$ECM\left(f_{n,h_0}(x)\right) = \frac{f(x)\,\|K\|_2^2}{n\,h_0} + \frac{h_0^4}{4}\left(f''(x)\,\mu_2(K)\right)^2 =$$

$$= \frac{\left(f(x)\,\|K\|_2^2\right)^{4/5}\left(f''(x)\,\mu_2(K)\right)^{2/5}}{n^{4/5}} + \frac{\left(f(x)\,\|K\|_2^2\right)^{4/5}}{4\,n^{4/5}}\left(f''(x)\,\mu_2(K)\right)^{2/5} =$$

$$= \frac{5}{4}\left(f(x)\,\|K\|_2^2\right)^{4/5}\left(f''(x)\,\mu_2(K)\right)^{2/5}\,n^{-4/5}$$

o sea el orden del ECM es $n^{-4/5}$