2.3. Propiedades Estadísticas

Pregunta 2.6

¿Podríamos imitar lo mismo que hicimos para el histograma?

Si. Las propiedades que vimos anteriormente son universales para estimadores.

Entonces:

$$\mathrm{MSE}(\hat{f}_h(x)) = \mathrm{Var}(\hat{f}_h(x)) + \mathrm{Sesgo}^2(\hat{f}_h(x))$$
$$\mathrm{MISE}(\hat{f}_h) = \int \mathrm{Var}(\hat{f}_h(x))dx + \int \mathrm{Sesgo}^2(\hat{f}_h(x))dx$$

donde

$\mathrm{Var}\left(\hat{f}_h(x)\right) = \mathbb{E}\left[\hat{f}_h(x) - \mathbb{E}\hat{f}_h(x)\right]^2$ and $\mathrm{Sesgo}\left(\hat{f}_h(x)\right) = \mathbb{E}\left[\hat{f}_h(x)\right] - f(x)$.

2.3.1. Varianza

X_i iid

?

$$
\begin{aligned}
\operatorname{Var}(\hat{f}_h(x)) &= \operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)\right) \\
&= \frac{1}{n^2h^2}\sum_{i=1}^{n}\operatorname{Var}\left(K\left(\frac{x-X_i}{h}\right)\right) \\
&= \frac{1}{nh^2}\operatorname{Var}\left(K\left(\frac{x-X}{h}\right)\right) \\
&= \frac{1}{nh^2}\left\{\mathbb{E}\left[K^2\left(\frac{x-X}{h}\right)\right] - \left\{\mathbb{E}\left[K\left(\frac{x-X}{h}\right)\right]\right\}^2\right\}.
\end{aligned}
$$

$$\hat{f}_h(x) := \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)$$

Usando que

$$
\begin{aligned}
\mathbb{E}\left[K^2\left(\frac{x-X}{h}\right)\right] &= \int K^2\left(\frac{x-s}{h}\right) f(s)ds \\
&= h\int K^2(u)\, f(uh+x)du \\
&= h\int K^2(u)\,\{f(x)+o(1)\}\, du \\
&= h\left\{\|K\|_2^2 f(x) + o(1)\right\}.
\end{aligned}
$$

• $u = s - x$

→ Rsp to h

• $K(u) = K(-u)$

TAYLOR.

$\|K\|_2^2 = \int K^2(u)\,du$

$\int K(u)du = 1$

$$
\begin{aligned}
\mathbb{E}\left[K\left(\frac{x-X}{h}\right)\right] &= \int K\left(\frac{x-s}{h}\right) f(s)ds \\
&= h\int K(u)\, f(uh+x)du \\
&= h\int \underline{K(u)\,\{f(x)+o(1)}\}\, du \\
&= h\,\{f(x)+o(1)\}\,.
\end{aligned}
$$

Por lo tanto se obtiene que

$$
\operatorname{Var}\left(\hat{f}_h(x)\right) = \boxed{\frac{1}{nh}\|K\|_2^2 f(x) + o\left(\frac{1}{nh}\right),} \text{ si } nh \to \infty.
$$

2.3.2. Sesgo

Para el sesgo tenemos

$$
\begin{aligned}
\operatorname{Sesgo}\left(\hat{f}_h(x)\right) &= \mathbb{E}\left[\hat{f}_h(x)\right] - f(x) \\
&= \frac{1}{nh}\sum_{i=1}^{n} \mathrm{E}\left[K\left(\frac{x - X_i}{h}\right)\right] - f(x) \\
&= \frac{1}{h}\mathrm{E}\left[K\left(\frac{x - X_1}{h}\right)\right] - f(x) \\
&= \int \frac{1}{h}K\left(\frac{x-u}{h}\right) f(u)du - f(x)
\end{aligned}
$$

$f(sh+x)$

Tarea 2.7

Usando el cambio de variable $s = \frac{u-x}{h}$ y las propiedades del kernel pruebe que

TAYLOR ORDEN 2

$$
\operatorname{Sesgo}\left(\hat{f}_h(x)\right) = \frac{h^2}{2} f'' \mu_2(K) + o(h^2), \text{ si } h \to 0
$$

$f''(x)$

donde $\mu_2 = \int s^2 K(s)ds$.

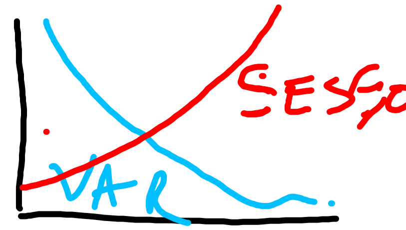
Nota: *En algunas pruebas más formales, se necesita además que f'' sea absolutamente continua y que $\int (f'''(x))dx < \infty$.*

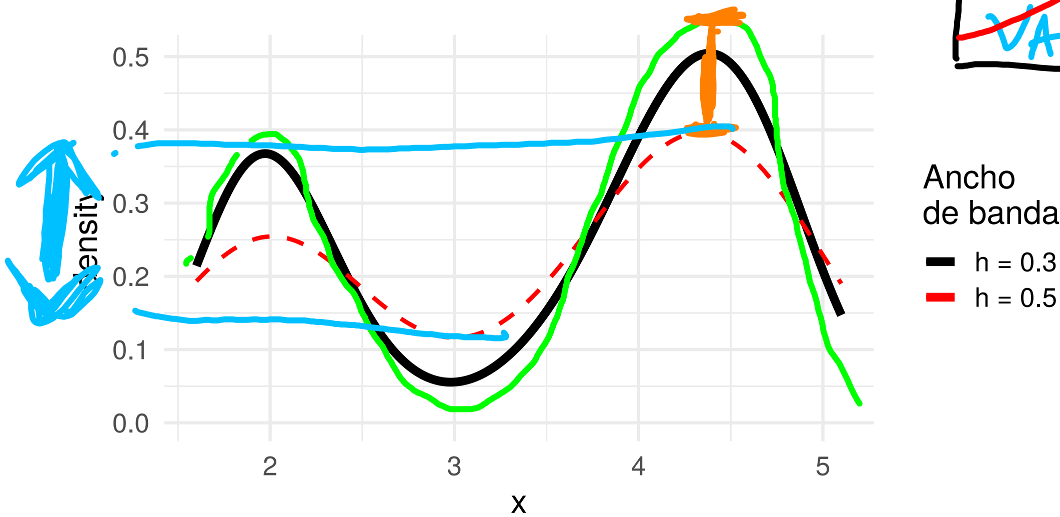
Mundo Real

Sesgo

Mundo Modelo

VAR (Incertidumbre)





Nota 2.8

Note como los cambios en el ancho de banda modifican la suavidad (sesgo) y el aplanamiento de la curva (varianza).

2.3.3. Error cuadrático medio y Error cuadrático medio integrado

El error cuadrático medio se escribe

$$\begin{aligned}\mathrm{MSE}(\hat{f}_h(x)) &= \mathrm{Sesgo}\left(\hat{f}_h(x)\right)^2 + \mathrm{Var}\left(\hat{f}_h(x)\right)\\ &= \frac{h^4}{4}\left(\mu_2(K)\underline{f''(x)}\right)^2 + \frac{1}{nh}\|K\|_2^2\underline{f(x)} + o(h^4) + o\left(\frac{1}{nh}\right).\end{aligned}$$

Y el error cuadrático medio integrado se escribe como,

$$\begin{aligned}\mathrm{MISE}\left(\hat{f}_h\right) &= \int \mathrm{MSE}\left(\hat{f}_h(x)\right)dx\\ &= \int \mathrm{Sesgo}\left(\hat{f}_h(x)\right)^2 + \mathrm{Var}\left(\hat{f}_h(x)\right)dx\\ &= \frac{h^4}{4}\mu_2^2(K)\left\|f''(x)\right\|_2^2 + \frac{1}{nh}\|K\|_2^2 + o(h^4) + o\left(\frac{1}{nh}\right).\end{aligned}$$

$$\frac{\partial \mathrm{MISE}}{\partial h} \ldots$$

2.3.4. Ancho de banda óptimo

Minimizando el MISE con respecto a h obtenemos

$C \cdot n^{-1/5}$

$$h_{opt} = \left(\frac{\|K\|_2^2}{\|f''\|_2^2 \left(\mu_2(K)\right)^2 n} \right)^{1/5} = O\left(n^{-1/5}\right).$$

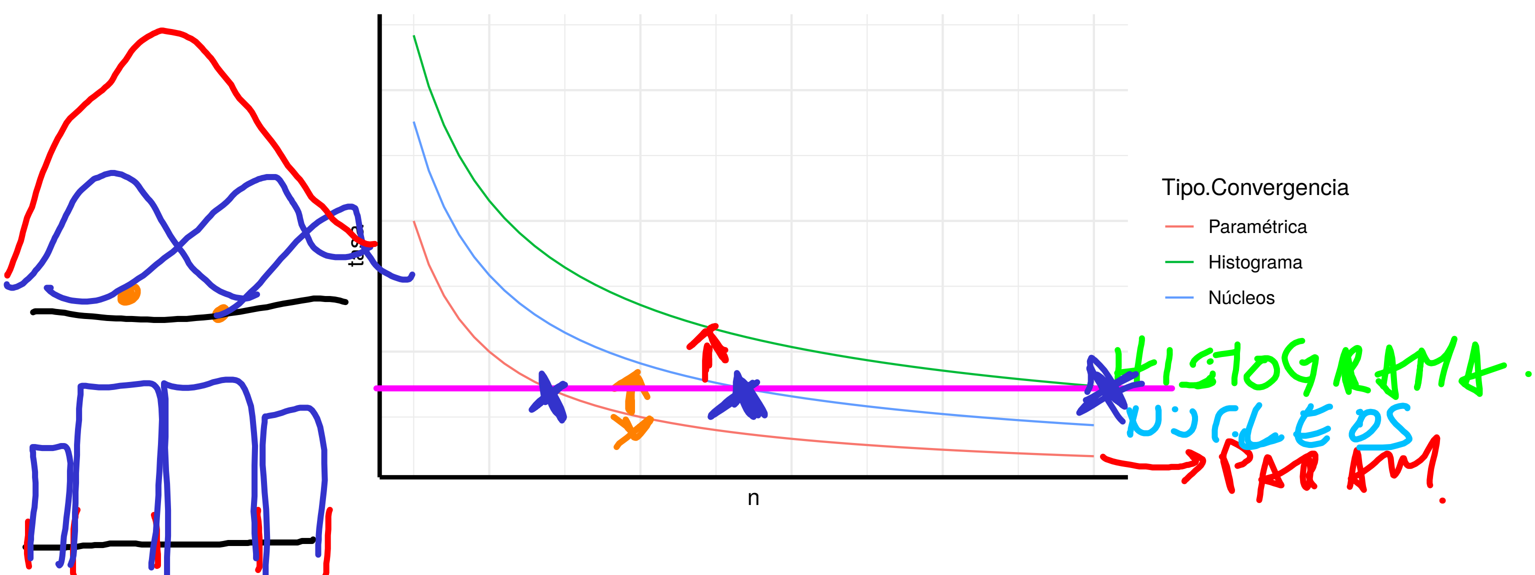
Nota 2.9

De forma práctica, h_{opt} no es un estimador útil de h porque depende de $\|f''\|_2^2$ que es desconocido.

Más adelante veremos otra forma de encontrar este estimador.

Evaluando h_{opt} en el MISE tenemos que

$$\mathrm{MISE}(\hat{f}_h) = \frac{5}{4}\left(\|K\|_2^2\right)^{4/5}\left(\|f''\|_2^2 \mu_2(K)\right)^{2/5} n^{-4/5} = O\left(n^{-4/5}\right).$$



Nota 2.10: Detalle técnico

Formalmente, es posible probar que si f es β veces continuamente diferenciable y $\int \left(f^{(\beta)}\right)^2 < \infty$, entonces se tiene que

$$\boxed{h_{opt} = O\left(n^{-\frac{1}{2\beta+1}}\right).}$$

Por lo tanto se podría aproximar a una tasa paramétrica de convergencia si β es grande.