ERROR CUADRATICO MEDIO

EL ERROR CUADRATICO MEDIO DE $\hat{f}_h(x)$ ES

$$MSE(\hat{f}_h(x)) = \underline{\mathbb{E}[\{\hat{f}_h(x) - f(x)\}^2]}$$

SE PUEDE DESCOMPONER EL MSE COMO

$$MSE\{\hat{f}_h(x)\} = \underline{Var\{\hat{f}_h(x)\}} + \underline{\{Bias\{\hat{f}_h(x)\}\}^2}.$$

TAREA

USANDO LO QUE HICIMOS ANTES SE TIENE QUE

$$MSE\{\hat{f}_h(x)\} = \underbrace{\boxed{\frac{1}{nh} f(x)}}_{Var(\hat{f}_n(x)) \to 0} + \underbrace{\boxed{f'\{(j-\tfrac{1}{2})h\}^2 \underbrace{\{(j-\tfrac{1}{2})h - x\}^2}}}_{Bias(\hat{f}_n(x))^2 \to 0}$$

$$\underbrace{+\, o(h^2) + o\left(\frac{1}{nh}\right)}$$

OJO

- $\boxed{\text{SI } h \to 0 \,,\; nh \to +\infty}$

$n \to +\infty$
$nh \to +\infty$

ENTONCES $MSE\{\hat{f}_h(x)\} \to 0$

USAMOS MAS OBSERVACIONES $(n \to +\infty)$ Y EL ANCHO DE BANDA MAS PEQUEÑO $(h \to +\infty)$

⚠ SI $MSE\{\hat{f}_h(x)\} \to 0$ (CONVERGENCIA EN L^2)

$\Rightarrow \hat{f}_h(x) \xrightarrow{P} f(x)$ $\therefore$ $\hat{f}_h(x)$ ES CONSISTENTE.

$$\mathrm{MSE}\{\hat{f}_n(x)\} = \underbrace{\boxed{\cancel{\frac{1}{nh} f(x)}}}_{}^{\mathrm{Var}(f_n(x)) \to 0} + \overbrace{\boxed{\cancel{f'\left\{\left(j-\tfrac{1}{2}\right)h\right\}^2} \left\{\left(j-\tfrac{1}{2}\right)h - x\right\}^2}}^{\mathrm{Bias}(\hat{f}_n(x)) \to +\infty}$$

$$+\, o(h^2) \quad + o\left(\frac{1}{nh}\right)$$

$h \to +\infty$

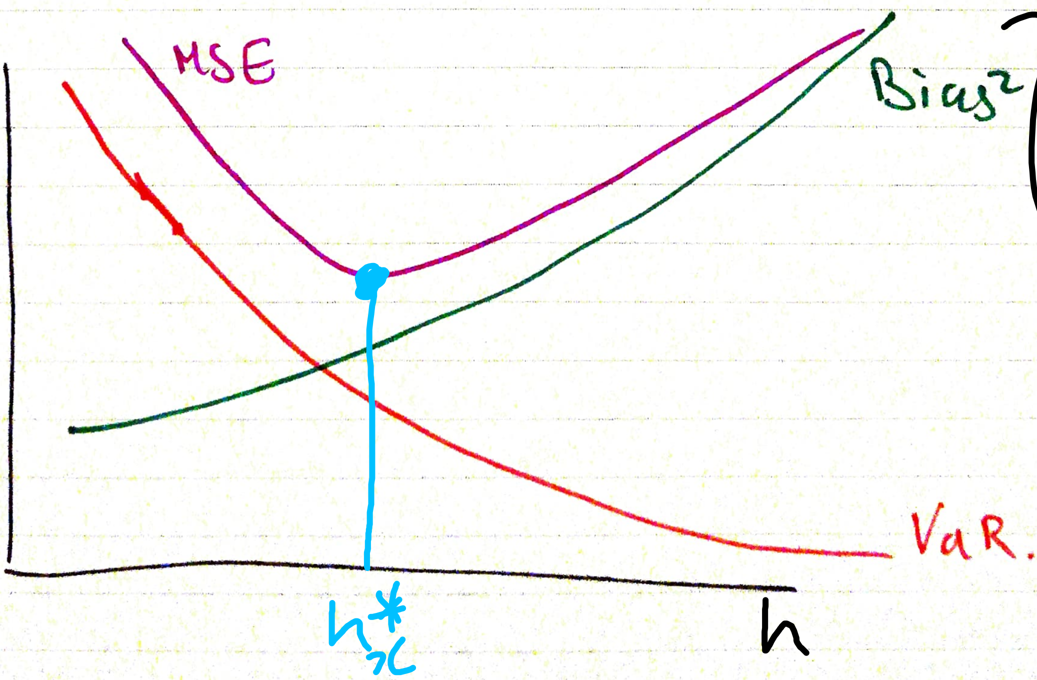
OJO Suponga n fijo. y $\boxed{x \text{ fijo}}$

- SI $h \to 0 \Rightarrow \mathrm{Var}(\hat{f}_n(x)) \to +\infty$

  $\mathrm{Bias}(\hat{f}_n(x))^2 \to f'(0)\, x^2$

- SI $h \to +\infty \Rightarrow \mathrm{Var}(\hat{f}_n(x)) \to 0$

  $\mathrm{Bias}(\hat{f}_n(x))^2 \to +\infty$



Para cada x

$\mathrm{MSE} = \mathrm{Var} + \mathrm{Bias}^2$

ERROR CUADRÁTICO MEDIO INTEGRADO.

EL PROBLEMA CON EL $MSE(\hat{f}_h(x))$ ES QUE DEPENDE DE UN SOLO PUNTO "x".

SOLUCIÓN

$$MISE(\hat{f}_h) = \mathbb{E}\left[\int_{-\infty}^{+\infty} \{\hat{f}_h(x) - f(x)\}^2 dx\right]$$

$$= \int_{-\infty}^{+\infty} \mathbb{E}[\{\hat{f}_h(x) - f(x)\}^2]\, dx$$

$$= \int_{-\infty}^{+\infty} MSE(\hat{f}_h(x))\, dx.$$

ADEMÁS:

$$MISE(\hat{f}_h) = \int \frac{1}{nh} f(x)\, dx +$$

$$+ \int \sum_j I(x \in B_j) \left\{\left(j-\tfrac{1}{2}\right)h - x\right\}^2 \left[f'\left(\left(j-\tfrac{1}{2}\right)h\right)\right]^2 dx.$$

$$= \frac{1}{nh} + \sum_j \underbrace{\int_{B_j} \left\{x - \left(j-\tfrac{1}{2}\right)h\right\}^2}_{\frac{h^2}{12}} \underbrace{\left\{f'\left(\left(j-\tfrac{1}{2}\right)h\right)\right\}^2}_{\text{NO DEPENDE DE "}x\text{"}} dx.$$

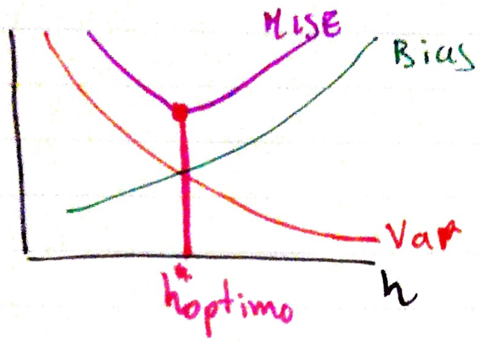
$$= \frac{1}{nh} + \frac{h^2}{12} \underbrace{\sum_j f'\left(\left(j-\tfrac{1}{2}\right)h\right)^2 \cdot h.}_{\text{SUMA DE RIEMMAN}}$$

No depende de x

$$\approx \frac{1}{nh} + \frac{h^2}{12} \int_{\mathbb{R}} \{f'(x)\}^2 dx = \boxed{\frac{1}{nh} + \frac{h^2}{12} \underbrace{\|f'\|_2^2}_{\text{DESCONOCIDO}}}$$

ANCHO DE BANDA ÓPTIMO PARA EL HISTOGRAMA.

EL MISE TIENE EL MISMO COMPORTAMIENTO QUE EL MSE



OJO: LO IMPORTANTE ES MINIMIZAR EL MISE.

$$\frac{\partial \, MISE(f_n)}{\partial h} = \frac{-1}{nh^2} + \frac{1}{6} h \|f'\|_2^2 = 0$$

$$\Rightarrow \boxed{h^*_{opt} = \left(\frac{6}{n\|f'\|_2^2}\right)^{1/3} \sim n^{-1/3}}$$

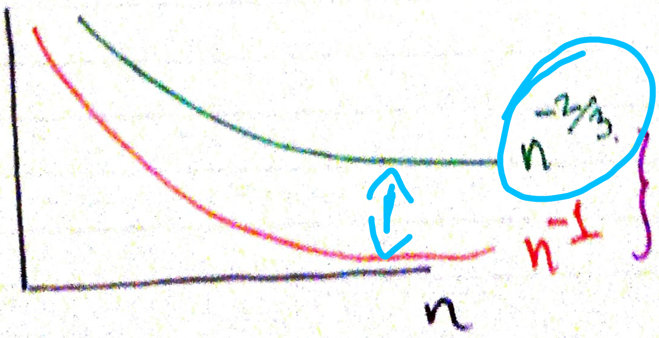
SI EVALUAMOS h_{opt} EN EL MISE SE OBTIENE.

$$MISE(f_{h_{opt}}) = C\, n^{-2/3},$$ CON $\underline{C}$ UNA CONSTANTE QUE DEPENDE DE f' → DESCONOCIDA

RECUERDE DE ESTADÍSTICA I
SI $X_1, \dots, X_n \sim N(\mu, \sigma^2)$, SE QUIERE ESTIMAR μ.

$$\hat{\mu} = \frac{1}{n}\sum X_i$$

$$MSE(\hat{\mu}) = Bias^2(\hat{\mu}) + Var(\hat{\mu}) = 0 + \frac{1}{n} Var(X) = Cn^{-1}$$



EL CAMBIO DE VELOCIDAD ES POR USAR UN HISTOGRAMA EN LUGAR DE UNA DENS. PARAMÉTRICA.