Solution du Concours de Statistique Non Paramétrique

Solution du Problème

Commençons par calculer le biais : en faisant le changement de variable suivant

$$y = x + uh, \quad dy = hdu$$

$$\begin{aligned}
\mathbb{E}\left(f_n(x)\right) &= \frac{1}{nh}\sum_{i=1}^{n}\mathbb{E}\left[K\left(\frac{X_i - x}{h}\right)\right] \\
&= \frac{1}{nh}\sum_{i=1}^{n}\int_{\mathbb{R}} K\left(\frac{y-x}{h}\right) f(y)dy \\
&= \frac{1}{h}\int_{\mathbb{R}} K\left(\frac{y-x}{h}\right) f(y)dy \\
&= \int_{\mathbb{R}} K(u)f(x+uh)du.
\end{aligned}$$

En effectuant un dévelopement limité à l'ordre 2, avec $\zeta_u \in [x, x+uh]$, il vient

$$\begin{aligned}
\mathbb{E}\left(f_n(x)\right) &= \int_{\mathbb{R}} K(u)f(x+uh)du \\
&= \int_{\mathbb{R}} K(u)[f(x) + (uh)f'(x) + \frac{(uh)^2}{2}f"(\zeta_u)]du \\
&= f(x)\underbrace{\int_{\mathbb{R}} K(u)du}_{=1} + hf'(x)\underbrace{\int_{\mathbb{R}} uK(u)du}_{=0} + \frac{h^2}{2}ds\int_{\mathbb{R}} u^2K(u)f"(\zeta_u)du. \\
&= \int_{\mathbb{R}} K(u)f(x+uh)du.
\end{aligned}$$

Il en résulte que

$$\begin{aligned}
|Biais\left(f_n(x)\right)| &= |\mathbb{E}[f_n(x)] - f(x)| \\
&\le \frac{h^2}{2}\left|\int_{\mathbb{R}} u^2K(u)f"(\zeta_u)du\right| \\
&\le \frac{h^2}{2}\int_{\mathbb{R}} u^2|K(u)||f"(\zeta_u)|du \\
&\le h^2\underbrace{\frac{\max_x |f"(x)|}{2}\int_{\mathbb{R}} u^2|K(u)|du}_{C_1}
\end{aligned}$$

d'où la première partie.

Pour prouver la seconde partie, on utilise le faite que les variables aléatoires $Y_i = K\left((X_i - x)/h\right)$, $i = 1 \ldots, n$ sont i.i.d. et que la variance de la somme de variables indépendantes coïncide avec la somme des variances :

$$\begin{aligned}
Var\left[f_n(x)\right] &= \frac{1}{(nh)^2} Var\left[\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)\right] \\
&= \frac{1}{(nh)^2} \sum_{i=1}^{n} Var\left[K\left(\frac{X_i - x}{h}\right)\right] \\
&= \frac{1}{(nh)^2} \times n \times Var\left[K\left(\frac{X_1 - x}{h}\right)\right] \\
&\leq \frac{1}{(nh)^2} \mathbb{E}\left[K\left(\frac{X_1 - x}{h}\right)^2\right] \\
&= \frac{1}{(nh)^2} \int_{\mathbb{R}} K\left(\frac{y - x}{h}\right)^2 f(y)dy \\
&= \frac{1}{nh} \int_{\mathbb{R}} K(u)^2 f(x = uh)du \\
&\leq \frac{1}{nh} \underbrace{\sum_{z} f(z) \int_{\mathbb{R}} K(u)^2 du}_{C_2} .
\end{aligned}$$

C'est exactement ce qu'il fallait démontrer.