We have seen how to consistently estimate a cont. CDF F.

What about the density?

Assume $X_1, \dots, X_n \overset{iid}{\sim} f$

$$f(x) = \frac{dF(x)}{dx}$$

$$\hat{f}_h(x) = \frac{\hat{F}(x+h) - \hat{F}(x-h)}{2h}$$

$$= \frac{\frac{1}{n}\sum^{n} I(x_i < x+h) - I(x_i < x-h)}{2h}$$

$$= \frac{1}{2nh}\sum_{i=1}^{n} I(x-h < x_i \leq x+h)$$

$$= \frac{1}{2nh}\sum_{i=1}^{n} I(|x-x_i| < h)$$

$$\underbrace{\qquad\qquad\qquad\qquad}_{N_h(x)}$$

$$\mathbb{E}\left[I(|x - x_i| < h) \right] = P_h(x)$$

$$\mathbb{E}\left[\hat{f}_n(x) \right] = \frac{1}{2nh} \mathbb{E}\left[N_h(x) \right]$$

$$= \frac{1}{2nh} \, n \, P_h(x)$$

$$= \frac{P_h(x)}{2h}$$

$$= \frac{1}{2h} \int_{x-h}^{x+h} f(t)\,dt$$

$$= \lim_{h \to 0} \frac{f(x+h) - (-f(x-h))}{2}$$

$$= f(x)$$

$$Var[\hat{f}_n(x)] = \frac{1}{4n^2h^2} P_n(x)(1-P_n(x))\, n$$

$$= \frac{P_n(x)(1-P_n(x))}{4nh^2}$$

So for bias we want

$$h \longmapsto 0$$

and for the variance we want

$$nh^2 \longmapsto \infty .$$