Machine Learning and Applications

Assignment 5

1. (3) Let $X_1, \ldots, X_n \sim f$ and let $\widehat{f}_n$ be the kernel density estimator using the boxcar kernel:

$$K(x) = \begin{cases} 1 & \frac{1}{2} < x < \frac{1}{2} \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Show that

$$\mathbb{E}(\widehat{f}_n(x)) = \frac{1}{h} \int\limits_{x-h/2}^{x+h/2} f(y)dy$$

   and

$$\mathbb{V}(\widehat{f}_n(x)) = \frac{1}{nh^2} \left[\int\limits_{x-h/2}^{x+h/2} f(y)dy - \left(\int\limits_{x-h/2}^{x+h/2} f(y)dy \right)^2 \right].$$

   (b) Show that if $h \to 0$ and $nh \to \infty$ as $n \to \infty$, then $\widehat{f}_n(x) \xrightarrow{P} f(x)$.

2. (3) Let us consider pairs of points $(x_1, Y_1), \ldots, (x_n, Y_n)$ related by

$$Y_i = r(x_i) + \epsilon_i,$$

   where $\mathbb{E}\epsilon_i = 0, \quad \mathbb{V}\epsilon = \sigma^2$. Let us assume that x_i are ordered in ascending order. Show that with suitable smoothness assumptions on function $r(x)$, an estimate $\widehat{\sigma}^2$

$$\widehat{\sigma}^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (Y_{i+1} - Y_i)^2$$

   is a consistent estimator of σ^2, i.e. $\widehat{\sigma}^2 \xrightarrow{P} \sigma^2$.