Exercise A (KNN)

Let $y_i = f(x_i) + \varepsilon_i$, and let $\hat{f}(x) = \frac{1}{K} \sum_{i \in \mathcal{N}_K(x)} y_i$. Then, the <u>bias</u> of the KNN estimator is

$$\mathbb{E}\left[\hat{f}(x)\right] = \frac{1}{K} \sum_{i \in \mathcal{N}_K(x)} \mathbb{E}[y_i] = \frac{1}{K} \sum_{i \in \mathcal{N}_K(x)} f(x_i)$$

Hence, the bias is

$$\text{Bias}\left\{\hat{f}(x)\right\} = \frac{1}{K} \sum_{i \in \mathcal{N}_K(x)} f(x_i) - f(x),$$

whereas the variance is

$$\text{var}\left\{\hat{f}(x)\right\} = \text{var}\left(\frac{1}{K} \sum_{i \in \mathcal{N}_K(x)} y_i\right) = \frac{1}{K^2} \sum_{i \in \mathcal{N}_K(x)} \text{var}(y_i) = \frac{\cancel{K}\sigma^2}{\cancel{K^2}} = \frac{\sigma^2}{K}$$

independent r.v!