STAT 5460: Homework I FALL 2025

Write your name on each page. This homework has 2 questions for a total of 30 points, due date is **Monday September 22, 2025**. Please turn in your solutions (CLEAN version) on Monday 22, 2025 during class.

1. (a) [3 points] Show that the kernel density estimate

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right),$$

with kernel K and bandwidth $h > 0$, is a valid density. What condition(s) did you require on K?

(b) [5 points] Show that the kernel density estimate

$$\hat{f}(x) = \frac{1}{nh(x)} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h(x)}\right),$$

with kernel K and bandwidth function $h(x) > 0\,, \forall x$, is a not valid density.

2. A natural estimator for the rth derivative $f^{(r)}(x)$ of $f(x)$ is

$$\widehat{f^{(r)}}(x) = \frac{1}{nh^{r+1}} \sum_{i=1}^{n} K^{(r)}\left(\frac{x - X_i}{h}\right)$$

assuming that K satisfies the necessary differentiability conditions.

(a) [5 points] Derive an asymptotic expression for the bias of $\widehat{f^{(r)}}(x)$. Also mention the assumptions you made to obtain this result.

(b) [5 points] Derive an asymptotic expression for the variance of $\widehat{f^{(r)}}(x)$ (mention the assumptions you made to obtain this result).

(c) [3 points] Calculate the mean squared error (MSE) of $\widehat{f^{(r)}}(x)$.

(d) [3 points] Calculate the mean integrated squared error (MISE) of $\widehat{f^{(r)}}$.

(e) [2 points] From all your previous results, can you conclude why density derivative estimation is becoming increasingly more difficult for estimating higher order derivatives?

(f) [4 points] Find an expression for the asymptotically optimal constant bandwidth.