Name:

ID#:

Section: 455   or   555

	2	8	10

Directions – *The quiz is closed book/notes. You have 10 minutes to complete it; use this paper only.*

Problem 1: Recall (2pts) (Answer in one sentence only.)

What is the role of bandwidth/window width in kernel density estimation?

Solution:
It controls the bias and variance in the resultant estimation.

Problem 2: Work (8 pts) (Show all derivations/work and explain.)
Suppose that we have a dataset $\mathbf{X} = \{0, 1, 1, 1, 2, 3, 4, 4\}$.

1. Estimate the probability $p(x = 1)$ using histogram of bin-width of 1

   Solution:

   $$p(x = 1) = \frac{3}{8}$$

2. Estimate the probability $p(x = 1)$ using the following kernel function, assume the bandwith parameter $h = 2$.

   $$k(t) = \begin{cases} 1 - |t| & \text{if } |t| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   Solution:

   $$p_n(x) = \frac{k_n}{nV_n} = \frac{1}{n} \sum_{i=1}^{N} \frac{1}{h_n^d} K(\frac{x - x_i}{h_n})$$

   $$\begin{aligned} p_0(x) &= \frac{1}{8} \sum_{i=1}^{8} \frac{1}{2^1} K(\frac{0 - x_i}{2}) \\ &= \frac{1}{8 \times 2} \left(K(\frac{1}{2}) + K(0) \times 3 + K(-\frac{1}{2}) + K(-1) + K(-\frac{3}{2}) \times 2 \right) \\ &= \frac{1}{8 \times 2} \left(\frac{1}{2} + 1 \times 3 + \frac{1}{2} + 0 \right) \\ &= \frac{1}{4} \end{aligned}$$