Name: _____

Practice Quiz - 2
CS 591Q/791V - Pattern Recognition
Posted on: April 23, 2010

1. Consider a set of one-dimensional values sampled from an unknown density $p(x)$: 1, 1.5, 1.75, 2, 2.5, 2.75, 3, 5, 6, 6.25, 6.5, 7, 7.5. Estimate the value of the density function, $\hat{p}(x)$, at x=0, 1, 3, 5, 7 and 9, using a *uniform* kernel function of window width 1.

2. Suppose n i.i.d samples are drawn from an unknown density $p(x)$. The basic approach to estimate $p(x)$ based on the n samples uses the following equation:

$$p_n(x) = \frac{k_n/n}{V_n}.$$

   Here, k_n is the number of samples falling in a region R_n of volume V_n placed around x.

   (a) List the conditions required for $p_n(x)$ to converge to $p(x)$.

   (b) What is the most critical parameter in the Parzen window approach to density estimation? What procedure would you adopt to select this parameter?

3. Let $D = \{\boldsymbol{x_1}, \ldots \boldsymbol{x_n}\}$ be a set of n independent labeled samples and let $D_k(\boldsymbol{x}) = \{\boldsymbol{x_1}', \ldots \boldsymbol{x_k}'\}$ be the k nearest neighbors of $\boldsymbol{x}$. Consider a two category problem with $P(\omega_1) = P(\omega_2) = 0.5$. Assume further that the conditional densities of $p(\boldsymbol{x}|\omega_i)$ are uniform within unit hyperspheres that are a distance of 10 units apart. Show that if k is odd the average probability of error of the k nearest neighbor classifier is given by

$$P_n(e) = \frac{1}{2^n} \sum_{j=0}^{(k-1)/2} \binom{n}{j}.$$