**Problem Set #1**

ECONOMICS 241A
SPRING 2010

Due April 7

1. Prove that $O_p(1) + o_p(1) = O_p(1)$ and $O_p(1) \cdot o_p(1) = o_p(1)$. [Hint: you may want to use the inequalities

$$\begin{aligned} \Pr\{A \cup B\} &\leq \Pr\{A\} + \Pr\{B\}, \\ \Pr\{A \cap B) &\leq \Pr\{A\} + \Pr\{B\}.] \end{aligned}$$

2. For the "two-sided" univariate density function estimator

$$\tilde{f}_n(x) = h^{-1}[\hat{F}(x + h/2) - \hat{F}(x - h/2)]$$

of a density $f(x)$ of a sequence of i.i.d. scalar random variables $\{x_i\}$, find the optimal sequence $h^*$ which asymptotically minimizes the mean-squared error of the estimator (pointwise). [Here $\hat{F}(x)$ is the empirical c.d.f. of the observed data.]. Assume the true density $f(x)$ has as many non-zero derivatives as you need. How do the convergence rates of $h^*$ and the MSE to zero compare to the corresponding rates for the "one-sided" density estimator

$$\hat{f}_n(x) = h^{-1}[\hat{F}(x + h) - \hat{F}(x)]?$$

3 For the same two-sided estimator $\tilde{f}_n(x)$ of problem #2, show that the optimal bandwidth sequence $h^*$ is scale equivariant; that is, if $y_i$ is related to $x_i$ by $y_i = \alpha \cdot x_i$ for some $\alpha > 0$, the optimal bandwidth $h_y^*$ for estimation of the density $f_y(y_0)$ at a point $y_0 = \alpha \cdot x_0$ is related to the optimal bandwidth $h_x^*$ for estimation of the density $f_x(x_0)$ by $h_y^* = \alpha \cdot h_x^*$.

4. Given an i.i.d. sample of a bivariate random variable $(x_i, y_i)$, a simple estimator of the conditional mean function $g(x) \equiv E[y_i|x_i = x]$ is the uniform kernel estimator

$$\hat{g}(x) \equiv \left[\sum_{i=1}^{n} w_{in}\right]^{-1} \cdot \left[\sum_{i=1}^{n} w_{in} y_i\right],$$

where the "weight" function $w_{in}$ is defined as

$$w_{in} \equiv h^{-1} 1\{|x - x| < h/2|\},$$

and $h = h_n$ is a nonrandom bandwidth sequence depending upon the sample size $n$.

Suppose the distribution of $x_i$ is discrete, with finite support. That is,

$$\Pr\{x_i = \xi_j\} = \pi_j, \quad j = 1, 2, ..., J$$

for some distinct constants $\xi_1, \xi_2, ..., \xi_J$ with $\sum_i \pi_i = 1$. Under what conditions on the bandwidth sequence $h$ will $\hat{g}(\xi_j)$ be consistent? Asymptotically normal? (Be sure to make your conditions as general as possible, but you need not verify any regularity conditions for the limit theorems you cite.) When it is consistent and asymptotically normal, derive the asymptotic distribution of $\hat{g}(\xi_j)$.

5. Consider the problem of estimating the conditional variance of a scalar random variable $y_i$ given a $p$-dimensional vector of regressors $x_i$. Since

$$\begin{aligned}\sigma^2(x) &\equiv Var\{y_i|x_i = x\} \\ &= E[y_i^2|x_i = x] - (E[y_i|x_i = x])^2 \\ &\equiv m(x) - (g(x))^2,\end{aligned}$$

we can use kernel regression to estimate $m(x)$ and $g(x)$, and then plug them into this expression to estimate $\sigma^2(x)$.

As usual, we'll assume $y_i$ and $x_i$ are i.i.d. across $i$, with as many bounded moments as needed, and that all the unknown density and regression functions have as many continuous derivatives as needed; we'll also assume that $y_i$ and $x_i$ are jointly continuously distributed, with positive marginal density $f(x)$ of $x_i$ at the value $x$.

Under these conditions, using the same (bounded and nonnegative) kernel function $K(u)$ and bandwidth sequence $h_n$ to estimate both the first and second conditional moments $g(x)$ and $m(x)$ of $y_i$, find the asymptotic distribution of an appropriately-normalized version of

$$\hat{\sigma}^2(x) \equiv \hat{m}(x) - (\hat{g}(x))^2,$$

assuming that the bandwidth $h_n$ tends to zero *faster* than the optimal rate, i.e.,

$$h_n = \frac{c}{n^\alpha},$$

where $\alpha \in (1/(p+4), 1/p)$. You need not explicitly check Liapunov conditions, etc. Instead, you should just show that $\hat{\sigma}^2(x)$ is asymptotically equivalent to a kernel regression of a certain function of $y_i$ and $x_i$ on $x_i$, and apply the known asymptotic distribution results for that latter kernel regression.