ECO519. Homework #3

(due: Wednesday May 3rd)

1. Using stochastic equicontinuity arguments, establish the asymptotic distribution of the Censored Least Absolute deviations estimator

$$\widehat{\beta} = \underset{\beta}{\text{argmin}} \frac{1}{n} \sum_{i=1}^{n} \Big| y_i - \max\{0, x_i'\beta\} \Big|.$$

   Keep all the assumptions in Powell's paper.

2. Go back to Problem 1 in Homework 2. For the estimator you proposed in part (b), use stochastic equicontinuity arguments to establish its asymptotic distribution. Use whatever assumptions you need.

   Hint: Recall that, as we saw in lecture, in both cases you need to apply the steps in Equations (3.3)-(3.9) in Andrews (1994) (Handbook of Econometrics). Essentially, you need to show that the empirical process in question satisfies Pollard's entropy condition. This will entail showing that the objective function in each of the two cases belongs to either Type I, II or III as defined by Andrews. You will need to invoke Theorem 2 (p.2272) and Theorem 3 (p.2273).

3. Let $m(x) = E[Y|X = x]$, where $X \in \mathbb{R}$. Consider a nonparametric estimator

$$\widehat{m}(x) = \frac{1}{Nh\widehat{f}(x)} \sum_{i=1}^{N} Y_i K\left(\frac{X_i - x}{h}\right), \quad \text{where } x \in \mathbb{S}(X) \text{ (the support of } X\text{)}$$

   Suppose $K(\cdot)$ is differentiable everywhere in $\mathbb{R}$. Now consider the object

$$\frac{d\widehat{m}(x)}{dx} \equiv \widehat{m'(x)}$$

   (a) Find conditions under which $\widehat{m'(x)}$ is a consistent estimator for $m'(x) \equiv \frac{dE[Y|X=x]}{dx}$. Make whatever smoothness assumptions needed.

   (b) Find conditions –if they exist– under which

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{n} \left(\widehat{m'(X_i)} - m'(X_i)\right) \xrightarrow{d} \mathcal{N}(0, V)$$

   What type of kernel would yield this result? Provide an example of such a kernel –if it exists–. If this result is valid, find the expression for the asymptotic variance V.

4. Let $\mu(W) = E[Z|W]$. Suppose the following model is true about Y, X, W (each of which is a real-valued random variable) and an unobservable ε:

$$Y = \beta_0 + \beta_1 X + \beta_2 \mu(W) + \varepsilon; \quad \text{where } E[\varepsilon|X] = 0 \text{ and } E[\varepsilon|\mu(W)] = 0.$$

Let

$$\widehat{\mu}(w) = \frac{1}{Nh}\sum_{i=1}^{N} Z_i K\left(\frac{W_i - w}{h}\right)$$

and suppose that $\mu(\cdot)$ is assumed to be smooth enough, and $K(\cdot)$ is chosen appropriately so that the following result is true for any $w \in \mathbb{S}(W)$:

$$\widehat{\mu}(w) - \mu(w) = \frac{1}{Nh}\sum_{i=1}^{N} \frac{[Z_i - \mu(w)]}{f_W(w)} K\left(\frac{W_i - w}{h}\right) + \xi_N(w),$$

where $\sup_{w\in\mathbb{S}(W)} (N^{1-\delta}h)|\xi_N(w)| = O_p(1)$ for any $\delta > 0$. With all this in mind, consider the estimator $\widehat{\beta} = (\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2)$ that solves:

$$\widehat{\beta} = \underset{\widehat{\beta}}{\text{argmin}} \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i - \widehat{\beta}_2 \mu(W_i)\right)^2.$$

Find conditions under which $\widehat{\beta}$ is $\sqrt{N}$-consistent, and characterize its asymptotic distribution. Explain why in general, it is more likely that we would have to use a trimmed-sum of squared residuals of the form $\frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i - \widehat{\beta}_2 \mu(W_i)\right)^2 \phi(W_i)$, where $\phi(W_i)$ is different from zero inside some compact set in the interior of $\mathbb{S}(W)$ and zero elsewhere.

5. This problem is about implementing the estimation proposed in Problem 4. First, generate data artificially as follows:

$$X_i \sim \mathcal{N}(0,1); \quad W_i \sim \mathcal{N}(0,4) \quad \text{and} \quad Z_i = 2W_i + 3W_i^2 - W_i^3 + \nu_i, \quad \text{with } \nu_i \sim \mathcal{N}(0,1).$$

Let $\beta_0 = 1$, $\beta_1 = -2$ and $\beta_2 = 3$. Let the sample size be $N = 200$.

Fix the bandwidth sequence to $h = c_1 N^{-\alpha}$ for some $c_1 > 0$ and $\alpha > 0$. Consider the family of kernels of polynomial form:

$$K(x) = \left(a_0 + a_1 x^2 + a_2 x^4 + \cdots + a_\ell x^{2\ell}\right)\mathbb{1}\{|x| < 1\}$$

(notice that this kernel has nonzero coefficients only for even powers of x). Find the values of $\ell, a_0, a_1, \ldots, a_\ell$ that yield the type of kernel you need. Now fix a value $\alpha > 0$ that –together with the kernel– satisfy the conditions you found in Problem 4.

The experiment is the following: Fix a grid of values for c_1, say from $c_1 \in [0.1, 5]$ with increments of 0.1 (you can use a different range of values). Let R be the number of elements in the grid of values for c_1. Generate the data R times and estimate $\widehat{\beta}$ each time. Which estimator(s) are more stable as you change c_1? Which are less stable? What neighborhood of values of c_1 yields the smallest bias in the estimates? Note: There may be a range of values of c_1 for which $\widehat{\beta}$ cannot be computed (if c_1 is such that $\widehat{\mu}(W_i)$ is nearly constant for all i). If this happens, try a different range of values