4th sheet of exercises
due: 13 December 2007

Exercise 13. Consider the kernel regression estimator in the following representation $\hat{g}(x) = \hat{u}(x)/\hat{f}_X(x)$ with

$$\hat{u}(x) = \frac{1}{nh^d}\sum_{j=1}^{n} Y_j \cdot K\big((x-X_j)/h\big)\,,$$
$$\hat{f}_X(x) = \frac{1}{nh^d}\sum_{j=1}^{n} K\big((x-X_j)/h\big)\,.$$

We study the random design case. Define the function $u := g \cdot f_X$ and assume that $f_X(x) \neq 0$ for some $x \in \mathbb{R}^d$. Show that weak consistency of $\hat{g}(x)$ (more precisely, $\hat{g}(x) \to g(x)$ in probability as $n \to \infty$ for that x) follows from

$$E\big|\hat{u}(x) - u(x)\big|^2 + E\big|\hat{f}_X(x) - f_X(x)\big|^2 \stackrel{n\to\infty}{\longrightarrow} 0\,. \tag{1}$$

Exercise 14. Adopt the setting of the previous exercise. Show that

$$E\hat{u}(x) = \big[K_h * u\big](x)\,.$$

Conclude that $E\hat{u}(x) \stackrel{h\to 0}{\longrightarrow} u(x)$ for appropriate kernels whenever u is continuous at x. Also see Lemma 3.1 from the lectures.

Exercise 15. In the setting of the Exercises 13 and 14, prove that

$$\operatorname{var}\hat{u}(x) \stackrel{n\to\infty}{\longrightarrow} 0\,,$$

for appropriate kernels whenever $nh^d \to \infty$.

Exercise 16. Finally, prove that condition (1) from Exercise 13 is satisfied by appropriate selection the kernel and the bandwidth. For the latter part, consider the density estimation problem as a regression problem with $Y_j = 1$ a.s.. Further, summarize the conditions which are required to prove consistency.