Statistics 583, Problem Set 9

Wellner; 5/23/2007

Reading: Wasserman, Chapters 4-9, pages 43-223
Due: Wednesday, May 30, 2006

1. (a) Verify the variance part of Theorem 6.9, Wasserman, page 129.
   (b) Verify the proof of (6.12) - (6.14) in Theorem 6.11, Wasserman, page 129. Two questions about the proof:
   (i) Is the hypothesis "f' is absolutely continuous" (i.e. $f'(x) = \int_0^x f''(y)dy$ for a function f'') used in the proof, or should the hypothesis really be that f is absolutely continuous; i.e. $f(x) = \int_0^x f'(x)dx$ for a function f' (satisfying $\int_0^1 f'(y)^2 dy < \infty$)?
   (ii) Is the $o(1/n)$ term on page 142, line -8, correct? Does this have any effect on (6.12)?
   (c) Can you find the limiting distribution of $n^{1/3}(\widehat{f}_n(x) - f(x))$ under hypotheses similar to those used in (a corrected version of) Theorem 6.11?

2. (a) Consider the kernel density estimator defined in (6.26), Wasserman, page 132. Show that if the density f and the kernel k satisfy the hypotheses of Wasserman's theorem 6.28, page 133, and $h = h_n$ satisfies the hypotheses of Theorem 6.27, then for fixed $x \in \mathbb{R}$,
$$\sqrt{nh_n}(\hat{f}_n(x) - E\hat{f}_n(x)) \rightarrow_d N\left(0, f(x)\int k^2(x)dx\right).$$
   (b) Under what restriction on h_n does it follow (from (a) together with further analysis of the bias) that
$$\sqrt{nh_n}(\hat{f}_n(x) - f(x)) \rightarrow_d N\left(0, f(x)\int k^2(x)dx\right)?$$
   (c) If $h_n = cn^{-1/5}$ and the hypotheses of (a) hold, find the limiting distribution of $\sqrt{nh_n}(\hat{f}_n(x) - f(x))$.
   (d) Under the same assumptions as in (c), find the limiting distribution of $\sqrt{nh_n}(\sqrt{\hat{f}_n(x)} - \sqrt{f(x)})$.
   (e) Suppose that $x, y \in \mathbb{R}$ with $x < y$. Find the joint limiting distribution of $(\sqrt{nh_n}(\hat{f}_n(x) - f(x)), \sqrt{nh_n}(\hat{f}_n(y) - f(y))$ under the assumptions in (b) and (c).

3. (a) Wasserman, problem 6.9.3, page 143.
   (b) Show that (6.35) on page 136 holds.

4. **Bonus optional problem.** In the context of the histogram estimator in problem 1 above, is there a choice of the bin width $h = 1/m$ that leads to the estimator being asymptotically Poisson rather than asymptotically Gaussian?

5. **Bonus optional problem.** (a) Wasserman, problem 6.9.3, page 143.
   (b) Verify the formula (6.16) of Wasserman, page 129.