Statistics 582, Problem Set 4

Wellner; 1/28/98

Reading: Chapter 4, section 6.
Due: Wednesday, February 4, 1998.

1. (MLE's and consistency via reparametrization/compactification).
   Suppose that $X_1, \ldots, X_n$ are i.i.d. P_{θ_0} $\theta_0 \in \Theta = R$ where $\mathcal{P} = \{P_\theta : p_\theta(x) = (dP_\theta/d\mu)(x) = g(x-\theta), \theta \in R\}$ and where $g(x) = \exp(-x)/(1+\exp(-x))^2$ is the logistic density function.
   A. Show that the MLE $\widehat{\theta}_n$ of θ exists and is unique.
   B. Show that the MLE $\widehat{\theta}_n$ of θ is consistent by compactifying the parameter space using the reparametrization shown in the following figure, and then applying Wald's theorem. [Hint: see theorem 4.2 and examples 4.3 and 4.4 in section 4.4.]

2. Use Jensen's inequality to extend the proof of theorem 4.6.2 given in class to the case where ties are possible. That is, suppose that $Y_1, \ldots, Y_m$ are the distinct values appearing in the sample $X_1, \ldots, X_n$ and let $m_j \equiv \#\{i \le n : X_i = Y_j\}$, $q_j \equiv Q(\{Y_j\})$ so that $\sum_{j=1}^k m_j = n$, and $\sum_{j=1}^m q_j \le 1$. Then show that

$$\prod_{j=1}^k q_j^{m_j} \le \prod_{j=1}^k (\frac{m_j}{n})^{m_j},$$

   and that the resulting maximizer yields the empirical measure $\mathbb{P}_n$.

3. (Interval censoring case 1 or "current status data".) Suppose that (X_i, Y_i), $i = 1, \ldots, n$ are i.i.d. pairs of nonnegative random variables with $X_i \sim F$, $Y_i \sim G$, and X_i, Y_i independent for each i. Suppose that we observe $(Y_i, \delta_i) \equiv (Y_i, 1_{[X_i \le Y_i]})$, $i = 1, \ldots, n$. as noted in class on 1/23, $(\delta_i | Y_i) \sim$Bernoulli$(F(Y_i))$, and if G has density g with respect to some dominating measure μ, the joint density of (Y_i, δ_i) is

$$p_{F,g}(y, \delta) = F(y)^\delta (1 - F(y))^{1-\delta} g(y).$$

   Suppose that we order the observed Y_i's as $0 \le Y_{(1)} \le \ldots \le Y_{(n)}$, assume no ties, and denote the corresponding δ's by $\delta_{(i)}$, and set $p_i \equiv F(Y_{(i)})$, $q_i \equiv G(\{Y_{(i)}\})$.

Then a nonparametric likelihood for estimation of F and G is given by

$$L_n(F, G|\underline{(Y, \delta)}) = \prod_{i=1}^{n} p_i^{\delta_{(i)}}(1 - p_i)^{1-\delta_{(i)}} q_i \equiv L_n(\underline{p}, \underline{q}) ,$$

and hence the log-likelihood is

$$l_n(\underline{p}, \underline{q})) = \sum_{i=1}^{n} \{\delta_{(i)} \log p_i + (1 - \delta_{(i)}) \log(1 - p_i)\} + \sum_{i=1}^{n} q_i .$$

We want to maximize this over $\underline{p} = (p_1, \ldots, p_n)$ and $\underline{q} = (q_1, \ldots, q_n)$ satisfying $0 \le p_1 \le \ldots \le p_n \le 1$ and $q_i \ge 0$, $\sum_{i=1}^{n} q_i \le 1$. the maximum over the q's is easy (since we have done it already) and yields the empirical distribution of the Y_i's as an estimator of G. Suppose that $n = 5$, $Y_{(1)} = 1.2$, $Y_{(2)} = 2.3$, $Y_{(3)} = 2.7$, $Y_{(4)} = 3.1$, $Y_{(5)} = 3.9$, $\delta_{(1)} = \delta_{(3)} = \delta_{(4)} = 1$, $\delta_{(2)} = \delta_{(5)} = 0$. Show that the vector $\underline{\widehat{p}}$ maximizing the loglikelihood is given by $\widehat{p}_1 = \widehat{p}_2 = 1/2$, $\widehat{p}_3 = \widehat{p}_4 = \widehat{p}_5 = 2/3$, and that this corresponds to the left-derivative of the greatest convex minorant of the points $\{(i, \sum_{j \le i} \delta_{(j)}), i = 0, \ldots, n\}$ where $(0, 0)$ corresponds to $i = 0$.

4. (Optional bonus problem: Problem 3.1, continued.) Suppose the same context as in Problem 3.1. Assuming that $\overline{\beta}_n \to_p \beta_0$, you used a uniform strong law of large numbers to show that

$$\overline{\alpha} = \left\{ \frac{1}{n} \sum_{i=1}^{n} X_i^{\overline{\beta}} \right\}^{1/\overline{\beta}} .$$

Now suppose in addition that $\overline{\beta}_n$ is asymptotically linear with influence function $\psi = \psi(x; \alpha, \beta)$ Thus $E_0\psi(X) = 0$, $E_0\psi^2(X) < \infty$, and

$$\sqrt{n}(\overline{\beta}_n - \beta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi(X_i) + o_p(1) \to_d N(0, E_0(\psi^2(X))) .$$

Show that $\overline{\alpha}_n$ is asymptotically linear and find it's influence function.
[Hint: what would you like as the appropriate generalization / extension of the uniform strong law of large numbers to solve this problem?]