**Statistics 581, Problem Set 6**

Wellner; 11/1/2000

**Reminder:** Midterm Exam, Monday, November 6.
**Reading:** Chapter 2, Sections 4-6; start reading Chapter 3.
**Due:** Wednesday, November 8, 2000.

1. Consider the tests of $H : p = p_0$ versus $K : p \neq p_0$ based on $Q_n$ and $4nH_n^2$ that we discussed in class and in problem set 3. These tests can be used to form confidence sets $C_n$ and $D_n$ for $p$ as follows: define $Q_n(p_0) = \sum_{j=1}^k (N_j - np_{j0})^2/np_{j0}$ and $H_n^2(p_0) \equiv 4n\sum_{j=1}^k (\sqrt{\hat{p}_j} - \sqrt{p_{j0}})^2$, so that both $Q_n(p_0)$ and $H_n^2(p_0)$ are asymptotically $\chi_{k-1}^2$ under the null hypothesis. Let
$$C_n = \{p_0 : Q_n(p_0) \le \chi_{k-1,\alpha}^2\}$$
and
$$D_n = \{p_0 : H_n^2(p_0) \le \chi_{k-1,\alpha}^2\} .$$
Show that these both yield asymptotic $1-\alpha$ confidence sets: e.g.
$$P_{p_0}(p_0 \in C_n) \to 1 - \alpha \qquad \text{as } \ n \to \infty .$$
Can you use our "Theorem 2" about $Q_n(p_0)$ and $H_n^2(p_0)$ under $p \neq p_0$ to prove some other property of these confidence sets?

2. Suppose that Suppose that $X_1, \ldots, X_n, \ldots$ are i.i.d. random vectors in $R^k$ with common distribution function $F$ and corresponding probability measure $P$ on $(R^k, \mathcal{B}_k)$. Let $\mathbb{P}_n$ be the empirical measure defined by
$$\mathbb{P}_n = n^{-1}\sum_{i=1}^n \delta_{X_i} ,$$
and consider $\mathbb{P}_n$ and the empirical process $\mathbb{G}_n$ as indexed by a class of sets $\mathcal{C} \subset \mathcal{B}_k$:
$$\{\mathbb{P}_n(C) : \ C \in \mathcal{C}\}, \qquad \{\mathbb{G}_n(C) : \ C \in \mathcal{C}\} ,$$
where
$$\mathbb{G}_n \equiv \sqrt{n}(\mathbb{P}_n - P) .$$
(a) Show that $\mathbb{G}_n \to_{f.d.} \mathbb{G}_P$ where $\mathbb{G}_P$ is a $P$-Brownian bridge process indexed by $\mathcal{C}$: i.e. show that for any integer $m$ and sets $C_1, \ldots, C_m \in \mathcal{C}$,
$$(\mathbb{G}_n(C_1), \ldots, \mathbb{G}_n(C_m)) \to_d (\mathbb{G}_P(C_1), \ldots, \mathbb{G}_P(C_m)) \sim N_m(0, \Sigma)$$
where $\Sigma = (\sigma_{jj'})$ is given by
$$\sigma_{jj'} = P(C_j \cap C_{j'}) - P(C_j)P(C_{j'}) .$$
(b) When $\mathcal{C} = \mathcal{O} \equiv \{(-\infty, x] : \ x \in R^k\}$ specialize the result in (a) and show that it gives the finite-dimensional convergence of the empirical distribution function $\mathbb{F}_n$: i.e.
(i) show that $\mathbb{P}_n((-\infty, x]) = \mathbb{F}_n(x)$;
(ii) show that $P((-\infty, x]) = F(x)$;
(iii) show that $\mathbb{Y}(x) \equiv \mathbb{G}_P((-\infty, x])$ has mean zero and covariance
$$E\{\mathbb{Y}(x)\mathbb{Y}(y)\} = F(x \wedge y) - F(x)F(y), \qquad x, y \in R^k .$$