Statistics 521, Problem Set 6

Wellner; 10/31/2007

Reminder: Midterm exam on Friday, November 9.
Reading: Shorack, PfS, Chapter 4, sections 4.1 - 4.4, pages 67 - 85; Chapter 5, sections 5.1 - 5.3, pages 87 - 97.

Due: Wednesday, November 7, 2007.

1. PfS, Exercise 3.5.7, page 61, modified as follows: Suppose that $f_0, f_1, \ldots$ are ≥ 0, defined on a sigma-finite measure space $(\Omega, \mathcal{A}, \mu)$. (a) Suppose that $\int_\Omega f_n d\mu = 1$ for $n = 0, 1, \ldots$, and $f_n \to_{a.e.} f_0$ with respect to μ. Show that

$$\sup_{A \in \mathcal{A}} \left| \int_A f_n d\mu - \int_A f_0 d\mu \right| \to 0 \qquad \text{as} \quad n \to \infty \, .$$

   (b) Show that the conclusion of (a) holds if just $f_n \to_\mu f_0$ and $\int_\Omega f_n d\mu \to \int_\Omega f_0 d\mu$.

2. Suppose that P, Q are two probability measures on the same measurable space $(\Omega, \mathcal{A})$ which are both absolutely continuous with respect to the measure μ with densities (Radon-Nikodym derivatives) p and q respectively. Thus $P(A) = \int_A p d\mu$ and $Q(A) = \int_A q d\mu$ for $A \in \mathcal{A}$. Show that

$$d_{TV}(P, Q) \equiv \sup_{A \in \mathcal{A}} |P(A) - Q(A)| = \frac{1}{2} \int |p - q| d\mu = \int (p - q)^+ d\mu \, .$$

3. Suppose that $X_n \sim \text{Binomial}(n, p_n)$ for $n = 1, 2, \ldots$ with $np_n \to \lambda > 0$, and let P_n be the induced distribution of X_n on R. Let $X_0 \sim \text{Poisson}(\lambda)$ and let P_0 be the corresponding induced distribution on R. Show that $d_{TV}(P_n, P_0) \to 0$ as $n \to \infty$.

4. PfS, Exercise 4.11, page 51: If X is a non-negative random variable satisfying $\int_0^\infty \sqrt{P(X > t)} dt < \infty$, then we say that $X \in \mathcal{L}_{2,1}$.
   (a) Show that if $X \in \mathcal{L}_{2,1}$, then $X \in \mathcal{L}_2$.
   (b) Show that if $X \in \mathcal{L}_r$ for some $r > 2$, then $X \in \mathcal{L}_{2,1}$.

5. PfS, Exercise 4.1.2, page 67: Identify ϕ^+, ϕ^-, $|\phi|$ and $|\phi|(\Omega)$ in the context of the prototypical situation of example 4.1.1, page 66. Be sure to specify Ω^+ and Ω^-.

6. **Optional bonus problem:**
Let $X_{n1}, \ldots, X_{nn}$ be independent, $X_{nk} \sim \text{Bernoulli}(p_{nk})$, and let $Y_n \sim \text{Poisson}(\sum_{k=1}^n p_{nk})$. Let P_n be the distribution of $\sum_{k=1}^n X_{nk}$ and let Q_n be the distribution of Y_n. Show that

$$d_{TV}(P_n, Q_n) \equiv \sup_{A \in \mathcal{B}} |P(S_n \in A) - P(Y_n \in A)| \leq \sum_{k=1}^n p_{nk}^2.$$

Note that when $p_{nk} = p_n \to 0$ for all k and $np_n \to \lambda$, then $\sum_{k=1}^n p_{nk}^2 = np_n^2 = (np_n)^2/n = O(n^{-1})$.

Hint: Construct S_n and Y_n on a common probability space as follows: let $T_{nk} \sim \text{Poisson}(p_{nk})$, $k = 1, \ldots, n$ be independent, and let $Z_{nk} \sim \text{Bernoulli}(1 - (1 - p_{nk})e^{p_{nk}})$, $k = 1, \ldots, n$ be independent and independent of the T_{nk}'s. Define

$$X_{nk} = 1_{[T_{nk} \geq 1]} + 1_{[T_{nk}=0]} 1_{[Z_{nk}=1]}.$$

Set $S_n = \sum_{k=1}^n X_{nk}$, $Y_n = \sum_{k=1}^n T_{nk}$. Check that $X_{nk} \sim \text{Bernoulli}(p_{nk})$, $Y_n \sim \text{Poisson}(\sum_1^n p_{nk})$, and

$$\begin{aligned} P(T_{nk} = 0, X_{nk} = 1) &= e^{-p_{nk}} - (1 - p_{nk}) \\ P(T_{nk} \geq 1, X_{nk} = 0) &= 0 \\ P(T_{nk} \geq 2) &= 1 - e^{-p_{nk}} - p_{nk} e^{-p_{nk}}. \end{aligned}$$

Show that

$$d_{TV}(P_n, Q_n) \leq P(S_n \neq Y_n) \leq \sum_{k=1}^n P(X_{nk} \neq T_{nk}) \leq \sum_{k=1}^n p_{nk}^2.$$