## Statistics 581, Problem Set 3

Wellner; 10/13/2010

**Reading:** Lehmann & Casella, TPE, pages 54-61 and pages 75-78.
Ferguson, ACILST, pages 1 - 60.
**Due:** Wednesday, October 20, 2010.

1. A sequence of random variables $Y_n$ is *bounded in probability* and we write $Y_n = O_p(1)$ if
$$\lim_{\lambda \to \infty} \limsup_{n \to \infty} P(|Y_n| > \lambda) = 0;$$
i.e. for each $\epsilon > 0$ there exist $\lambda_\epsilon$ and $N_\epsilon$ such that $P(|Y_n| > \lambda_\epsilon) < \epsilon$ for all $n > N_\epsilon$.
(a) Show that if $Y_n \to_d Y$ for some random variable $Y$, then $Y_n$ is bounded in probability. (This is Lehmann and Casella, problem 8.24, page 77.)
(b) Give an example of a sequence of random variables $Y_n$ that is bounded in probability, but does not converge in distribution.
(c) Lehmann and Casella, problem 8.25, page 77.

2. Suppose that $X$ is a random variable with finite fourth moment; $E|X|^4 < \infty$. Then $\mu_4 = E(X - \mu)^4$ is the fourth central moment of $X$. The ratio $\mu_4/\sigma^4 \equiv \kappa$ is the *kurtosis* of $X$ (or of the distribution function $F$ of $X$), and $\gamma_2 \equiv \mu_4/\sigma^4 - 3$ is called the *excess of kurtosis*; note that for any $N(\mu, \sigma^2)$ random variable, $\gamma_2 = 0$. Investigate the value of $\gamma_2$ for various classical distributions ($t_r$, uniform, bernoulli, Poission($\lambda$), ... ). How big can $\gamma_2$ be? How small can $\gamma_2$ be?

3. Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with mean $\mu$, variance $\sigma^2$, and $E|X|^4 < \infty$.
(a) Show that the sample variance $S_n^2 = \sum_{i=1}^n (X_i - \overline{X}_n)^2/(n-1)$ satisfies
$$\sqrt{n}(S_n^2 - \sigma^2)/\sqrt{2}\sigma^2 \to_d N(0, 1 + \gamma_2/2).$$
(b) Suppose that you want to test $H : \sigma \le \sigma_0^2$ versus $K : \sigma^2 > \sigma_0^2$ for $\sigma_0$ a fixed number, and you base your test on normal theory, but in fact the $X$'s are *not normal* with $\gamma_2 \ne 0$. What effect does this have on the level (or size or actual type one error) of the normal theory test?

4. Suppose that $X_1, \ldots, X_n$ are independent Poisson($\lambda$) random variables (so $P(X_1 = k) = e^{-\lambda}\lambda^k/k!$, $k = 0, 1, \ldots$).
(a) Show that $\sqrt{n}(\overline{X}_n - \lambda) \to_d N(0,$ "*something*"$)$.
(b) Show that the sequence $\{\sqrt{n}|\overline{X}_n - \lambda|\}$ is uniformly integrable and find $\lim_{n\to\infty} E(\sqrt{n}|\overline{X}_n - \lambda|)$.
(c) Let $g(x) = x^\gamma$ for $x \ge 0$ and $0 < \gamma < \infty$. Show that $\sqrt{n}(g(\overline{X}_n) - g(\lambda)) \to_d N(0, V^2)$ and compute $V^2$ explicitly in terms of $\lambda$ and $\gamma$. For what $\gamma$ is $V^2$ constant in $\lambda$? Is this the value of $\gamma$ that makes $g(\overline{X}_n)$ "most nearly normal"?

5. Suppose that $X_1, X_2, \ldots$ are i.i.d. positive random variables, and define $\overline{X}_n \equiv n^{-1}\sum_{i=1}^n X_i$, $H_n \equiv 1/(n^{-1}\sum_{i=1}^n (1/X_i))$, and $G_n \equiv \{\prod_{i=1}^n X_i\}^{1/n}$ to be the *arithmetic, harmonic, and geometric* means respectively. We know that $\overline{X}_n \to_{a.s.} E(X_1) = \mu$ if and only if $E|X_1| < \infty$.

(a) Use the SLLN together with appropriate additional hypotheses to show that $H_n \to_{a.s.} 1/\{E(1/X_1)\} \equiv h$, and $G_n \to_{a.s.} \exp(E\{\log X_1\}) \equiv g$.
(c) Use the multivariate CLT and the delta method to find the joint limiting distribution of $\sqrt{n}(\overline{X}_n - \mu, H_n - h, G_n - g)$. You will need to impose or assume additional moment conditions to be able to prove this. Specify these additional assumptions carefully.

6. **Optional bonus problem 1:** Ferguson, ACILST, problem 5, page 18 (section 3): Let $X_{n,1}, \ldots, X_{n,n}$ be independent, $X_{n,k}$ ~Bernoulli($p_{n,k}$), and let $Y_n$ ~Poisson($\sum_{k=1}^n p_{n,k}$). Let $P_n$ be the distribution of $\sum_{k=1}^n X_{n,k}$ and let $Q_n$ be the distribution of $Y_n$. Show that

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

Specialize this to the case when $p_{n,k} = p_n = \lambda/n$ for all $1 \le k \le n$.

7. **Optional bonus problem 2:** Let $P_r$ denote the $t_r-$distribution with $r$ "degrees of freedom" with $r > 0$, $r$ real, defined for Borel subsets $A$ of $\mathbb{R}$ by

$$P_r(A) = \int_A p_r(x)dx = \int_A \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\,\Gamma(r/2)} \frac{1}{(1 + x^2/r)^{(r+1)/2}} dx,$$

and let $Q$ denote the standard normal distribution with d.f. $\Phi$ and density $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$. Investigate $d_{TV}(P_r, Q)$ as a function of $r$. For what sequences $C_r \to 0$ does it hold that $d_{TV}(P_r, Q) \le C_r$?