Statistics 581, Problem Set 3

Wellner; 10/11/00

Reading: Lehmann & Casella, TPE, pages 54-61 and pages 75-78.
Ferguson, ACILST, pages 1 - 60.
Due: Wednesday, October 18, 2000.

1. Suppose that $Y \sim \chi_n^2(\delta)$. Compute $E(Y)$ and $Var(Y)$.
Hint: use the formulas $E(Y) = E\{E(Y|K)\}$ and
$$Var(Y) = Var(E(Y|K)) + E\{Var(Y|K)\}\,.$$

2. Lehmann and Casella, TPE, problem 8.25, page 77 Note problem 8.24 on the same page.

3. Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with mean μ, variance σ^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
$$\sqrt{n}(S_n^2 - \sigma^2)/\sqrt{2}\sigma^2 \to_d N(0, 1 + \gamma_2/2).$$
where $\mu_4 \equiv E(X-\mu)^4$ and $\gamma_2 \equiv \mu_4/\sigma^4 - 3$ is called the *excess of kurtosis*.
(b) Suppose that you want to test $H : \sigma \le \sigma_0^2$ versus $K : \sigma^2 > \sigma_0^2$ for σ_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 $N(0,1)$ random variables, and let $Y_i = X_i^2$, for $i = 1, \ldots, n$. Thus $\sum_1^n Y_i \sim \chi_n^2$.
(a) Show that $\sqrt{n}(\overline{Y}_n - 1) \to_d N(0, \text{“}something''\text{)}$, and find “something”.
(b) Show that for each $r > 0$, $\sqrt{n}(\overline{Y}_n^r - 1) \to_d N(0, V^2(r))$ and find $V^2(r)$ as a function of r.
(c) Show that
$$\frac{\sqrt{n}(\overline{Y}_n^{1/3} - (1 - 2/(9n)))}{\sqrt{2/9}} \to_d N(0,1)\,.$$
Does this agree with your result in (b)?
(d) Make normal probablility plots to compare the approximations in (a) and (c). [The transformation in (c) is called the “Wilson-Hilferty” transformation of a χ^2 random variable.

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:** Ferguson, ACILST, problem 5, page 18:
Let $X_{n1}, \ldots, X_{nn}$ be independent, X_{nk} $\sim$Bernoulli(p_{nk}), and let Y_n $\sim$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 $T_{nk} \sim$ Poisson(p_{nk}), $k = 1, \ldots, n$ be independent, and let $Z_{nk} \sim$ 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 $X_{nk} \sim$ Bernoulli(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

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