Statistics 581, Problem Set 6

Wellner; 10/29/2008

Reminder: Midterm exam: Monday, November 3.
Reading: Lecture Notes Chapter 3, sections 1-2;
Ferguson, ACILST, chapters 19-20, pages 126 - 139;
Lehmann and Casella, TPE, Sections 2.5 and 2.6, pages 113 - 129;
and Section 6.2, pages 437 - 443.
Due: Wednesday, November 5, 2008.

1. Chapter 2, Exercise 5.3, page 25. [Hint: One approach uses the fact that $\mathbb{S}_n(t_j) - \mathbb{S}_n(t_{j-1}) = n^{-1/2}\sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i$, $j = 1, \ldots, t_k$ with $t_0 \equiv 0$ are independent random variables.]

2. Ferguson, ACILST, problem 4, page 93 (modified slightly): suppose that $X_1, \ldots, X_n$ are i.i.d. F with continuous and positive density f in neighborhoods of $F^{-1}(p)$, $F^{-1}(1/2)$, and $F^{-1}(1-p)$ for some $0 < p < 1/2$. (Ferguson takes $p = 1/4$.)
(a) Find the asymptotic distribution of the mid-p-quantile range $R_n(p) \equiv (X_{(n(1-p))} + X_{(np)})/2$; i.e. find the asymptotic distribution of $\sqrt{n}(R_n(p) - r(p))$ where $r(p) = (F^{-1}(1-p) + F^{-1}(p))/2$.
(b) Find the asymptotic distribution of the median.
(c) For a general distribution function F, the mid-p-quantile range and median estimate different parameters, the population mid-p-quantile range $r(p)$ and the population median $F^{-1}(1/2)$ respectively, but in the case of a distribution function F that is symmetric about some point μ (so $1 - F(x+\mu) = F(x-\mu)$), they both estimate the point of symmetry, μ. Compute the asymptotic relative efficiency of the mid-p-quantile range relative to the median as a function of p when: (i) F is Cauchy(μ, σ); (ii) F is Uniform$(0, 2\mu)$.

3. Suppose that $X, X_1, X_2, \ldots, X_n$ are independent Exponential(λ) random variables:
$$P(X \ge x) = \exp(-\lambda x), \qquad x > 0.$$
(a) Show that the r-th moment of X, $\mu_r \equiv \mu_r(\lambda)$ is given by
$$\mu_r(\lambda) = EX^r = \frac{\Gamma(r+1)}{\lambda^r}.$$
(b) Use the moment calculation in (a) to show that
$$\frac{\mu_r(\lambda)}{\mu_{r+1}(\lambda)} = \frac{\lambda}{r+1}$$
and hence that the family of estimators $\{\hat\lambda_n^{(k)}\}_{k\ge 0}$ given by
$$\hat\lambda_n^{(k)} \equiv (k+1)\frac{\overline{X^k}_n}{\overline{X^{k+1}}_n} \equiv (k+1)\frac{n^{-1}\sum_1^n X_i^k}{n^{-1}\sum_1^n X_i^{k+1}}$$

are all consistent estimators of λ: $\hat{\lambda}_n^{(k)} \to_p \lambda$ for each $k = 0, 1, 2, \ldots$.
(c) Show that

$$\sqrt{n}(\hat{\lambda}_n^{(k)} - \lambda) \to_d N(0, \sigma_k^2(\lambda)) \quad \text{as} \quad n \to \infty$$

and compute $\sigma_k^2(\lambda)$ explicitly as a function of k and λ.
(d) What is the asymptotic relative efficiency of $\hat{\lambda}_n^{(k)}$ to $\widehat{\lambda}_n \equiv \hat{\lambda}_n^{(0)} = 1/\overline{X}_n$ for $k > 1$?
(e) Now suppose that $X, X_1, \ldots, X_n$ are i.i.d. with distribution function F on $(0, \infty)$ where F is not an exponential distribution function. Specify hypotheses on F (or X) which guarantee that $\hat{\lambda}_n^{(k)} \to_p$ some natural parameter, say $\lambda_k(F)$ defined in terms of F. What hypothesis will be needed to guarantee that $\sqrt{n}(\hat{\lambda}_n^{(k)} - \lambda_k(F)) \to_d N(0, V^2)$ for some V^2?

4. **Optional bonus problem:** Consider a function $T : \mathcal{F} \to \mathbb{R}$ where $\mathcal{F}$ is some (sub) class of distribution functions F (examples include the mean, $T(F) = \mu(F) = \int x dF(x)$, the variance $T(F) = \sigma^2(F) = \int (x - \int y dF(y))^2 dF(x)$, the median $T(F) = F^{-1}(1/2)$, linear combinations of order statistics $T(F) = \int_0^1 F^{-1}(u) w(u) du$, the *mean residual life function* at $x > 0$ $T(F) \equiv e(x, F) \equiv \int_{(x,\infty)} (1 - F(u)) du/(1 - F(x)) = E(X - x | X > x)$, and so forth). [The mean residual life function gives the mean life conditional on surviving beyond x.] The "principle of substitution" says that $T(F)$ can be estimated by $T(\widehat{F}_n)$. for some estimator $\widehat{F}_n$ of F. If T is sufficiently "smooth", then frequently the empirical distribution function $\mathbb{F}_n$ can be taken as the estimator $\hat{F}_n$ of F.

   Give a treatment of consistency and asymptotic normality of the estimator $e(x, \mathbb{F}_n)$ of $e(x, F)$ based on our results from sections 2.4 and 2.6. You may assume that with $X \sim F$ on $(0, \infty)$ we have $E_F X < \infty$, $E_F X^2 < \infty$, and $1 - F(x) > 0$ (as well as any other additional assumptions you need).