Statistics 581, Problem Set 6

Wellner; 11/3/2010

Reminder: Midterm exam: Monday, November 8.
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 10, 2010.

1. Suppose that $X, X_1, X_2, \ldots, X_n$ are i.i.d. with exponential$(1/\theta)$ distribution given by $1 - F_\theta(x) = \exp(-x/\theta)$, $x \ge 0$. Thus the density of X is $f(x;\theta) = (1/\theta)\exp(-x/\theta)1_{[0,\infty)}(x)$.
(a) Let $M_n = \mathbb{F}_n^{-1}(1/2)$ be the sample median. Find a constant c so that $cM_n \to_p \theta$, and then show that $\sqrt{n}(cM_n - \theta) \to_d N(0, \sigma^2)$ for some σ^2 and find σ^2 in terms of θ.
(b) Do the same with M_n replaced by the $p-$th sample quantile $\mathbb{F}_n^{-1}(p)$ where $0 < p < 1$. Find p which minimizes the asymptotic variance.

2. Suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution function F which has positive density f at its quartiles $F^{-1}(1/4)$ and $F^{-1}(3/4)$ and at its median $F^{-1}(1/2)$.
(a) Let $Q_n = (X_{(3n/4)} + X_{(n/4)})/2 = (\mathbb{F}_n^{-1}(3/4) + \mathbb{F}_n^{-1}(1/4))$, the mid-quartile range. Find the asymptotic distribution of Q_n as an estimator of the population mid-quartile range $Q = Q(F) = (F^{-1}(3/4) + F^{-1}(1/4))/2$. That is, prove that
$$\sqrt{n}(Q_n - Q) \to_d \text{“}\textit{something}\text{”}$$
and find “something”.
(b) Assuming that the underlying distribution is Cauchy(μ, σ) ($X = \sigma Y + \mu$ where $Y \sim$ Cauchy$(0,1)$), compare the variances of the mid-quartile range Q_n and the median M_n as estimators of μ.

3. **Optional bonus problem 1:** 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$(0, 2\mu)$.

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