Statistics 581, Problem Set 6 Solutions

Wellner; 11/3/2010

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 $cM_n \rightarrow_p \theta$, and then show that $\sqrt{n}(cM_n - \theta) \rightarrow_d N(0,\sigma^2)$ 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.

   Solution: (a) First note that $F_\theta^{-1}(t) = \theta\log[1/(1-t)]$ and hence $M_n = \mathbb{F}_n^{-1}(1/2) \rightarrow_{a.s.} F_\theta^{-1}(1/2) = \theta\log 2$. It follows that $\mathbb{F}_n^{-1}(1/2)/\log 2 \rightarrow_{a.s.} \theta$, and thus $c = 1/\log 2$ works. Furthermore, since $Q'(1/2) = 1/f_\theta(F_\theta^{-1}(1/2)) = 2\theta$, it follows that

$$\sqrt{n}(\mathbb{F}_n^{-1}(1/2) - F_\theta^{-1}(1/2)) \rightarrow_d Q'(1/2)N(0,1/4) = 2\theta N(0,1/4) = N(0,\theta^2)$$

   and hence

$$\sqrt{n}(cM_n - \theta) \rightarrow_d cN(0,\theta^2) = N(0,c^2\theta^2) = N(0,\theta^2/(\log 2)^2)\,.$$

   B. When we use $\mathbb{F}_n^{-1}(p)$ for $p \in (0,1)$, then we have

$$\mathbb{F}_n^{-1}(p) \rightarrow_{a.s.} F_\theta^{-1}(p) = \theta\log[1/(1-p)]\,,$$

   so with $c = c_p \equiv -1/\log(1-p)$, $c_p\mathbb{F}_n^{-1}(p) \rightarrow_{a.s.} \theta$. Since $Q'(p) = 1/f_\theta(F_\theta^{-1}(p)) = \theta/(1-p)$, it follows that

$$\sqrt{n}(\mathbb{F}_n^{-1}(p) - F_\theta^{-1}(p)) \rightarrow_d Q'(p)N(0,p(1-p)) = \frac{\theta}{1-p}N(0,p(1-p)) = N\left(0,\theta^2\frac{p}{1-p}\right)\,,$$

   and hence

$$\begin{aligned}\sqrt{n}(c_p\mathbb{F}_n^{-1}(p) - \theta) &= c_p\sqrt{n}(\mathbb{F}_n^{-1}(p) - F_\theta^{-1}(p)) \\ &\rightarrow_d c_pN\left(0,\theta^2\frac{p}{1-p}\right) = N\left(0,c_p^2\theta^2\frac{p}{1-p}\right) \\ &= N\left(0,\theta^2\frac{p}{((1-p)\{\log(1-p))^2\}}\right) \equiv N(0,\theta^2V^2(p))\,.\end{aligned}$$

   The function $V^2(p)$ is minimized by $p \approx .797... \equiv p_0$ and the minimum value is $V^2(p_0) \approx 1.544 < 1/(\log 2)^2 \approx 2.08....$

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

$$\sqrt{n}(Q_n - Q) \rightarrow_d \text{“}\textit{something}\text{”}$$

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.

Solution: (a) First, note that

$$Q_n = \frac{1}{2}(\mathbb{F}_n^{-1}(1/4) + \mathbb{F}_n^{-1}(3/4)) \rightarrow_p \frac{1}{2}(F^{-1}(1/4) + F^{-1}(3/4)) \equiv Q\,.$$

Furthermore

$$\sqrt{n}\begin{pmatrix} \mathbb{F}_n^{-1}(1/4) - F^{-1}(1/4) \\ \mathbb{F}_n^{-1}(3/4) - F^{-1}(3/4) \end{pmatrix} \rightarrow_d - \begin{pmatrix} Q'(1/4)\mathbb{U}(1/4) \\ Q'(3/4)\mathbb{U}(3/4) \end{pmatrix} \sim N_2(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} \frac{3}{16}Q'(1/4)^2 & \frac{1}{16}Q'(1/4)Q'(3/4) \\ \frac{1}{16}Q'(1/4)Q'(3/4) & \frac{3}{16}Q'(3/4)^2 \end{pmatrix}\,.$$

Thus

$$\begin{aligned} \sqrt{n}(Q_n - Q) &= \frac{1}{2}\underline{1}^T \sqrt{n}\begin{pmatrix} \mathbb{F}_n^{-1}(1/4) - F^{-1}(1/4) \\ \mathbb{F}_n^{-1}(3/4) - F^{-1}(3/4) \end{pmatrix} \\ &\rightarrow_d -\frac{1}{2}(Q'(1/4)\mathbb{U}(1/4) + Q'(3/4)\mathbb{U}(3/4)) \\ &\sim N(0, V^2) \end{aligned}$$

where

$$\begin{aligned} V^2 &= \frac{1}{4}\left\{Q'(1/4)^2\frac{3}{16} + 2Q'(1/4)Q'(3/4)\frac{1}{16} + Q'(3/4)^2\frac{3}{16}\right\} \\ &= \frac{1}{64}\left\{3Q'(1/4)^2 + 2Q'(1/4)Q'(3/4) + 3Q'(3/4)^2\right\}. \end{aligned}$$

(b) If $F(x) = F_0((x-\mu)/\sigma)$, where $F_0(x) = (1/\pi)(\arctan(x) + \pi/2)$ is the standard Cauchy$(0, 1)$ distribution function, then

$$F^{-1}(t) = \sigma F_0^{-1}(t) + \mu$$

where $F_0^{-1}(t) = \tan(\pi(t - 1/2))$ from problem 5.2. Furthermore

$$Q = \frac{1}{2}(F^{-1}(3/4) + F^{-1}(1/4)) = \mu + \sigma\frac{1}{2}(F_0^{-1}(3/4) + F_0^{-1}(1/4)) = \mu$$

by symmetry of the Cauchy$(0, 1)$ distribution about 0 (so that $F_0^{-1}(1/4) = \tan(\pi/4) = -\tan(-\pi/4) = F_0^{-1}(3/4)$). Note that $Q'(t) = \sigma Q_0'(t)$ where $Q_0'(t) = \pi/\cos^2(\pi(t-1/2))$. Also note that $Q_0'(1/4) = Q_0'(3/4)$. Thus V^2 in part (a) becomes

$$\begin{aligned} V^2 &= \frac{\sigma^2}{64}\left\{3Q_0'(1/4)^2 + 2Q_0'(1/4)Q_0'(3/4) + 3Q_0'(3/4)^2\right\} \\ &= \frac{\sigma^2}{64}\left\{8Q_0'(1/4)^2\right\} \\ &= \sigma^2\frac{\pi^2}{2} \end{aligned}$$

using $Q_0'(1/4) = \pi/(\cos(\pi/4)^2 = 2\pi$. On the other hand, the median $F^{-1}(1/2) = \mu$, and

$$\sqrt{n}(M_n - \mu) \rightarrow_d N(0, \sigma^2 Q_0'(1/2)^2 \frac{1}{4}) = N(0, \sigma^2 \frac{\pi^2}{4}),$$

so the asymptotic relative efficiency of the mid-quartile range with respect to the median for estimating the parameter μ in the Cauchy(μ, σ) family is 1/2. [As discussed in chapter 3, section 2, example 2.2, since the Cramér-Rao bound for estimation of μ in the Cauchy location family is $2\sigma^2$ (or $2\sigma^2/n$ depending on how one states the bound), the asymptotic relative efficiency of the median with respect to the asymptotically best possible estimator is $2\sigma^2/(\sigma^2(\pi^2/4)) = 8/\pi^2 \doteq .81\ldots$.]