Statistics 581, Problem Set 6 Solutions

Wellner; 11/05/2008

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.]

   Solution: Note that

$$\begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) \\ \vdots \\ \mathbb{S}_n(t_k) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) - \mathbb{S}_n(t_1) \\ \vdots \\ \mathbb{S}_n(t_k) - \mathbb{S}_n(t_{k-1}) \end{pmatrix}.$$

   where the components of the vector on the right side are independent and

$$\begin{aligned} \mathbb{S}_n(t_j) - \mathbb{S}_n(t_{j-1}) &= n^{-1/2} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i \\ &= \sqrt{\frac{[nt_j] - [nt_{j-1}]}{n}} \frac{1}{\sqrt{[nt_j] - [nt_{j-1}]}} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i \\ &\to_d \mathbb{S}(t_j) - \mathbb{S}(t_{j-1}) \sim \sqrt{t_j - t_{j-1}} N(0,1) = N(0, t_j - t_{j-1}). \end{aligned}$$

   Thus it follows that

$$\begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) - \mathbb{S}_n(t_1) \\ \vdots \\ \mathbb{S}_n(t_k) - \mathbb{S}_n(t_{k-1}) \end{pmatrix} \to_d \begin{pmatrix} \mathbb{S}(t_1) \\ \mathbb{S}(t_2) - \mathbb{S}(t_1) \\ \vdots \\ \mathbb{S}(t_k) - \mathbb{S}(t_{k-1}) \end{pmatrix}$$

   where the coordinates of the vector on the right side are independent. (This justifies the notation, since Brownian motion has independent increments.) Then the continuous mapping (or Mann-Wald) theorem yields

$$\begin{aligned} \begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) \\ \vdots \\ \mathbb{S}_n(t_k) \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) - \mathbb{S}_n(t_1) \\ \vdots \\ \mathbb{S}_n(t_k) - \mathbb{S}_n(t_{k-1}) \end{pmatrix} \\ &\to_d \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \mathbb{S}(t_1) \\ \mathbb{S}(t_2) - \mathbb{S}(t_1) \\ \vdots \\ \mathbb{S}(t_k) - \mathbb{S}(t_{k-1}) \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{S}(t_1) \\ \mathbb{S}(t_2) \\ \vdots \\ \mathbb{S}(t_k) \end{pmatrix} \sim N_k(0, (t_j \wedge t_{j'})_{j,j'=1}^k). \end{aligned}$$

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$(0, 2\mu)$.

Solution: (a) Now

$$W_n(p) \equiv \sqrt{n} \left(\begin{array}{c} \mathbb{F}_n^{-1}(p) - F^{-1}(p) \\ \mathbb{F}_n^{-1}(1-p) - F^{-1}(1-p) \end{array} \right) \rightarrow_d \left(\begin{array}{c} Q'(p)\mathbb{V}(p) \\ Q'(1-p)\mathbb{V}(1-p) \end{array} \right) \equiv W,$$

so, with $R_n(p) \equiv (1/2)(\mathbb{F}_n^{-1}(p) + \mathbb{F}_n^{-1}(1-p))$, $r(p) = (1/2)(F^{-1}(p) + F^{-1}(1-p))$, it follows that

$$\begin{aligned} \sqrt{n}(R_n(p) - r(p)) &= (1/2)\underline{1}^T W_n(p) \rightarrow_d (1/2)\underline{1}^T W(p) \\ &= (1/2)(Q'(p)\mathbb{V}(p) + Q'(1-p)\mathbb{V}(1-p)) \\ &\sim N(0, \frac{1}{4}(Q'(p)^2 p(1-p) + Q'(1-p)^2 p(1-p) + 2Q'(p)Q'(1-p)p^2)). \end{aligned}$$

When F is symmetric, $f(F^{-1}(p)) = f(F^{-1}(1-p)$, so $Q'(p) = Q'(1-p)$ and the asymptotic variance of $\sqrt{n}(R_n(p) - r(p))$ becomes $(p/2)Q'(p)^2$.
(b) For the median we have

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

(c) It follows from (a) and (b) that the asymptotic efficiency of R_n relative to the median is given by

$$ARE_{R_n, Med}(F) = \frac{(1/4)Q'(1/2)^2}{(p/2)Q'(p)^2} = \frac{1}{2p}\frac{Q'(1/2)^2}{Q'(p)^2} = \frac{1}{2p}\frac{f(F^{-1}(p))^2}{f(F^{-1}(1/2))^2}. \qquad (0.1)$$

When F is Cauchy(μ, σ) we have

$$f(x) = \frac{1}{\sigma} f_0\left(\frac{x-\mu}{\sigma}\right), \qquad F(x) = F_0\left(\frac{x-\mu}{\sigma}\right)$$

where

$$f_0(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \qquad F_0(x) = \frac{1}{2} + \frac{1}{\pi}\arctan(x).$$

Thus $F_0^{-1}(t) = \tan(\pi(t - 1/2))$, $F^{-1}(t) = \mu + \sigma F_0^{-1}(t)$, and it follows that $f(F^{-1}(t) = f_0(F_0^{-1}(t))/\sigma$. Therefore we compute $F_0^{-1}(p) = \tan(\pi(p - 1/2))$,

$F_0^{-1}(1/2) = \tan(0) = 0$, and $f_0(F_0^{-1}(p)) = 1/(2\pi)$, $f_0(F_0^{-1}(1/2)) = 1/\pi$. Thus the ARE computed in (0.1) above becomes

$$ARE_{R_n(p),Med}(Cauchy) = \frac{1}{2p} \frac{\left(\frac{1}{\sigma\pi}\frac{1}{(1+\tan^2(\pi(p-1/2)))}\right)^2}{\left(\frac{1}{\sigma\pi}\right)^2} = \frac{1}{2p(1+\tan^2(\pi(p-1/2)))^2}.$$

At the Cauchy distribution, the asymptotic variance of the median varies from 0 times the asymptotic variance of the mid-quartile range (when $p = 0$) to $1.xx$ of the asymptotic variance of the mid-p range (when $p =$); see Figure 3. It seems that for sampling from the Cauchy distribution the mid-p-quantile range has a variance which is smaller than the median for $p \in (0.387653, 1/2)$ with a maximum ARE of 1.05819 at $p = 0.443485$.

When F is Uniform$(0, 2\mu)$, $f(x) = (2\mu)^{-1}1_{[0,2\mu]}(x) = (2\mu)^{-1}f_0(x/2\mu)$ where $f_0(x) = 1_{[0,1]}(x)$. Therefore $f(F^{-1}(t)) = f_0(F_0^{-1}(t))/(2\mu) = 1/(2\mu)$ for all t. Thus the ARE computed in (0.1) above becomes

$$ARE_{R_n(p),Med}(Uniform) = \frac{1}{2p}\frac{(2\mu)^{-2}}{(2\mu)^{-2}} = \frac{1}{2p} \geq 1 \qquad \text{for all} \;\; 0 < p \leq 1/2$$

with equality at $p = 1/2$. See Figure 4.

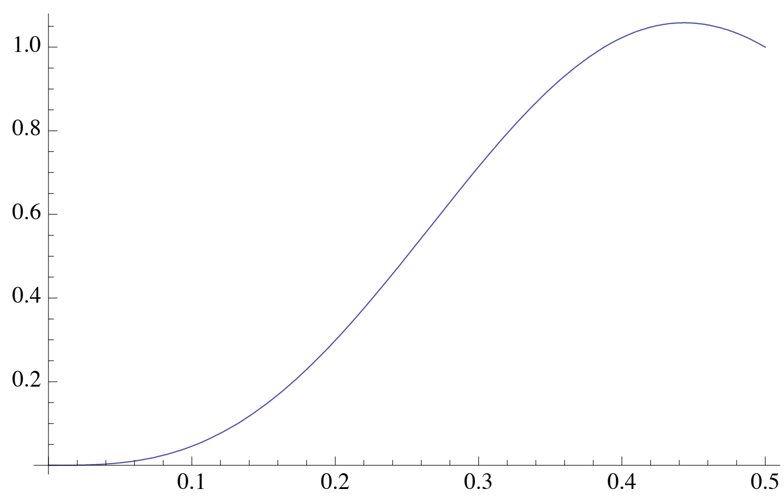

Figure 1: Asymptotic relative efficiency of $R_n(p)$ with respect to $\mathbb{F}_n^{-1}(1/2)$, $F =$Cauchy

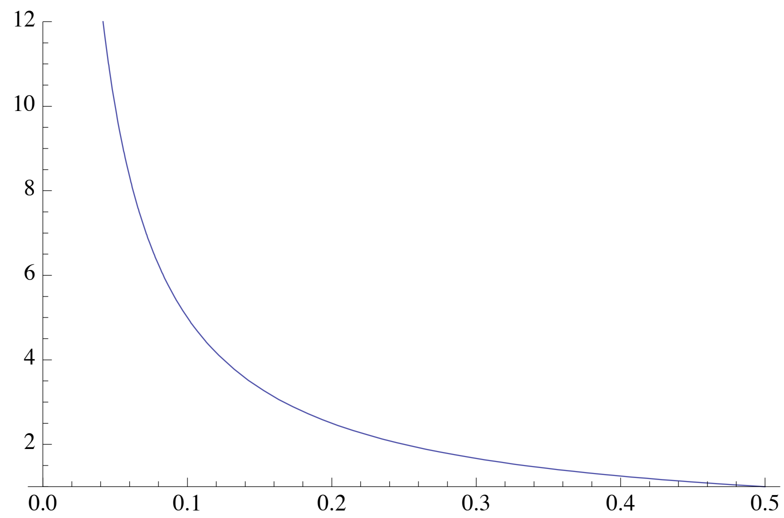

Figure 2: Asymptotic relative efficiency of $R_n(p)$ with respect to $\mathbb{F}_n^{-1}(1/2)$, $F=$Uniform

3. Suppose that $X, X_1, X_2, \ldots, X_n$ are independent Exponential(λ) random variables:

$$P(X \geq x) = \exp(-\lambda x), \qquad x > 0.$$

(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 $\{\hat{\lambda}_n^{(k)}\}_{k\geq 0}$ 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?

Solution:

(a) We compute

$$
\begin{aligned}
E(X^r) &= \int_0^\infty x^r \lambda e^{-\lambda x} dx = \lambda^{-r} \int_0^\infty (\lambda x)^r e^{-\lambda x} \lambda dx \\
&= \lambda^{-r} \int_0^\infty y^{(r+1)-1} e^{-y} dy = \lambda^{-r} \Gamma(r+1).
\end{aligned}
$$

(b) It follows from (a) that

$$
\frac{\mu_r(\lambda)}{\mu_{r+1}(\lambda)} = \frac{\lambda}{r+1}
$$

and hence

$$
\begin{aligned}
\hat{\lambda}_n^{(k)} &\equiv (k+1) \frac{\overline{X^k}_n}{\overline{X^{k+1}}_n} \equiv (k+1) \frac{n^{-1} \sum_1^n X_i^k}{n^{-1} \sum_1^n X_i^{k+1}} \\
&\to_p (k+1) \frac{\mu_k(\lambda)}{\mu_{k+1}(\lambda)} = \lambda.
\end{aligned}
$$

(c) Now by the multivariate CLT it follows that

$$
\sqrt{n} \begin{pmatrix} \overline{X^k}_n - \mu_k \\ \overline{X^{k+1}}_n - \mu_{k+1} \end{pmatrix} \to_d \underline{Z} \sim N_2(0, \Sigma)
$$

where

$$
\begin{aligned}
\Sigma &= \begin{pmatrix} \frac{\Gamma(2k+1) - \Gamma(k+1)^2}{\lambda^{2k}} & \frac{\Gamma(2k+2) - \Gamma(k+1)\Gamma(k+2)}{\lambda^{2k+1}} \\ \frac{\Gamma(2k+2) - \Gamma(k+1)\Gamma(k+2)}{\lambda^{2k+1}} & \frac{\Gamma(2k+3) - (\Gamma(k+2)^2}{\lambda^{2k+2}} \end{pmatrix} \\
&= \frac{1}{\lambda^{2k}} \begin{pmatrix} \Gamma(2k+1) - \Gamma(k+1)^2 & \frac{\Gamma(2k+2) - \Gamma(k+1)\Gamma(k+2)}{\lambda^1} \\ \frac{\Gamma(2k+2) - \Gamma(k+1)\Gamma(k+2)}{\lambda^1} & \frac{\Gamma(2k+3) - (\Gamma(k+2)^2}{\lambda^2} \end{pmatrix}.
\end{aligned}
$$

Thus by the delta method with $g(u, v) = u/v$, so that $\dot{g}(u, v) = v^{-1}(1, -u/v)$

$$
\begin{aligned}
\sqrt{n}(\hat{\lambda}_n^{(k)} - \lambda) &= (k+1)\sqrt{n}(g(\overline{X^k}_n, \overline{X^{k+1}}_n) - g(\mu_k(\lambda), \mu_{k+1}(\lambda)) \\
&\to_d (k+1)\dot{g}(\mu_k(\lambda), \mu_{k+1}(\lambda))\underline{Z} \\
&= \frac{k+1}{\mu_{k+1}(\lambda)} \left(Z_1 - \frac{\mu_k}{\mu_{k+1}} Z_2 \right) \\
&= \frac{1}{\mu_{k+1}} \left((k+1) Z_1 - \lambda Z_2 \right) \\
&\sim \frac{1}{\mu_{k+1}} N(0, \lambda^{-2k} C_k) = N(0, \frac{1}{\lambda^{2k} \mu_{k+1}^2} C_k) \\
&= N(0, \lambda^2 \frac{C_k}{\Gamma(k+2)^2}) \equiv N(0, \lambda^2 D_k)
\end{aligned}
$$

where

$$
C_k = (k+1)^2 \left\{ \Gamma(2k+1) - \Gamma(k+1)^2 \right\} - 2(k+1)\{\Gamma(2k+2) - \Gamma(k+1)\Gamma(k+2)\} + \Gamma(2k+3) - \Gamma(k+2)^2.
$$

and (after a bit of algebra)

$$D_k = \frac{\Gamma(2k+1)}{\Gamma(k+1)^2}\left\{1 - 2\frac{2k+1}{k+1} + \frac{(2k+2)(2k+1)}{(k+1)^2}\right\} = \frac{\Gamma(2k+1)}{\Gamma(k+1)^2}.$$

(c) When $k = 0$, we compute $D_k = 1$. Thus the asymptotic relative efficiency of $\hat{\lambda}_n^{(k)}$ with respect to $\hat{\lambda}_n^{(0)}$ is D_0/D_k. These estimators become inefficient relative to the mean very rapidly as k increases, as is shown by the following plot of the relative efficiency.

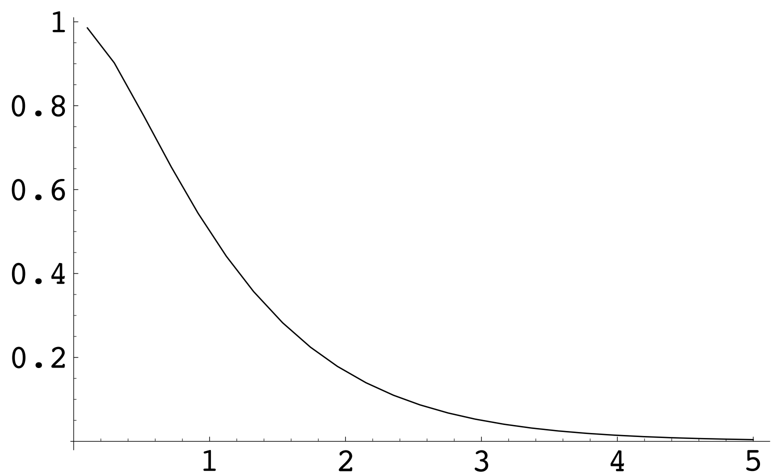

Figure 3: Asymptotic relative efficiency of $\hat{\lambda}_n^{(k)}$ with respect to $\hat{\lambda}_n^{(0)}$