Statistics 581
Problem Set 5 Solutions
Wellner; 11/2/2005

1. Suppose that $X_1, \ldots, X_n$ are independent with common mean μ, but with variances $\sigma_1^2, \ldots, \sigma_n^2$ respectively.
(a) Show that $\overline{X}_n$ is a consistent estimator of μ if $\sum_{i=1}^n \sigma_i^2 = o(n^2)$.
(b) Now suppose that $X_i = \mu + \sigma_i \epsilon_i$ where $\epsilon_1, \ldots, \epsilon_n$ are i.i.d. with some distribution function F with $E(\epsilon_1) = 0$ and $Var(\epsilon_1) = 1 < \infty$. Show that if

$$\max_{1 \le i \le n} \sigma_i^2 / \sum_{i=1}^n \sigma_i^2 \to 0 \tag{0.1}$$

then with $\overline{\sigma_n^2} \equiv n^{-1} \sum_1^n \sigma_i^2$,

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sqrt{\overline{\sigma_n^2}}} \to_d N(0,1)\,. \tag{0.2}$$

Hence show that if both (0.1) and

$$\overline{\sigma_n^2} \to \text{“something”} \equiv \sigma_0^2\,, \tag{0.3}$$

then

$$\sqrt{n}(\overline{X}_n - \mu) \to_d N(0, \sigma_0^2)\,.$$

(c) Show that (0.1) holds but that (0.3) fails if $\sigma_i^2 = Ai^r$ with $r < 1$. Hence show that in this case $n^{(1-r)/2}(\overline{X}_n - \mu) = O_p(1)$.

Solution: (a) Let $\epsilon > 0$. Note that

$$P(|\overline{X}_n - \mu| > \epsilon) \le \frac{Var(\overline{X}_n)}{\epsilon^2} = \frac{\sum_{i=1}^n \sigma_i^2}{n^2 \epsilon^2} \to 0$$

if $\sum_1^n \sigma_i^2 = o(n^2)$.
(b) For clarity, change notation by letting $\sigma_i \equiv a_i$. Set $X_{ni} \equiv X_i - \mu = a_i \epsilon_i$ for $i = 1, \ldots, n$. Then $E(X_{ni}) = 0$, $\sigma_{ni}^2 = Var(X_{ni}) = a_i^2$, and $\sigma_n^2 = \sum_{i=1}^n a_i^2$. To check the Lindeberg condition we compute

$$\begin{aligned}
\frac{1}{\sigma_n^2} \sum_{i=1}^n E|X_{ni}|^2 1_{[|X_{ni}| > \delta \sigma_n]} &= \frac{1}{\sigma_n^2} \sum_{i=1}^n E\{a_i^2 \epsilon_i^2 1_{a_i |\epsilon_i| \ge \delta \sigma_n]}\} \\
&\le E\{\epsilon_1^2 1_{[|\epsilon_1| > \delta \sigma_n / \max_{1 \le i \le n} a_i]}\} \to 0
\end{aligned}$$

if (0.1) holds. Thus (0.2) follows from the Lindeberg-Feller CLT.
(c) If $a_i^2 = Ai^r$, then

$$\sum_{i=1}^n a_i^2 = A \sum_{i=1}^n i^r \sim \frac{An^{r+1}}{r+1} \qquad \text{as } n \to \infty\,,$$

(since $n^{-1}\sum_1^n (i/n)^r \to \int_0^1 x^r dx = 1/(r+1)$). Thus

$$\frac{\max_{i\le n} a_i^2}{\sum_1^n a_i^2} = \frac{O(n^r)}{n^{r+1}} \to 0\,,$$

but

$$\frac{1}{n}\sum_{i=1}^n \sigma_i^2 \sim \frac{An^r}{r+1} \to \infty\,.$$

It follows that $n^{(1-r)/2}(\overline{X}_n - \mu) = O_p(1)$ in this case.

2. Suppose that $X_{n,i} \sim \text{Bernoulli}(p_{n,i})$, $i = 1,\ldots,n$ are independent.
(a) Show that if

$$\sum_{i=1}^n p_{n,i}(1-p_{n,i}) \to \infty\,, \tag{0.4}$$

then

$$\frac{\sqrt{n}(\overline{X}_n - \overline{p}_n)}{\sqrt{n^{-1}\sum_{i=1}^n p_{n,i}(1-p_{n,i})}} \to_d N(0,1)\,.$$

Give one example $\{p_{n,i}\}_{i\ge 1}$ for which (0.4) holds and another example for which it fails.
(b) Compare the condition (0.4) to the condition for "Poisson approximation" given by $\sum_{i=1}^n p_{n,i}^2 \to 0$ given by Le Cam's inequality (see Ferguson, ACILST, problem 5, page 18). Can both conditions hold?

Solution: With $Y_{ni} \equiv X_{n,i} - p_{n,i}$, $i = 1,\ldots,n,\ldots$ we have $E(X_{ni}) = 0$, $\sigma_{ni}^2 = Var(Y_{ni}) = p_{n,i}(1-p_{n,i})$, and

$$\begin{aligned}\gamma_{ni} &= E|Y_{ni}|^3 = E|X_{n,i}-p_{n,i}|^3 = |1-p_{n,i}|^3 p_{n,i} + |0-p_{n,i}|^3(1-p_{n,i})\\ &\le p_{n,i}(1-p_{n,i})\{(1-p_{n,i})^2 + p_{n,i}^2\} \le 2p_{n,i}(1-p_{n,i})\,,\end{aligned}$$

so that $\sigma_n^2 = \sum_{i=1}^n p_{n,i}(1-p_{n,i})$ and $\gamma_n \le 2\sum_{i=1}^n p_{n,i}(1-p_{n,i})$. hence

$$\frac{\gamma_n}{\sigma_n^3} \le \frac{2}{\{\sum_{i=1}^n p_{n,i}(1-p_{n,i})\}^{1/2}} \to 0$$

if $\sum_1^n p_{n,i}(1-p_{n,i}) \to \infty$. Hence it follows from the Liapunov CLT that

$$\frac{\sum_{i=1}^n (X_{n,i}-p_{n,i})}{\sqrt{\sum_1^n p_{n,i}(1-p_{n,i})}} \to_d N(0,1)\,,$$

and this is equivalent to the stated conclusion.
If $p_{n,i} = 1/i^r$ with $r > 1$, then the assumption fails:

$$\sum_{i=1}^n p_{n,i}(1-p_{n,i}) = \sum_{i=1}^n i^{-r} - \sum_{i=1}^n i^{-2r} \to \sum_{i=1}^\infty i^{-r} - \sum_{i=1}^\infty i^{-2r} < \infty\,.$$

On the other hand, if $p_{n,i} = 1/i$, then it holds:

$$\sum_{i=1}^n p_{n,i}(1-p_{n,i}) = \sum_{i=1}^n i^{-1} - \sum_{i=1}^n i^{-2} \to \infty - \sum_{i=1}^\infty i^{-2} = \infty\,.$$

(b) Suppose that $p_{n,i} = i^\alpha/n$, $i = 1, \ldots, n$ for some $\alpha > 0$. Then

$$\sum_{i=1}^n p_{n,i}^2 = \frac{1}{n^2}\sum_{i=1}^n i^{2\alpha} \sim \frac{1}{2\alpha+1} n^{2\alpha-1} \to 0$$

if $\alpha < 1/2$. On the other hand for $p_{n,i} = i^\alpha/n$ and $0 < \alpha < 1/2$

$$\begin{aligned}\sum_{i=1}^n p_{n,i}(1-p-n,i) &= \sum_{i=1}^n p_{n,i} - \sum_{i=1}^n p_{n,i}^2 \\ &= n^{-1}\sum_{i=1}^n i^\alpha - o(1) \\ &\sim \frac{1}{\alpha+1} n^\alpha - o(1) \to \infty,\end{aligned}$$

and thus both conditions hold for these $p_{n,i}$.

3. Ferguson, ACILST, problem 3, 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 $F^{-1}(1/4)$ and $F^{-1}(3/4)$.
(a) Find the asymptotic distribution of the mid-quartile range $R_n \equiv (X_{(3n/4)} + X_{(n/4)})/2$; i.e. find the asymptotic distribution of $\sqrt{n}(R_n - r)$ where $r = (F^{-1}(3/4) + F^{-1}(1/4))/2$.
(b) For a symmetric distribution F, $r = r(F)$ equals the center of symmetry. In this case what is the asymptotic efficiency of R_n relative to the median as an estimator of the center of symmetry?
(c) Determine the efficiency in (b) explicitly when F is the Cauchy distribution with center of symmetry μ and scale σ.

Solution: (a) Now

$$W_n \equiv \sqrt{n}\begin{pmatrix} \mathbb{F}_n^{-1}(1/4) - F^{-1}(1/4) \\ \mathbb{F}_n^{-1}(3/4) - F^{-1}(3/4)\end{pmatrix} \to_d \begin{pmatrix} Q'(1/4)\mathbb{V}(1/4) \\ Q'(3/4)\mathbb{V}(3/4)\end{pmatrix} \equiv W,$$

so, with $R_n \equiv (1/2)(\mathbb{F}_n^{-1}(1/4) + \mathbb{F}_n^{-1}(3/4))$, $r = (1/2)(F^{-1}(1/4) + F^{-1}(3/4))$, it follows that

$$\begin{aligned}\sqrt{n}(R_n - r) &= (1/2)\underline{1}^T W_n \to_d (1/2)\underline{1}^T W \\ &= (1/2)(Q'(1/4)\mathbb{V}(1/4) + Q'(3/4)\mathbb{V}(3/4)) \\ &\sim N(0, \frac{1}{4}(Q'(1/4)^2\frac{3}{16} + Q'(3/4)^2\frac{3}{16} + 2Q'(1/4)Q'(3/4)\frac{1}{16})) \\ &= N(0, \frac{1}{64}\left\{3Q'(1/4)^2 + 2Q'(1/4)Q'(3/4) + 3Q'(3/4)^2\right\}).\end{aligned}$$

When F is symmetric, $f(F^{-1}(1/4)) = f(F^{-1}(3/4)$, so $Q'(1/4) = Q'(3/4)$ and the asymptotic variance of $\sqrt{n}(R_n - r$ becomes $Q'(1/4)^2/8$.
(b) For the median we have

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

Thus the asymptotic efficiency of R_n relative to the median is given by

$$\frac{(1/4)Q'(1/2)^2}{(1/8)Q'(1/4)^2} = 2\frac{Q'(1/2)^2}{Q'(1/4)^2} = 2\frac{f(F^{-1}(1/4))^2}{f(F^{-1}(1/2)^2}.$$

(c) 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 $F_0^{-1}(1/4) = \tan(-\pi/4) = -1$, $F_0^{-1}(1/2) = \tan(0) = 0$, and $f_0(F_0^{-1}(1/4) = 1/(2\pi)$, $f_0(F_0^{-1}(1/2)) = 1/\pi$. Thus the ARE computed in (b) above becomes

$$2\frac{\left(\frac{1}{\sigma}\frac{1}{2\pi}\right)^2}{\left(\frac{1}{\sigma}\frac{1}{\pi}\right)^2} = \frac{1}{2}.$$

At the Cauchy distribution, the asymptotic variance of the median is 1/2 of the asymptotic variance of the mid-quartile range.

4. Suppose that $X_1, \ldots, X_n$ are i.i.d. random vectors with values in R^k with $E(X_1) = \mu$ and $E(X_1^T X_1) < \infty$ so that $\Sigma = E(X_1 - \mu)(X_1 - \mu)^T$ is well-defined. Thus

$$Z_n \equiv \sqrt{n}(\overline{X}_n - \mu) \rightarrow_d Z \sim N_k(0, \Sigma).$$

Suppose that $g : R^k \rightarrow R$ is a function, and suppose that $\nabla g = \dot{g}$ exists at μ. Then the delta-method (or g' theorem) tells us that

$$\sqrt{n}(g(\overline{X}_n) - g(\mu)) \rightarrow_d \nabla g(\mu)^T Z \sim N(0, \nabla g(\mu)^T \Sigma \nabla g(\mu)).$$

Show that we can strengthen this as follows: Suppose that $\nabla g = \dot{g}$ is continuous at μ. Then $\sqrt{n}(g(\overline{X}_n) - g(\mu))$ is *asymptotically linear* at μ:

$$\begin{aligned} \sqrt{n}(g(\overline{X}_n) - g(\mu)) &= \nabla g(\mu)^T \sqrt{n}(\overline{X}_n - \mu) + o_p(1) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(X_i) + o_p(1) \end{aligned}$$

where

$$\psi(x) = \nabla g(\mu)^T (x - \mu) \tag{0.5}$$

which is called the *influence function* of $g(\overline{X}_n)$ as an estimator of $g(\mu)$, has mean $E\psi(X_i) = 0$ and $Var(\psi(X_i)) = \nabla g(\mu)^T \Sigma \nabla g(\mu)$.

Solution: By Taylor's theorem, for some Y_n^* satifying $|Y_n^* - \mu| \le |\overline{X}_n - \mu| \to_p 0$ it follows that

$$\begin{aligned}
\sqrt{n}(g(\overline{X}_n) - g(\mu)) &= \nabla g(Y_n^*)\sqrt{n}(\overline{X}_n - \mu) \\
&= \nabla g(\mu)\sqrt{n}(\overline{X}_n - \mu) \\
&\qquad + \{\nabla g(Y_n^*) - \nabla g(\mu)\}\sqrt{n}(\overline{X}_n - \mu) \\
&= \nabla g(\mu)\sqrt{n}(\overline{X}_n - \mu) + o_p(1)
\end{aligned}$$

since $\nabla g(Y_n^*) \to_p \nabla g(\mu)$ by continuity of ∇g at μ and since $\sqrt{n}(\overline{X}_n - \mu) = O_p(1)$. Now note that

$$\nabla g(\mu)\sqrt{n}(\overline{X} - \mu) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \nabla g(\mu)(X_i - \mu) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi(X_i)$$

with ψ as in (0.5).

5. Suppose that $X_1, \ldots, X_n$ are i.i.d. exponential(θ); i.e. with density $p_\theta(x) = \theta e^{-\theta x} 1_{[0,\infty)}(x)$. Let $X_{(n)} = X_{n:n}$ be the largest order statistic of $X_1, \ldots, X_n$.
(a) Find constants c_n so that $Y_n = X_{(n)} - c_n \to_d Y$ for some random variable Y and find the limiting distribution F_Y.
(b) Compute the density of Y_n and show that it converges to the density f_Y of Y.
(c) What can you conclude from the result of (b) and Scheffé's theorem (chap. 2 notes, prop. 1.14, page 9).

Solution: Let $c_n = \theta^{-1}\log n$. Then

$$\begin{aligned}
F_n(y) = P(Y_n \le y) &= P(X_{(n)} - c_n \le y) = P(X_{(n)} \le y + c_n) \\
&= P(X_j \le y + c_n \text{ for all } 1 \le j \le n) \\
&= P(X_1 \le y + c_n)^n = (1 - \exp(-\theta(y + c_n)))^n \\
&= \left(1 - \frac{e^{-\theta y}}{n}\right)^n \to \exp(-e^{-\theta y}) \\
&\equiv F_Y(y);
\end{aligned}$$

this is an extreme - value distribution of the the "double-exponential" or "Gumbel" type; see part (c) of Theorem 14, Ferguson, ACILST page 95.
(b) The density of Y_n is found easily by differentiating in the previous display. The result is that

$$f_n(y) = (1 - n^{-1}\exp(-\theta y))^{n-1}\exp(-\theta y) \to \exp(-e^{-\theta y})\exp(-\theta y) = f_Y(y) \equiv (y).$$

(c) Since the densities f_n converge pointwise to the limiting density, we conclude by Scheffé's theorem that with P_n begin the probability measure on $\mathbb{R}$ corresponding to F_n and P the corresponding probability measure on $\mathbb{R}$ corresponding to F_Y,

$$d_{TV}(P_n, P) = \frac{1}{2}\int |f_n(y) - f(y)|dy \to 0.$$