Statistics 523, Problem Set 1 Solutions

Wellner; 4/10/13

1. PfS Course Notes, Exercise 10.1.5, page 228. (Empirical process, Doob) Let $\mathbb{U}_n \equiv \sqrt{n}(\mathbb{G}_n - I)$ be the uniform empirical process; here $\mathbb{G}_n(t) \equiv n^{-1}\sum_{i=1}^n 1_{[0,t]}(\xi_i)$ where ξ_i are i.i.d. Uniform$(0,1)$ random variables and $I(t) = t$ for $0 \le t \le 1$. Let $\mathbb{U}$ denote a standard Brownian bridge process on $[0,1]$. Show that $\mathbb{U}_n \to_{fd} \mathbb{U}$ as $n \to \infty$; that is, show that for any set of points $0 < t_1 < t_2 < \cdots < t_k < 1$ we have

$$(\mathbb{U}_n(t_1), \ldots, \mathbb{U}_n(t_k)) \to_d (\mathbb{U}(t_1), \ldots, \mathbb{U}(t_k)) \quad \text{as} \quad n \to \infty.$$

Solution: Note that $\mathbb{U}_n(t) = n^{-1/2}\sum_{i=1}^n (1_{[0,t]}(\xi_i) - t) \equiv n^{-1/2}\sum_{i=1}^n Y_i(t)$ where $EY_i(t) = 0$ and $Var(Y_i(t)) = t(1-t)$. Thus for any $0 < t_1 < \cdots < t_k < 1$ we have

$$\begin{pmatrix} \mathbb{U}_n(t_1) \\ \cdot \\ \cdot \\ \cdot \\ \mathbb{U}_n(t_k) \end{pmatrix} = n^{-1/2}\sum_{i=1}^n \begin{pmatrix} Y_i(t_1) \\ \cdot \\ \cdot \\ \cdot \\ Y_i(t_k) \end{pmatrix} \equiv n^{-1/2}\sum_{i=1}^n \underline{X}_i$$

where the $\underline{X}_i$'s are i.i.d. random vectors with $E\underline{X}_i = \underline{0}$ and

$$Cov(\underline{X}_i, \underline{X}_i) = E(\underline{X}_i \underline{X}_i^T) = (t_j \wedge t_{j'} - t_j t_{j'})_{j,j'=1}^k.$$

Thus by the multivariate CLT it follows that

$$n^{-1/2}\sum_{i=1}^n \sum_{i=1}^n \underline{X}_i \to_d N_k(\underline{0}, (t_j \wedge t_{j'} - t_j t_{j'})_{j,j'=1}^k).$$

Hence we conclude that

$$\begin{pmatrix} \mathbb{U}_n(t_1) \\ \cdot \\ \cdot \\ \cdot \\ \mathbb{U}_n(t_k) \end{pmatrix} \to_d \begin{pmatrix} \mathbb{U}(t_1) \\ \cdot \\ \cdot \\ \cdot \\ \mathbb{U}(t_k) \end{pmatrix} \sim N_k(\underline{0}, (t_j \wedge t_{j'} - t_j t_{j'})_{j,j'=1}^k).$$

2. PfS Course Notes, Exercise 10.1.6, page 228. (Partial sum process of i.i.d. random variables) Let $\mathbb{S}_n$ denote the partial sum process of i.i.d. $(0,1)$ random variables X_i (that is, $E(X_i)=0$ and $Var(X_i)=1$: thus $\mathbb{S}_n(t) \equiv n^{-1/2}\sum_{i=1}^{[nt]} X_i$ for $0 \le t \le 1$. Let $\mathbb{S}$ denote standard Brownian motion on $[0,1]$. Show that $\mathbb{S}_n \to_{f.d.} \mathbb{S}$.

Solution: For $0 \equiv t_0 < t_1 < \cdots < t_k \le 1$, let

$$D_{n,j} \equiv \mathbb{S}_n(t_j) - \mathbb{S}_n(t_{j-1}) = n^{-1/2} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i$$

for $j = 1, \ldots, k$, and consider the random vector $\underline{D}_n = (D_{n,1}, \ldots, D_{n,k})^T$. Now by independence of the X_i's, $D_{n1}, \ldots, D_{nk}$ are independent random variables and

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

where the Z_j's are independent $N(0,1)$ random variables. Hence

$$\underline{D}_n^T = (D_{n1}, \ldots, D_{nk})^T \to_d (\sqrt{t_1 - t_0} Z_1, \ldots, \sqrt{t_k - t_{k-1}} Z_k)^T \equiv \underline{D}.$$

Thus, adding the increments back up and applying the Mann -Wald theorem yields

$$\begin{aligned} \begin{pmatrix} \mathbb{S}_n(t_1) \\ \cdots \\ \cdots \\ \cdots \\ \mathbb{S}_n(t_k) \end{pmatrix} &= \begin{pmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 1 & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 1 & 1 & 1 & 1 & \cdot & 1 \end{pmatrix} \underline{D}_n = \begin{pmatrix} D_{n1} \\ D_{n1} + D_{n2} \\ \cdot \\ \cdot \\ \cdot \\ D_{n1} + \cdots + D_{n,k} \end{pmatrix} \\ &\to_d \begin{pmatrix} D_1 \\ D_1 + D_2 \\ \cdot \\ \cdot \\ \cdot \\ D_1 + \cdots + D_k \end{pmatrix} \\ &\sim \ N_k(\underline{0}, (t_j \wedge t_{j'})_{j,j'=1}^k). \end{aligned}$$

Thus $\mathbb{S}_n \to_{f.d.} \mathbb{S}$.

3. PfS Course Notes, Exercise 10.2.1, page 236. (Characterization of "uan") Show that the following are equivalent:
(a) $|X_{n,k}|$'s are uan; that is, $\max_{1\le k\le n} P(|X_{n,k}|\ge \epsilon)\to 0$ for all $\epsilon>0$.
(b) $\max_{1\le k\le n}|\phi_{nk}(t)-1|\to 0$ uniformly on every finite interval of t's.
(c) $\max_{1\le k\le n} E(X_{n,k}^2\wedge 1)=\max_{1\le k\le n}\int(x^2\wedge 1)dF_{nk}(x)\to 0$.

Solution: Suppose that (a) holds. We first show that something more general than (c) holds. Fix $\epsilon>0$, $r>0$ and $\tau>\epsilon$. Then we can write

$$\begin{aligned} E|X_{nk}|^r 1_{[|X_{nk}|\le\tau]} &= E|X_{nk}|^r 1_{[|X_{nk}|\le\epsilon]} + E|X_{nk}|^r 1_{[\epsilon<|X_{nk}|\le\tau]} \\ &\le \epsilon^r + \tau^r P(|X_{nk}|>\epsilon). \end{aligned}$$

Thus

$$\begin{aligned} \max_{k\le n} E|X_{nk}|^r 1_{[|X_{nk}|\le\tau]} &\le \epsilon^r + \tau^r \max_{k\le n} P(|X_{nk}|>\epsilon) \\ &\to \epsilon^r \end{aligned}$$

for every $0<\epsilon<\tau$. But since ϵ is arbitrary, this yields

$$\max_{k\le n} E|X_{nk}|^r 1_{[|X_{nk}|\le\tau]}\to 0$$

Hence for every $r>0$ and $\tau>0$

$$\max_{k\le n} E\{|X_{nk}|^r\wedge\tau^r\} = \max_{k\le n}\left\{\tau^r P(|X_{nk}|>\tau) + E\{|X_{nk}|^r 1_{[|X_{nk}|\le\tau}\}\right\}\to 0$$

as $n\to\infty$. In particular with $r=2$ and $\tau=1$ we have $\max_{k\le n} E\{|X_{nk}|^2\wedge 1\}\to 0$.

Conversely, suppose (c) holds and let $0<\epsilon<\tau$. Then

$$\begin{aligned} P(|X_{nk}|>\epsilon) &= P(|X_{nk}|>\epsilon, |X_{nk}|>\tau) + P(|X_{nk}|>\epsilon, |X_{nk}|\le\tau) \\ &\le P(|X_{nk}|>\tau) + \epsilon^{-2}E\{X_{nk}^2 1_{[|X_{nk}|\le\tau]}\} \\ &\to 0+0 \end{aligned}$$

as $n\to\infty$, so (a) holds.

Now we will show that (a) is equivalent to (b). Suppose that (a) holds. Then (b) also holds by the equivalence of (a) and (b) already proved. By Lemma 9.6.1 with $m=0$ it follows that

$$|e^{it}-1|\le 2^{1-\delta}|t|^\delta$$

for all $t\in\mathbb{R}$ and $\delta\in[0,1]$. We will use the two extreme cases $\delta=0$ and $\delta=1$. Then it follows that for each fixed $\epsilon>0$ we have

$$\begin{aligned} |\phi_{nk}(t)-1| &\le E|e^{itX_{nk}}-1| = E|e^{itX_{nk}}-1|1_{[|X_{nk}|\le\epsilon]} + E|e^{itX_{nk}}-1|1_{[|X_{nk}|>\epsilon]} \\ &\le |t|E\{|X_{nk}|1_{[|X_{nk}|\le\epsilon]}\} + 2P(|X_{nk}|>\epsilon) \end{aligned}$$

where we used the previous display with $\delta = 1$ to handle the first term and the previous display with $\delta = 0$ to handle the second term. Thus

$$\begin{aligned}\max_{k\le n} |\phi_{nk}(t) - 1| &\le |t| \max_{k\le n} E\{|X_{nk}| 1_{[|X_{nk}|\le\epsilon]}\} + 2\max_{k\le n} P(|X_{nk}| > \epsilon) \\ &\to 0 + 0 = 0\end{aligned}$$

uniformly for $|t| \le T$ by (b) and (a) respectively.

Now we show that (b) implies (a). To see this we use Inequality 9.5.1: for each $\epsilon > 0$ we have

$$\max_{k\le n} P(|X_{nk}| > \epsilon) \le 7\epsilon \max_{k\le n} \int_0^{1/\epsilon} (1 - \mathrm{Re}\phi_{nk}(t))dt \to 0$$

since (b) holds. This completes the proof of equivalence of (a), (b), and (c).

4. PfS Course Notes, Exercise 10.2.8, page 237.
(i) Show that Lindeberg's condition that $LF_n(\epsilon) \to 0$ for all $\epsilon > 0$ implies Feller's condition that $\max_{1\le k\le n} \sigma_{n,k}^2/\sigma_n^2 \to 0$.
(ii) Let $X_{n1}, \ldots, X_{nn}$ be row independent Poisson(λ/n) random variables with $\lambda > 0$. Discuss which of the Lindeberg-Feller, Liapunov, and Feller conditions holds in this context. [The Liapunov $(2 + \delta)$ condition is as follows: for some $0 < \delta \le 1$ we have
$$\sum_{k=1}^n E|X_{nk} - \mu_{nk}|^{2+\delta}/\sigma_n^{2+\delta} \to 0.]$$
(iii) Repeat part (ii) when $X_{n1}, \ldots, X_{nn}$ are row independent and all have the probability density $cx^{-3}(\log x)^2$ on $x \ge 3$ (for some constant $c > 0$).
(iv) Repeat part (ii) when $P(X_{nk} = a_k) = P(X_{nk} = -a_k) = 1/2$ for row-independent X_{nk}'s. Discuss this for general a_k's and present two or three interesting examples for which the various conditions differ (i.e. hold or fail to hold).

Solution: (i) Suppose that $LF_n(\epsilon) \to 0$ for all $\epsilon > 0$. Fix $\epsilon > 0$. Then

$$\begin{aligned}\frac{\sigma_{nk}^2}{\sigma_n^2} &= \frac{1}{\sigma_n^2}\left\{EX_{nk}^2 1_{[|X_{nk}|\le\epsilon\sigma_n]} + EX_{nk}^2 1_{[|X_{nk}|>\epsilon\sigma_n]}\right\} \\ &\le \epsilon^2 + \frac{1}{\sigma_n^2} E\{X_{nk}^2 1_{[|X_{nk}|>\epsilon\sigma_n]}\}\end{aligned}$$

and hence

$$\begin{aligned}
\max_{k\le n}\frac{\sigma_{nk}^2}{\sigma_n^2} &\le \epsilon^2+\max_{k\le n}\frac{1}{\sigma_n^2}E\{X_{nk}^2 1_{[|X_{nk}|>\epsilon\sigma_n]}\}\\
&\le \epsilon^2+\frac{1}{\sigma_n^2}\sum_{k=1}^n E\{X_{nk}^2 1_{[|X_{nk}|>\epsilon\sigma_n]}\}\\
&= \epsilon^2+LF_n(\epsilon)\to\epsilon^2
\end{aligned}$$

as $n\to\infty$. Since $\epsilon>0$ is arbitrary we conclude that

$$\max_{k\le n}\frac{\sigma_{nk}^2}{\sigma_n^2}\to 0.$$

(ii) When the X_{nk}'s are i.i.d. Poisson(λ/n), then $\mu_{nk}=\lambda/n=\sigma_{nk}^2$, so $\sum_{k=1}^n \mu_{nk}=n(\lambda/n)=\lambda$ and $\sum_{k=1}^n\sigma_{nk}^2=\lambda$. Since $\sum_{k=1}^n X_{nk}\stackrel{d}{=}N_\lambda\sim$ Poisson(λ), we have

$$Z_n\equiv\frac{1}{\sigma_n}(\sum_{k=1}^n(X_{nk}-\mu_{nk})\stackrel{d}{=}\frac{N_\lambda-\lambda}{\sqrt{\lambda}}$$

does not converge in distribution to $N(0,1)$. This implies that the Lindeberg condition fails (since if it holds then it would follow that $Z_n\to_d Z\sim N(0,1)$). It further follows that all the Liapunov-$2+\delta$ conditions fail, since they all imply that the Lindeberg condition holds. Here the Feller condition

$$\max_{k\le n}\frac{\sigma_{nk}^2}{\sigma_n^2}=\frac{\lambda/n}{\lambda}=\frac{1}{n}\to 0$$

holds.
(iii) If $f(x)=cx^{-3}(\log x)^{-2}1_{[e,\infty)}(x)$, then

$$\begin{aligned}
EX^2 &= \int_e^\infty cx^2x^{-3}(\log x)^{-2}dx=c\int_e^\infty x^{-1}(\log x)^{-2}dx\\
&= e\int_1^\infty v^{-2}dv=c<\infty
\end{aligned}$$

by the change of variable $v=\log x$. Thus $\sigma^2=Var(X_{nk})<\infty$ and $EX_{nk}=\mu$ is well-defined. Hence $\sigma_n^2=n\sigma^2$ and $\max_{k\le n}\sigma_{nk}^2/\sigma_n^2=1/n\to 0$, so the Feller condition holds. Since $\sigma^2<\infty$, we know from the ordinary CLT that

$$\frac{\sum_{k=1}^n X_{nk}-\mu_{nk}}{\sigma_n}=\sqrt{n}(\overline{X}_n-\mu)/\sigma\to_d Z\sim N(0,1).$$

Thus by the Lindeberg-Feller CLT the Lindeberg condition holds. This can also be checked directly since

$$LF_n(\epsilon) = \frac{1}{n\sigma^2} n E(X-\mu)^2 1_{[|X-\mu|>\epsilon\sigma\sqrt{n}]} = \frac{1}{\sigma^2} E(X-\mu)^2 1_{[|X-\mu|>\epsilon\sigma\sqrt{n}]} \to 0$$

for every $\epsilon > 0$ by the Dominated Convergence Theorem (with dominating function $(X-\mu)^2$). On the other hand all the Liapunov $2+\delta$ conditions fail: Note that by the C_r inequality we have, with $r = 2+\delta$,

$$|x|^r \le C_r |x-\mu|^r + |\mu|^r,$$

so that $|x-\mu|^{2+\delta} \ge |x|^{2+\delta}/C_{2+\delta} - |\mu|^{2+\delta}$. Thus

$$\begin{aligned} E|X-\mu|^{2+\delta} &\ge E|X|^{2+\delta}/C_{2+\delta} - \mu^{2+\delta} \\ &= \int_e^\infty x^{2+\delta} f(x)/C_{2+\delta} - \mu^{2+\delta} \\ &= c\int_e^\infty x^{-1+\delta}(\log x)^{-2} dx - \mu^{2+\delta} = \infty. \end{aligned}$$

5. PfS Course Notes, Exercise 10.1.10, page 230. (Special cases of Gnedenko's theorem) Let $X_{n:n}$ be the maximum of an i.i.d. sample $X_1, \ldots, X_n$ from F. Then

   (a) If $1-F(x) = e^{-x}$ for $x \ge 0$, then $P(X_{n:n} - \log n \le y) \to e^{-e^{-y}}$ for all $y \in \mathbb{R}$.

   (b) If $1-F(x) = |x|^b$ for $-1 \le x \le 0$ with $b > 0$, then $P(n^{1/b} X_{n:n} \le y) \to \exp(-|y|^b)$ for all $y < 0$.

   (c) If $1-F(x) = 1/x^a$ for $x \ge 1$ with $a > 0$, then $P(X_{n:n}/n^{1/a} \le y) \to \exp(-y^{-a})$ for all $y > 0$.

   Solution: (a) If $1-F(x) = e^{-x}$, then, using $[X_{n:n} \le t] = \cap_{i=1}^n [X_i \le t]$ and independence of the X_i's we find that

$$\begin{aligned} P(X_{n:n} - \log n \le y) &= P(X_{n:n} \le y + \log n) = P(X_1 \le y+\log n, \ldots, X_n \le y + \log n) \\ &= P(X_1 \le y + \log n)^n = \{1 - \exp(-y+\log n)\}^n \\ &= \{1 - n^{-1} e^{-y}\}^n \to \exp(-e^{-y}) \end{aligned}$$

   for all $y \in \mathbb{R}$. Thus $X_{n:n} - \log n \to_d Y \sim$ Gumbel type III, the double exponential extreme value distribution.
   (b) If $1-F(x) = |x|^b$, $-1 \le x \le 0$ with $b > 0$, then, again using $[X_{n:n} \le t] = \cap_{i=1}^n [X_i \le t]$ and independence of the X_i's we find that

$$\begin{aligned} P(n^{1/b} X_{n:n} \le y) &= P(X_{n:n} \le n^{-1/b} y) = P(X_1 \le n^{-1/b} y, \ldots, X_n \le n^{-1/b} y) \\ &= P(X_1 \le n^{-1/b} y)^n = \{1 - |n^{-1/b} y|^b\}^n \\ &= \{1 - n^{-1}|y|^b\}^n \to \exp(-|y|^b) \end{aligned}$$

for $y < 0$. Thus $n^{1/b} X_{n:n} \to_d$ Gumbel type II.
(c) If $1 - F(x) = 1/x^a$, $x \geq 1$ with $a > 0$, then, again using $[X_{n:n} \leq t] = \cap_{i=1}^{n} [X_i \leq t]$ and independence of the X_i's we find that

$$\begin{aligned} P(n^{-1/a} X_{n:n} \leq y) &= P(X_{n:n} \leq n^{1/a} y) = P(X_1 \leq n^{1/a} y, \ldots, X_n \leq n^{1/a} y) \\ &= P(X_1 \leq n^{1/a} y)^n = \{1 - (n^{1/a} y)^{-a}\}^n \\ &= \{1 - n^{-1} y^{-a}\}^n \to \exp(-y^{-a}) \end{aligned}$$

for $y > 0$. Thus $n^{-1/a} X_{n:n} \to_d$ Gumbel type I.

Fisher and Tippett (1928) and Gnedenko (1943) showed that only these three types of distributions can occur as the limiting distribution of the largest order statistic of a i.i.d. sample. This is a starting point for the theory of extreme value theory; see for example Galambos (1978) or de Haan and Fereira (2006).