Statistics 522, Problem Set 9 Solutions

Wellner; 3/8/17

1. (a) Let $X_1, X_2, \ldots,$ be i.i.d. random variables and let $Z_n \equiv n^{-1/2}\sum_{i=1}^n X_i$. For another sequence of i.i.d. random variables $X_1', X_2', \ldots,$ with each $X_i' \stackrel{d}{=} X_i$ and all X_i''s independent of the X_i's, let $X_i^s \equiv X_i - X_i'$ and set $Z_n^s \equiv n^{-1/2}\sum_{i=1}^n X_i^s$. Note that nothing has been assumed about finiteness of moments of the X_i's (or X_i's). Prove or disprove the following statement: $Z_n \rightarrow_d N(0,1)$ if and only if the symmetrized random variables $Z_n^s \rightarrow_d N(0,2)$.
(b) Now suppose that $X_1, X_2, \ldots$ are i.i.d. as in part (a), and suppose that $Z_n \equiv Z_{n,a,b} \equiv n^{1/2}(\overline{X} - a)/b$ for some $a \in \mathbb{R}$ and $b > 0$. What can you say about a and b if it is known that $Z_n \rightarrow_d N(0,1)$?

Solution: (a) Suppose that $Z_n \rightarrow_d N(0,1)$. Then it follows (from e.g. Exercise 9.1.3) that $\{Z_n\}$ is tight, and thus from the converse classical CLT (Theorem 10.7.1, PfS course notes page 265) as proved in class that $E(X_1) = 0$ and $E(X_1^2) < \infty$. Hence $E(X_1 - X_1') = 0$ and $E(X_1 - X_1')^2 = EX_1^2 + E(X_1')^2 = 2$. Then by the classical CLT we have $(Z_n, Z_n') \rightarrow_d (Z, Z')$ where Z, Z' are independent. Thus we have

$$Z_n^s \ = Z_n - Z_n' \rightarrow_d Z - Z' \sim N(0,2).$$

Now Suppose that $Z_n^s \rightarrow_d N(0,2)$. Then $\{Z_n^s\}$ is tight, and by the converse to the classical CLT we have $0 = E(X_1 - X_1')$ and $E(X_1 - X_1')^2 < \infty$. By independence of the X_i's together with their being identically distributed this implies $2EX_1^2 < \infty$ and hence $EX_1^2 < \infty$. This entails that $E|X_1| < \infty$ and hence that $\mu \equiv E(X_1) = E(X_1')$ is well-defined. This does *not* imply that $E(X_1) = 0$. But it does imply that $W_n \equiv n^{1/2}(\overline{X}_n - \mu) = n^{-1/2}\sum_{i=1}^n (X_i - \mu)$ satisfies $W_n \rightarrow_d Z \sim N(0,1)$. Thus the two statements are not quite equivalent, but they are equivalent up to a centering of Z_n.

(b) If we know that $Z_n(a,b) \rightarrow_d N(0,1)$ for some constants a, b with $b > 0$, then $\{Z_n(a,b)\}$ is tight. Since

$$Z_n(a,b) = n^{-1/2}\sum_{i=1}^n (X_i - a)/b \equiv n^{-1/2}\sum_{i=1}^n Y_i,$$

we know from the converse CLT that $E(Y_1^2) = E(X_1 - a)^2/b^2 < \infty$ and that $0 = E(Y_1) = E(X_1 - a)/b$. This yields $a = E(X_1) = \mu$, and hence $EY_1^2 = E(X_1 - \mu)^2/b^2 = \sigma^2/b^2$. But then by the classical CLT it follows that $Z_n(a, b) \to_d N(0, \sigma^2/b^2)$. But by our hypothesis $Z_n(a, b) \to_d N(0, 1)$, and thus $\sigma^2/b^2 = 1$, or $\sigma^2 = b^2$.

2. Suppose that $X_1, \ldots, X_m$ are i.i.d. with $E(X_1) = \mu$ and $Var(X_1) = \sigma^2 < \infty$; suppose that $Y_1, \ldots, Y_n$ are i.i.d. and independent of the X_i's with $E(Y_1) = \nu$ and $Var(Y_1) = \tau^2 < \infty$.
(a) Use the classical CLT (Theorem 11.2.2, W Chapter 11) to show that $\sqrt{m}(\overline{X}_m - \mu)/\sigma \to_d N(0, 1)$ as $m \to \infty$ and that $\sqrt{n}(\overline{Y}_n - \nu)/\tau \to_d N(0, 1)$ as $n \to \infty$. Do these two sequences converge jointly in distribution as $m \wedge n \to \infty$?
(b) Let $N = m + n$ and set

$$D_{m,n} \equiv \sqrt{\frac{mn}{N}} \left(\overline{Y}_n - \overline{X}_m - (\nu - \mu) \right).$$

Use (a) to show that if $\lambda_N \equiv m/N \to \lambda \in [0, 1]$ then

$$D_{m,n} \to_d \sqrt{\lambda}\tau Z - \sqrt{1-\lambda}\sigma Z' \sim N(0, \lambda\tau^2 + (1-\lambda)\sigma^2)$$

where $Z, Z' \sim N(0, 1)$ are independent.
(c) Let $S^2_{m,n} \equiv \lambda_N S^2_Y + (1 - \lambda_N) S^2_X$ where $S^2_X \equiv m^{-1} \sum_{i=1}^m (X_i - \overline{X}_m)^2$ and $S^2_Y \equiv n^{-1} \sum_{j=1}^n (Y_j - \overline{Y}_n)^2$. Show that if $\lambda_N \to \lambda \in [0, 1]$ then $S^2_{m,n} \to_p \lambda\tau^2 + (1 - \lambda)\sigma^2$.
(d) Use (b) and (c) to show that if $\lambda_N \to \lambda \in [0, 1]$, then $T_{m,n} \equiv D_{m,n}/S_{m,n} \to_d Z \sim N(0, 1)$.
(e) Use the result of (d) and the Helly selection theorem to show that $T_{m,n} \to_d Z \sim N(0, 1)$ whenever $m \wedge n \to \infty$.

Solution: (a) The classical CLT yields the claimed convergences:

$$Z_m \equiv \sqrt{m}(\overline{X}_m - \mu) \to_d \sigma Z \sim N(0, \sigma^2), \qquad \text{and}$$
$$Z'_n \equiv \sqrt{n}(\overline{Y}_n - \nu) \to_d \tau Z' \sim N(0, \tau^2).$$

By independence of the X's and Y's we also have joint convergence: the joint distribution function $F_{m,n}$ of (Z_m, Z'_n) factors, and each factor converges:

$$\begin{aligned} F_{m,n}(x,y) &\equiv P(Z_m \le x, Z_n' \le y) = P(Z_m \le x)P(Z_n' \le y) \\ &\to P(\sigma Z \le x)P(\tau Z' \le y) = P(\sigma Z \le x, \tau Z' \le y) \end{aligned}$$

for all $(x,y) \in \mathbb{R}^2$ where Z and Z' are independent $N(0,1)$ random variables. Reasoning via characteristic functions yields the same conclusion. This can be written as

$$\begin{pmatrix} \sqrt{m}(\overline{X}_m - \mu) \\ \sqrt{n}(\overline{Y}_n - \nu) \end{pmatrix} \to_d \begin{pmatrix} \sigma Z \\ \tau Z' \end{pmatrix}$$

where Z, Z' are independent $N(0,1)$ rv's.
(b) We can write

$$\begin{aligned} D_{m,n} &\equiv \sqrt{\frac{mn}{N}} \left(\overline{Y}_n - \overline{X}_m - (\nu - \mu)\right) \\ &= \sqrt{m/N}\sqrt{n}(\overline{Y}_n - \nu) - \sqrt{n/N}\sqrt{m}(\overline{X}_m - \mu) \\ &\to_d \sqrt{\lambda}\tau Z' - \sqrt{1-\lambda}\sigma Z \\ &\sim N(0, \lambda\tau^2 + (1-\lambda)\sigma^2) \end{aligned}$$

if $\lambda_N \equiv m/N \to \lambda$ by the continuous mapping theorem and the joint convergence in (a).
(c) Now $S_X^2 \to_p \sigma^2$ and $S_Y^2 \to_p \tau^2$ by the weak law of large numbers (applied to $(X_i - \mu)^2$ and $(Y_i - \nu)^2$ respectively) and the continuous mapping for convergence in probability. Therefore if $\lambda_N \to \lambda \in [0,1]$ we have

$$S_{m,n}^2 \equiv \lambda_N S_Y^2 + (1-\lambda_N)S_X^2 \to_p \lambda\sigma^2 + (1-\lambda)\tau^2$$

by continuous mapping for convergence in probability.
(d) Now (b) and (c) yield

$$T_{m,n} \equiv \frac{D_{m,n}}{S_{m,n}} \to_d \frac{\sqrt{\lambda}\tau Z' - \sqrt{1-\lambda}\sigma Z}{\sqrt{\lambda\tau^2 + (1-\lambda)\sigma^2}} \sim N(0,1)$$

by Slutsky's theorem, assuming that $\lambda_N \to \lambda \in [0,1]$.
(e) If $m \to \infty$ and $n \to \infty$ in any way, it remains true that $\lambda_N = m/N \in [0,1]$ for all m, n. Thus for any subsequences $\{m'\}$ and $\{n'\}$ there

exist further subsequences $\{m''\}$ and $\{n''\}$such that $\lambda_{N''} \equiv m''/N'' \to$ some $\lambda \in [0,1]$. For these subsequences we conclude from (a)-(d) that $T_{m'',n''} \to_d N(0,1)$. Since the limit is the same for all such initial subsequences $\{m'\}$ and $\{n'\}$, we conclude that the full sequence $T_{m,n}$ satisfies $T_{m,n} \to_d N(0,1)$.

3. Suppose that $\xi_1, \xi_2, \ldots$ are i.i.d. Uniform$[0,1]$ random variables. Let $\mathbb{G}_n(t) \equiv n^{-1}\sum_{i=1}^n 1_{[0,t]}(\xi_i)$ be the empirical distribution function of the ξ_i's, and let $\mathbb{U}_n(t) \equiv \sqrt{n}(\mathbb{G}_n(t) - t)$ for $0 \le t \le 1$. Use the multivariate CLT to show that all the finite-dimensional distributions of the sequence $\{\mathbb{U}_n : \ n \ge 1\}$ converge in distribution to the corresponding finite-dimensional distributions of a Brownian bridge process $\mathbb{U}$ on $[0,1]$; i.e. show that for any integer $k \ge 1$ and any points $0 < t_1 < \cdots < t_k < 1$

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

Solution: Suppose that $k \ge 1$ and and $0 < t_1 < \cdots < t_k < 1$. Then

$$\begin{aligned}(\mathbb{U}_n(t_1), \ldots, \mathbb{U}_n(t_k))^T &= n^{-1/2}\sum_{i=1}^n (1_{[0,t_1]}(\xi_i) - t_1, \ldots, 1_{[0,t_k]}(\xi_i) - t_k)^T \\ &= n^{-1/2}\sum_{i=1}^n \underline{Y}_i = \sqrt{n}(\overline{\underline{Y}}_n - 0)\end{aligned}$$

where $\underline{Y}_i = (1_{[0,t_1]}(\xi_i) - t_1, \ldots, 1_{[0,t_k]}(\xi_i))^T$, $i = 1, \ldots, n$, are i.i.d. random vectors in $\mathbb{R}^k$ with

$$E\{\underline{Y}_i\} = (E1_{[0,t_1]}(\xi_i) - t_1, \ldots, E1_{[0,t_k]}(\xi_i) - t_k)^T = (0, \ldots, 0)^T = \underline{0}$$

and $|1_{[0,t_j]}(\xi_1) - t_j| \le 1$ so that $E\{\underline{Y}_1^T \underline{Y}_1\} < \infty$, and

$$E\{\underline{Y}_1 \underline{Y}_1^T\} = (t_j \wedge t_{j'} - t_j t_{j'})_{j,j'=1}^k \equiv \Sigma_Y.$$

Thus by the multivariate CLT $\sqrt{n}(\overline{\underline{Y}}_n - 0) \to_d N_k(0, \Sigma_Y)$. But this is just the distribution of the vector $(\mathbb{U}(t_1), \ldots, \mathbb{U}(t_k))^T$ where $\mathbb{U}$ is a standard Brownian bridge process on $[0,1]$. Thus $(\mathbb{U}_n(t_1), \ldots, \mathbb{U}_n(t_k))^T \to_d (\mathbb{U}(t_1), \ldots, \mathbb{U}(t_k))^T$ as claimed: all the finite-dimensional distributions of $\mathbb{U}_n$ converge to the corresponding finite-dimensional distributions of $\mathbb{U}$.

4. Exercise 14.1.3, PfS, page 400-401: The partial sum process $\overline{\mathbb{S}}_n$ defined in Example 11.5.2 satisfies $\overline{\mathbb{S}}_n \to_d \mathbb{S}$ in $C[0,1]$ where $\mathbb{S}$ is a standard Brownian motion process. Consider the following four functions defined on $C[0,1]$: for $x \in C[0,1]$, let (a) $g(x) \equiv \sup_{0 \le t \le 1} x(t)$; (b) $g(x) \equiv \int_0^1 x(t)dt$; (c) $g(x) \equiv \lambda\{t \in [0,1] : x(t) > 0\}$ (where λ denotes Lebesgue measure); (d) $g(x) \equiv \inf\{t > 0 : x(t) = b\}$ with $b > 0$ fixed (where the infimum is take to be 1 if the set is empty).
For each of these real-valued functions (or "functionals" since they are real), find the discontinuity set D_g of g. Do we have $g(\mathbb{S}_n) \to_d g(\mathbb{S})$ for the g's in (a) - (d)?

Solution: (a) This function g is everywhere continuous: for $x, y \in C[0,1]$,

$$|g(y) - g(x)| = |\|y^+\|_\infty - \|x^+\|_\infty| \le \|y - x\|_\infty.$$

(b) In this case g is everywhere continuous as well: for $x, y \in C[0,1]$,

$$\begin{aligned} |g(y) - g(x)| &= \left| \int_0^1 y(t)dt - \int_0^1 x(t)dt \right| = \left| \int_0^1 (y(t) - x(t))dt \right| \\ &\le \int_0^1 |y(t) - x(t)|dt \le \|y - x\|_\infty \cdot 1 = \|y - x\|_\infty. \end{aligned}$$

(c) This functional g is not everywhere continuous: if $x(t) = 0$ for all $0 \le t \le 1$ and $y(t) = \epsilon$ for all $0 \le t \le 1$, then $\|y - x\|_\infty = \epsilon$, but $g(y) = 1$ while $g(x) = 0$. But we claim that if $C_g \equiv \{x \in C[0,1] : g \text{ is continuous at } x\}$, then $P(\mathbb{S} \in C_g^c) = 0$. A nice proof of the fact that $P(\mathbb{S} \in C_g) = 1$ is given by Durrett (2010), Example 8.6.4, page 387, and by Billingsley (1999), *Convergence of Probability Measures*, pages 246 - 247.
(d) This first passage time functional g is also not everywhere continuous on $C[0,1]$ for much the same reason as in (c): if we let $x(t) = 0$ for all t and let $y(t) = b1_{[t_0,1]}(t)$ for some (very small) b, then $\|y - x\|_\infty = b$, but $g(x) = 1$, while $g(y) = t_0$. Only slightly more complicated examples work for an arbitrary $b > 0$: take

$$x(t) \equiv x_\epsilon \equiv t_0^{-1}(b - \epsilon)t1_{[0,t_0]}(t) + (b - \epsilon)1_{[t_0,1]}(t),$$

and

$$y(t) = (bt/t_0)t1_{[0,t_0]}(t) + b1_{[t_0,1]}(t).$$

Then $\|y - x_\epsilon\|_\infty \le \epsilon$ but $g(y) - g(x_\epsilon) = t_0 - 1 \nrightarrow 0$ as $\epsilon \to 0$.

In any case, since $P(\mathbb{S} \in C_g) = 1$, the continuous mapping theorem yields $g(\mathbb{S}_n) \to_d g(\mathbb{S})$ in examples (a) - (d). The limiting distributions are as follows:
(a) For this g,

$$\begin{aligned} P(g(\mathbb{S}) > x) &= P(\sup_{0\le t\le 1} \mathbb{S}(t) > x) = P(\tau_x \le 1) \\ &= 2P(\mathbb{S}(1) > x) = 2(1 - \Phi(x)) \end{aligned}$$

where Φ is the standard normal distribution function.
(b) In this case $g(\mathbb{S}) = \int_0^1 \mathbb{S}(t)dt \sim N(0, 1/3)$ since $g(\mathbb{S})$ is Gaussian (as a linear combination of jointly normal random variables) with $E\left(\int_0^1 \mathbb{S}(t)dt\right) = \int_0^1 E(\mathbb{S}(t))dt = \int_0^1 0 \cdot dt = 0$ and

$$\begin{aligned} E\left(\int_0^1 \mathbb{S}(t)dt\right)^2 &= E\left(\int_0^1 \mathbb{S}(s)ds\right)\left(\int_0^1 \mathbb{S}(t)dt\right) \\ &= \int_0^1\int_0^1 E\left(\mathbb{S}(s)\mathbb{S}(t)\right) dsdt \quad \text{by Fubini's theorem} \\ &= \int_0^1\int_0^1 (s \wedge t)\, dsdt \\ &= 1/3. \end{aligned}$$

(c) In this case, $g(\mathbb{S}) = \int_0^1 1_{[\mathbb{S}(t)>0]}dt \sim \text{Beta}(1/2, 1/2)$ which is also known as the "arcsin distribution" because the distribution function of $g(\mathbb{S})$ is $F(u) = P(g(\mathbb{S}) \le u) = (2/\pi)\text{arcsin}(\sqrt{u})$. See Durrett (2010) for a proof via simple random walk, or see Billingsley (1999), pages 97 -100, for a proof of this via a computation of a joint distribution of this functional with other functionals of Brownian motion.
(d) In this case the "more natural" version of g is $\tilde{g}$ given by $\tilde{g}(x) \equiv \inf\{t > 0 : \ x(t) \ge b\}$ defined on $C[0, \infty)$. Then for $\mathbb{S}$ a Brownian motion defined on $[0, \infty)$,

$$\begin{aligned} P(\tilde{g}(\mathbb{S}) < t) &= P(\tau_b < t) = P(\sup_{0\le s\le t} \mathbb{S}(s) \ge b) \\ &= 2P(\mathbb{S}(t) > b) = 2(1 - \Phi(b/\sqrt{t})) \end{aligned}$$

by a reflection argument.