Statistics 523, Problem Set 3 Solutions

Wellner; 4/23/2020

1. (a) Give an example of a random variable Y with distribution function F on $\mathbb{R}^+ = [0, \infty)$ for which $EY^r = \infty$ for all $r > 0$.
   (b) Does your example in (a) have $Eg(Y) < \infty$ for some measurable function g with $g(y) \to \infty$ as $y \to \infty$?

   Solution: (a) Suppose that F is defined by

   $$1 - F(x) = \begin{cases} (\log x)^{-\gamma}, & x \ge e, \\ 1, & x < e, \end{cases}$$

   where $\gamma > 0$. Note that F has density f given by $f(x) = \gamma x^{-1} (\log x)^{-1-\gamma} 1_{[e,\infty)}(x)$. Then if $Y \sim F$,

   $$\begin{aligned} EY^r &= \int_0^\infty r x^{r-1}(1 - F(x))dx = \int_0^e r x^{r-1} dx + \int_e^\infty r x^{r-1} (\log x)^{-\gamma} dx \\ &= \infty \quad \text{for all} \quad r > 0 \end{aligned}$$

   since $\lim_{x\to\infty} x^r/(\log x)^\gamma = \infty$ for all $r, \gamma > 0$.
   (b) Consider $g(x) = (\log x)^\delta$. Then $g(x) \to \infty$ if $\delta > 0$. Moreover,

   $$\begin{aligned} Eg(Y) &= \int_e^\infty (\log x)^\delta \frac{\gamma}{x(\log x)^{\gamma+1}} dx \\ &= \gamma \int_1^\infty v^{-(1+\gamma-\delta)} dv = \frac{\gamma}{\gamma - \delta} < \infty \end{aligned}$$

   if $0 < \delta < \gamma$.

2. Stein's method for convergence in distribution to the Poisson distribution depends on the following characterization: $X \sim \text{Poisson}(\lambda)$ if and only if

   $$E[Xf(X)] = \lambda E[f(X + 1)]$$

   for all functions f for which the expectations exist. Show that if $X \sim \text{Poisson}(\lambda)$ then the identity in the display holds for any bounded function f.

Proof. Suppose that $X \sim \text{Poisson}(\lambda)$ and let $f : \mathbb{N} \to \mathbb{R}$ be bounded. Then

$$\begin{aligned} E[Xf(X)] &= \sum_{k=0}^{\infty} k f(k) e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} f(k) e^{-\lambda} \frac{\lambda^k}{(k-1)!} \\ &= \sum_{m=0}^{\infty} f(m+1) e^{-\lambda} \frac{\lambda^{m+1}}{m!} = \lambda \sum_{m=0}^{\infty} f(m+1) e^{-\lambda} \frac{\lambda^m}{m!} \\ &= \lambda E f(X+1). \end{aligned}$$

The reverse argument goes as follows: Suppose that for any $A \subset \mathbb{N}$ we can construct a function $g_{A,\lambda} : \mathbb{N} \to \mathbb{R}$ satisfying

$$\lambda g(k+1) - k g(k) = 1_A(k) - \text{Poiss}_\lambda(A) \tag{1}$$

for all $k \geq 0$. Then if W takes values in $\mathbb{N}$ we have

$$\lambda E\{\lambda g(W+1) - W g(W)\} = P(W \in A) - \text{Poiss}_\lambda(A)$$

If the left side is zero, then we conclude that W has a Poisson(λ) distribution. The solution of (1) can be found recursively, starting from $k = 0$ and working up. See Barbour, Holst, and Janson (1992), *Poisson Approximation*.

3. PfS, Exercise 10.1.5, page 228. Use the multivariate CLT to show that $\mathbb{U}_n \to_{f.d.} \mathbb{U}$ where $\mathbb{U}$ is a standard Brownian bridge process on $[0, 1]$.

   Solution: First write $\mathbb{U}_n(t) = \sqrt{n}(\mathbb{G}_n(t) - t) = n^{-1/2} \sum_{i=1}^n (1_{[0,t]}(\xi_i) - t)$. Then for $0 < t_1 < \cdots < t_k < 1$

   $$(\mathbb{U}_n(t_1), \ldots, \mathbb{U}_n(t_k))^T = n^{-1/2} \sum_{i=1}^n \underline{V}_i$$

   where $\underline{V}_i \equiv (V_i(t_1), \ldots, V_i(t_k))^T$, $i \in \{1, \ldots, n\}$ are i.i.d. random vectors with $E(\underline{V}_i) = E(1_{[0,t_1]}(\xi_1) - t_1, \ldots 1_{[0,t_k]}(\xi_1) - t_k)^T = \underline{0}$, and

   $$\begin{aligned} E\underline{V}_i \underline{V}_i^T &= E\{1_{[0,t_j]}(\xi_1) 1_{[0,t_j']}(\xi_1)\} - \underline{t}\,\underline{t}^T \\ &= (t_j \wedge t_{j'} - t_j t_{j'})_{j,j'=1}^k \equiv \Sigma. \end{aligned}$$

   Thus by the multivariate CLT it follows that

   $$(\mathbb{U}_n(t_1), \ldots, \mathbb{U}_n(t_k))^T = n^{-1/2} \sum_{i=1}^n \underline{V}_i \to_d N_k(\underline{0}, \Sigma).$$

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

4. PfS, Exercise 10.1.6, page 228. Use the multivariate CLT to show that $\mathbb{S}_n \to_{f.d.} \mathbb{S}$ where $\mathbb{S}_n$ is the partial sum process of i.i.d. $(0,1)$ rv's and $\mathbb{S}$ is a standard Brownian motion.

   Solution: Let $0 < t_1 < \cdots < t_k \le 1$. We want to show that

   $$(\mathbb{S}_n(t_1), \ldots, \mathbb{S}_n(t_k))^T \to_d N_k(\underline{0}, \{t_j \wedge t_{j'}\}_{j,j'=1}^k).$$

   To show this we first consider the vectors $\underline{V}_n \equiv (V_{n,1}, \ldots, V_{n,k})^T$ where

   $$V_{n,j} \equiv \mathbb{S}_n(t_j) - \mathbb{S}_n(t_{j-1}), \quad j \in \{1, \ldots, k\}$$

   with $\mathbb{S}_n(t_0) \equiv 0$. Note that the components of the vectors $\underline{V}_n$ are independent since $V_{n,j} = n^{-1/2} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i$ where the X_i are independent. Furthermore

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

   where the V_j are i.i.d. $N(0,1)$ by the ordinary CLT for each $1 \le j \le k$. Now note that

   $$(\mathbb{S}_n(t_1), \ldots, \mathbb{S}_n(t_k))^T = g(\underline{V}_n) \equiv (V_{n,1}, V_{n,1} + V_{n,2}, \ldots, V_{n,1} + \cdots + V_{n,k})$$

   where $g(\underline{v}) \equiv (v_1, v_1 + v_2, \ldots, v_1 + \cdots + v_k)$is a continuous map from $\mathbb{R}^k$ to $\mathbb{R}^k$. Thus by the Mann-Wald (or continuous mapping) theorem,

   $$\begin{aligned} (\mathbb{S}_n(t_1), \ldots, \mathbb{S}_n(t_k))^T &= g(\underline{V}_n) \\ &\to_d g(\underline{V}) = (V_1, V_1 + V_2, \ldots, V_1 + \cdots + V_k) \\ &\sim N_k(\underline{0}, \{t_j \wedge t_{j'}\}_{j,j'=1}^k) \stackrel{d}{=} (\mathbb{S}(t_1), \ldots, \mathbb{S}(t_k))^T. \end{aligned}$$

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

5. Goldstein's probabilistic proof of the Lindeberg-Feller CLT relies on the following lemma, which is a kind of converse for Slutsky's lemma.
   Let $\{U_n\}$ and $\{V_n\}$ be sequences of random variables such that U_n and V_n are independent for every n. Then $U_n \to_d U$ and $U_n + V_n \to_d U$ implies that $V_n \to_p 0$. Prove this lemma. (This is Lemma 5.1 in Goldstein (2009).)

   Solution: Independence of U_n and V_n for each n together with convergence in distribution of both U_n and $U_n + V_n$ to U yields

   $$\phi_U(t) = Ee^{itU} \leftarrow Ee^{it(U_n+V_n)} = Ee^{itU_n} \cdot Ee^{itV_n} \to \phi_U(t) \cdot \lim_{n\to\infty} Ee^{itV_n},$$

and hence $\phi_U(t)\lim_{n\to\infty} Ee^{itV_n} = \phi_U(t)$ for all $t \in \mathbb{R}$.

This can be turned into a proof by arguing as follows: Since ϕ_U is the characteristic function (of a proper random variable), there is a neighborhood of 0, say $|t| < \delta$, such that $\phi_U(t) \neq 0$ for all $|t| < \delta$; this follows from $\phi_U(0) = 1$ and the continuity of ϕ_U. This leads to the conclusion that the limit $\phi_V(t) = \lim_{n\to\infty} Ee^{itV_n} = 1$ for $|t| < \delta$ for some (perhaps small) $\delta > 0$. But this implies that $E(V) = 0$ and $E|V|^2 = 0$ by Durrett (2010), exercise 3.3.19. This implies $V = 0$ with probability 1, and hence $\phi_V(t) = 1$ for all $t \in \mathbb{R}$. Thus $V_n \to_d 0$ and this implies that $V_n \to_p 0$.

Alternatively, Use Exercise 3.3.20, Durrett (2010): $V_n \to_d 0$ if and only if $\phi_{V_n}(t) \to 1$ for $|t| < \delta$ for some $\delta > 0$.

Of course the point of Goldstein's proof is to (completely!) avoid the use of characteristic functions. See Goldstein (2009) pages 59-60.

Here is the statement and solution of Durrett's exercise. If $\lim_{t\searrow 0} t^{-2}(\phi(t)-1) = c > -\infty$, then $E(X) = 0$ and $E|X|^2 = -2c < \infty$. In particular, if $\phi(t) = 1 + o(t^2)$, then $\phi(t) \equiv 1$.
Solution: $E|X|^2 < \infty$ follows from Theorem 3.3.9, Durrett (2010). By comparison with $\phi(t) = 1 + it\mu - (1/2)t^2\sigma^2 + o(t^2)$ (Theorem 3.3.8, Durrett (2010)), it follows that $\mu = 0$ and $\sigma^2 = -2c$. If $\phi(t) = 1 + o(t^2)$, then $c = 0$ and $X \equiv 0$.