Solution to Homework 4 DIY Problems.

1. ⋆⋆⋆ Suppose $F_n, n \geq 1$ is a tight sequence of distribution functions. Show that their characteristic functions ψ_n are equi-continuous, i.e., for any $\epsilon > 0$, there exists $\delta > 0$, such that $|\psi_n(t) - \psi_n(s)| < \epsilon$ for all n and all $|t - s| < \delta$.

   Solution. For all t, s such that $|t - s| \leq \delta$,

$$\begin{aligned} & \sup_n |\psi_n(t) - \psi_n(s)| \leq \sup_n |E(e^{itX_n} 1_{\{|X_n| \leq M\}}) - E(e^{isX_n} 1_{\{|X_n| \leq M\}})| + 2 \sup_n P(|X_n| > M) \\ \leq \quad & M|t - s| + 2 \sup_n P(|X_n| > M) \leq M\delta + 2 \sup_n P(|X_n| > M) \to 0, \end{aligned}$$

   as $\delta \to 0$ and then $M \to 0$. □

2. ⋆⋆ The Levy concentration function is defined, for a r.v. X (or essentially its distribution) as

$$Q_X(a) = \sup_x P(X \in [x, x + a)), \qquad a \geq 0$$

   Prove that $Q_X(a/2)Q_Y(a/2) \leq Q_{X+Y}(a) \leq \min(Q_X(a), Q_Y(a))$ for two independent r.v.s. X and Y.

   Solution. By the independence of X and Y,

$$\begin{aligned} & Q_X(a/2)Q_Y(a/2) = \sup_x P(X \in [x, x + a/2)) \sup_y P(Y \in [y, y + a/2)) \\ \leq \quad & \sup_{x,y} P(X \in [x, x + a/2)),\ Y \in [y, y + a/2)) \\ \leq \quad & \sup_t P(X + Y \in [t, t + a)) = Q_{X+Y}(a); \end{aligned}$$

   and

$$\begin{aligned} & Q_{X+Y}(a) = \sup_x P(X + Y \in [x, x + a)) = \sup_x \int P(X + t \in [x, x + a)) dF_Y(t) \\ \leq \quad & \sup_y P(X \in [y, y + a)) = Q_X(a). \end{aligned}$$

   Likewise, $Q_{X+Y}(a) \leq Q_Y(a)$. □

3. ⋆⋆ The Levy metric is defined, between two distribution functions F and G, as

$$\rho(F, G) = \inf\{\epsilon > 0 : F(x - \epsilon) - \epsilon \leq G(x) \leq F(x + \epsilon) + \epsilon \quad \text{for all } x\}$$

   Show that F_n converges to F weakly if and only if $\rho(F_n, F) \to 0$.

   Solution. "$\Longleftarrow$" $\rho(F_n, F) \to 0$ ensures the existence of $\epsilon_n \downarrow 0$ such that $F(x - \epsilon_n) - \epsilon_n \leq F_n(x) \leq F(x + \epsilon_n) + \epsilon_n$ for all x. For any x that is continuity point of F,

$$\begin{aligned} & 0 \leftarrow F(x - \epsilon_n) - \epsilon_n - (F(x + \epsilon) + \epsilon) \leq F_n(x) - (F(x + \epsilon) + \epsilon) \\ \leq \quad & F_n(x) - F(x) \\ & \leq F_n(x) - (F(x - \epsilon) - \epsilon) \leq F(x + \epsilon_n) + \epsilon_n - (F(x - \epsilon_n) - \epsilon_n) \to 0. \end{aligned}$$

   "$\Longrightarrow$" Suppose $\rho(F_n, F) \nrightarrow 0$. There exist $\epsilon > 0$ and x_n such that $F_n(x_n) \notin [F(x_n - \epsilon) - \epsilon, F(x_n + \epsilon) + \epsilon]$. Without loss generality, suppose $F_n(x_n) > F(x_n + \epsilon) + \epsilon$. There exists a subsequence of x_n tending to either $-\infty$ or a constant, say x. The former implies F_n has probability mass of at least ϵ at $-\infty$, which is impossible for our definition of cdf. As continuity points are dense, there exists one in $(x, x + \epsilon)$, say y. Then, $F_n(y) > F_n(x_n) > F(y) + \epsilon$, meaning that $F_n(y) \nrightarrow F(y)$. □

4. ⋆⋆ Suppose X and Y are independent, and $X+Y$ and X have same distribution. Show that $Y=0$ with probability 1.

Solution. $\psi_{X+Y}(t)=\psi_X(t)\psi_Y(t)=\psi_X(t)$. Hence $\psi_Y(t)=1$ for t in a neighborhood of 0, implying $E(cos(tY))=1$ for t in a neighborhood of 0. Hence $P(Y=0)=1$. □

5. ⋆⋆ Suppose X_n are independent with mean 0 and $E(|X_n|^\delta)\le 1$ for some $\delta>2$. Raise a counter-example to show $S_n/s_n \nrightarrow N(0,1)$, where $S_n=X_1+\cdots X_n$ and $s_n^2=\mathrm{var}(S_n)$.

Solution. Let $P(X_n=1)=4^{-n}=P(X_n=-1)$ and $P(X_n=0)=1-2/4^n$. Then,

$$P(S_n=0)\ge 1-\sum_{i=1}^n P(X_i\ne 0)=1-\sum_{i=1}^n 2/4^i\ge 1-2/3=1/3.$$

Then S_n/s_n does not converge to $N(0,1)$. □

6. ⋆⋆ Suppose $X_1,...,X_n$ are iid with mean 0 and finite positive variance. Use central limit theorem and Kolmogorov's 0-1 law to show $\limsup S_n/\sqrt{n}=\infty$ a.s., where $S_n=X_1+\cdots+X_n$.

Solution. By Kolmogorov's 0-1 law, $\limsup S_n/\sqrt{n}$ must be c, which is either a constant or ∞ or $-\infty$. By the CLT, for any constant a,

$$P(\limsup_n \frac{S_n}{\sqrt{n}}>a)=\lim_n P(\sup_{k\ge n}\frac{S_k}{\sqrt{k}\sigma}>\frac{a}{\sigma})\ge \lim_n P(\frac{S_n}{\sqrt{n}\sigma}>\frac{a}{\sigma})\to P(N(0,1)>\frac{a}{\sigma})>0.$$

Therefore, $\limsup S_n/\sqrt{n}=\infty$ with probability 1. □

7. ⋆⋆⋆ For positive values of α, if any, does CLT hold for iid symmetric r.v.s with distribution $F(x)=1-1/(2x^\alpha)$ for $x>1$ and $F(x)=1/2$ for $x\in[0,1]$?

Solution. For $x>1$, $P(|X|>x)=x^{-\alpha}$ and

$$E(|X|^2 1_{\{|X|\le x\}\}})=2\int_1^x t^2 dF(t)=\begin{cases}2\log x & \alpha=2\\ \frac{\alpha}{2-\alpha}(x^{-\alpha+2}-1) & \alpha\ne 2\end{cases}$$

Therefore, as $x\to\infty$.

$$\frac{x^2P(|X|>x)}{E(|X|^2 1_{\{|X|\le x\}\}}))}\to\begin{cases}0 & \text{for } \alpha\ge 2\\ \alpha/(2-\alpha) & \text{for } \alpha<2\end{cases}$$

It follows from Theorem 2.3 that when $\alpha\ge 2$, CLT holds. □

8. ⋆ Show Lindeberg condition implies $\max_{1\le i\le n}\sigma_i/s_n\to 0$, which is equivalent to $\sigma_n/s_n\to 0$ where $\sigma_i>0$ and $s_n^2=\sigma_1^2+\cdots+\sigma_n^2$.

Solution. For any $\epsilon>0$,

$$\begin{aligned}\frac{1}{s_n^2}\max_{1\le i\le n}\sigma_i^2 &= \frac{1}{s_n^2}\max_{1\le i\le n}E(X_i^2 1_{\{|X_i|\le\epsilon s_n\}}+\frac{1}{s_n^2}\max_{1\le i\le n}E(X_i^2 1_{\{|X_i|>\epsilon s_n\}}\\ \le\ & \epsilon^2+\frac{1}{s_n^2}\sum_{i=1}^n E(X_i^2 1_{\{|X_i|>\epsilon s_n\}}\to 0, \quad \text{as } n\to\infty.\end{aligned}$$

Since $\epsilon>0$ is arbitrary, we have $\max_{1\le i\le n}\sigma_i/s_n\to 0$.

Obviously, $\max_{1\le i\le n}\sigma_i/s_n\to 0$ implies $\sigma_n/s_n\to 0$. On the other hand, if $\sigma_n/s_n\to 0$, for any integer M,

$$\max_{1\le i\le n}\sigma_i/s_n\le\max_{1\le i\le M}\sigma_i/s_n+\max_{M<i\le n}\sigma_i/s_i\to\max_{M<i}\sigma_i/s_i, \quad \text{as } n\to\infty.$$

Let $M\to\infty$, we have $\max_{1\le i\le n}\sigma_i/s_n$. □

9. ⋆⋆ Suppose X_n are iid with mean μ and finite variance. Then,

$$\frac{\sqrt{n}(\bar{X}-\mu)}{\sqrt{1/(n-1)\sum_{i=1}^{n}(X_i-\bar{X})^2}} \to N(0,1)$$

where $\bar{X}=(1/n)\sum_{i=1}^{n}X_i$. In particular, t-distribution with n degree of freedom tends to standard normal distribution as $n\to\infty$.

Solution. The classical central limit theorem implies $\sqrt{n}(\bar{X}-\mu)/\sigma \to N(0,1)$. Law of large numbers ensures $1/(n-1)\sum_{i=1}^{n}(X_i-\bar{X})^2 \to E(X-E(X))^2=\sigma^2$ a.s.. Then, the convergence follows from Slutsky's theorem. The above statistic follows t-distribution with degree of freedom $n-1$ when X_i follows normal distribution. □

10. ⋆⋆⋆ Raise an example of independent random variables X_n with mean 0 and finite positive variance σ_n^2 such that $S_n/s_n \to N(0,1)$, but the Lindeberg condition does not hold, where $S_n=X_1+\cdots+X_n$ and $s_n^2=\sigma_1^2+\cdots+\sigma_n^2$.

Solution. X_n are independent $\sim N(0,2^n)$, which is the same as $\sqrt{2^n}N(0,1)$. Then, $s_n^2=2+2^2+\cdots+2^n=2^{n+1}-2$. And, for $n\geq 2$,

$$E(X_n^2 1_{\{|X_n|>\epsilon s_n\}}) = 2^n E(Z^2 1_{\{|Z|>\epsilon\sqrt{2-2^{1-n}}\}}) \geq .5 s_n^2 E(Z^2 1_{\{|Z|>\epsilon\sqrt{2}\}}) \geq .5 s_n^2 E(Z^2 1_{\{|Z|>\epsilon\sqrt{2}\}}),$$

where Z is a standard normal r.v.. Then, Lindeberg condition does not hold. But S_n/s_n follows exactly $N(0,1)$. □