SDS 387, Fall 2024
Homework 1

Due September 17, by midnight on Canvas.

1. Let $\{x_n\}$ be a sequence of numbers. Describe the mathematical statements: $x_n = \Omega(1)$, $x_n = \omega(1)$ and $x_n = \Theta(1)$.

   $x_n = \Omega(1)$ is equivalent to the statement that $\inf_n |x_n| \geq C$ for some $C > 0$. $x_n = \omega(1)$ is equivalent to the statement that, for any $M > 0$ (arbitrarily large) there exists a $N \in \mathbb{N}$ (which depends on M) such that $|x_n| \geq M$ for all $n \geq N$. $x_n = \Theta(1)$ is equivalent to the statement that there exists a $C \geq 1$ such that $\frac{1}{C} \leq |x_n| \leq C$.

2. **Limit superior and limit inferior.**

   (a) Let $\{A_n\}$ be a sequence of events (an event is a collection of outcomes). Argue that an outcome belongs to $\limsup_n A_n$ if and only if it belongs to infinitely many events A_n's and that it belongs to $\liminf_n A_n$ if and only if there exists an integer N such that the outcome belongs to all the events A_n with $n \geq N$. Conclude that $\liminf_n A_n \subseteq \limsup_n A_n$.
   Recalling the definition $\limsup_n A_n = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m$, if a point x belongs to $\limsup_n A_n$ then, for every n, it belongs to the set $\cup_{m=n}^{\infty} A_m$. Equivalently, for every n, there exists a $k \geq n$ such that $x \in A_k$. That is x belongs to infinitely many events A_n's. Similarly, since $\liminf_n A_n = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m$, if x belongs to $\liminf_n A_n$, there exists a N such that x belongs to each A_m with $m \geq N$.

   (b) Consider the same setting above. De Morgan's Laws state that $(\cup_n A)^c = \cap_n A_n^c$ and $(\cap_n A)^c = \cup_n A_n^c$, where A^c is the complement of the set A. Use De Morgan's law to show that $(\liminf_n A_n)^c = \limsup_n A_n^c$.
   This follows directly from DeMorgan's Law.

   (c) Let A_n be $(-1/n, 1]$ if n is odd and $(-1, 1/n]$ if n is even. Find $\limsup_n A_n$ and $\liminf_n A_n$.
   Note that for any $k \in \mathbb{N}$, $A_k \cup A_{k+1} = (-1, 1]$. Hence $\bigcup_{k=n}^{\infty} A_k = (-1, 1]$ for all $n \in \mathbb{N}$, and hence

   $$\limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \bigcap_{n=1}^{\infty} (-1, 1] = (-1, 1].$$

   Also, note that for any $m \in \mathbb{N}$, $\bigcap_{k=m}^{\infty} A_{2k-1} = [0, 1]$ and $\bigcap_{k=m}^{\infty} A_{2k} = (-1, 0]$. Hence $\bigcap_{k=n}^{\infty} A_k = \{0\}$ for any $n \in \mathbb{N}$, and hence

   $$\liminf_n A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \bigcup_{n=1}^{\infty} \{0\} = \{0\}.$$

(d) **Bonus Problem**. Let A_n the interior of the ball in $\mathbb{R}^2$ with unit radius and center $\left(\frac{(-1)^n}{n}, 0\right)$. Find $\limsup_n A_n$ and $\liminf_n A_n$.

Let $D := \{x \in \mathbb{R}^2 : \|x\|_2 < 1\}$ and $B := \{x = (x_1, x_2) \in \mathbb{R}^2 : \|x\|_2 = 1,\, x_1 \neq 0\}$. We will show that $\liminf_n A_n = D$ and $\limsup_n A_n = D \cup B$.

For $\liminf_n A_n$, note that $x \in \liminf_n A_n$ if and only if $x \in A_n$ for all but finite n. Suppose $x \in D$. Then $\|x\|_2 < 1$, so choose N large enough so that $\frac{1}{N} < 1 - \|x\|_2$. Then for all $n \geq N$,

$$\begin{aligned}\left\|x - \left(\frac{(-1)^n}{n}, 0\right)\right\|_2 &\leq \|x\|_2 + \left\|\left(\frac{(-1)^n}{n}, 0\right)\right\|_2 \\ &= \|x\|_2 + \frac{1}{n} \leq \|x\|_2 + \frac{1}{N} < 1.\end{aligned}$$

Then $x \in A_n$ for all $n \geq N$, and hence $x \in \liminf_n A_n$, which implies $D \subset \liminf_n A_n$. Now, suppose $x \notin D$ and $x_1 \geq 0$. Then for all odd n,

$$\left\|x - \left(\frac{(-1)^n}{n},\, 0\right)\right\|_2 = \left\|\left(x_1 - \frac{1}{n},\, x_2\right)\right\|_2 > \|(x_1,\, x_2)\|_2 \geq 1,$$

Hence $x \notin A_n$ for all odd n, and hence $x \notin \liminf_n A_n$. Similarly, when $x \notin D$ and $x_1 \leq 0$, then $x \notin A_n$ for all even n, and hence $x \notin \liminf_n A_n$. These imply $\liminf_n A_n \subset D$, and hence

$$\liminf_n A_n = D.$$

For $\limsup_n A_n$, note that $x \in \limsup_n A_n$ if and only if $x \in A_n$ for infinitely many n. Suppose $x \in D \cup B$. We have already shown that $D = \liminf_n A_n \subset \limsup_n A_n$, and hence if $x \in D$ then $x \in \limsup_n A_n$. Now, suppose $x \in B$ and $x_1 > 0$. Then $\|x\|_1 = 1$. Choose N large enough so that $\frac{1}{N} < |x_1|$. Then for all even n with $n \geq N$, $\left|x_1 - \frac{1}{n}\right| \leq |x_1|$, and hence

$$\begin{aligned}\left\|x - \left(\frac{(-1)^n}{n}, 0\right)\right\|_2 &= \left\|\left(x_1 - \frac{1}{n},\, x_2\right)\right\|_2 \\ &< \left\|(x_1,\, x_2)\right\|_2 = 1.\end{aligned}$$

Hence $x \in A_n$ for all even n with $n \geq N$, and hence $x \in \limsup_n A_n$. Similarly, when $x \in B$ and $x_1 < 0$, $x \in A_n$ for all odd n with $n \geq N$, and hence $x \in \limsup_n A_n$. These imply that $D \cup B \subset \limsup_n A_n$. Now, suppose $x \notin D \cup B$. Then $\|x\|_2 > 1$ or $x = (0, \pm 1)$. When $\|x\|_2 > 1$, choose N large enough so that $\frac{1}{N} < 1 - \|x\|_2$. Then for all $n \geq N$,

$$\begin{aligned}\left\|x - \left(\frac{(-1)^n}{n}, 0\right)\right\|_2 &\geq \|x\|_2 - \left\|\left(\frac{(-1)^n}{n}, 0\right)\right\|_2 \\ &= \|x\|_2 - \frac{1}{n} \geq \|x\|_2 - \frac{1}{N} > 1.\end{aligned}$$

Then $x \notin A_n$ for all n with $n \geq N$, and hence $x \notin \limsup_n A_n$. Also, when $x = (0, \pm 1)$, then for all n,

$$\left\| x - \left(\frac{(-1)^n}{n}, 0 \right) \right\|_2 = \left\| \left(-\frac{(-1)^n}{n}, \pm 1 \right) \right\|_2 = \sqrt{1 + \frac{1}{n^2}} > 1,$$

Then $x \notin A_n$ for all n, and hence $x \notin \limsup_n A_n$. These show $\limsup_n A_n \subset D \cup B$, and hence

$$\limsup_n A_n = D \cup B.$$

3. Let $X_1, X_2, \ldots$ be a sequence of 0-1 Bernoulli random variables such $X_n \sim \text{Bernoulli}(1/n^2)$. Let $X = \sum_{n=1}^{\infty} X_n$. What is $\mathbb{P}(X < \infty)$?
Use Borel Cantelli's first Lemma. Define the events $A_n = \{X_n = 1\}$, $n = 1, 2, \ldots$. Then, $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} 1/n^2 = \frac{\pi^2}{6} < \infty$. So $\mathbb{P}(\limsup A_n) = 0$. Now, $\limsup A_n = \mathbb{P}(X = \infty)$.

4. Ferguson, problem 5, page 12.
$X_n \xrightarrow{p} 0$ for all values of α. By Borel-Cantelli's Second Lemma, if $\alpha \geq 1$ then, for any $\epsilon > 0$ $|X_n| > \epsilon$ infinitely often with probability 1, by independence and because $\sum_n \frac{1}{n} \sim \log n \to \infty$. On the other hand, when $\alpha < 1$, Borel-Cantelli's First Lemma will imply that the probability that $|X_n| > \epsilon$ infinitely often is equal to 0 for any ϵ. Therefore, $X_n \overset{w.p.\,1}{\to} 0$ if and only if $\alpha < 1$. Finally, by direct calculation

$$\mathbb{E}[|X_n|^p] = \frac{n^{\alpha p}}{n} \to 0$$

if and only if $\alpha\, p < 1$.

5. Prova Markov's inequality: if X is a non-negative random variable, then for any $\epsilon > 0$

$$\mathbb{P}(X \geq \epsilon) \leq \frac{\mathbb{E}[X]}{\epsilon}.$$

Markov's inequality is almost always a loose upper bound, but there are rare cases when it is sharp. Find an example in which it holds exactly. *Hint: take X to be the indicator function of a set and select the right ϵ.*
Prove the PaleyZygmund inequality, a reverse Markov inequality of sort: if X is a non-negative random variable with two or more moments, then, for any $\alpha \in (0, 1)$,

$$\mathbb{P}(X \geq \alpha \mathbb{E}[X]) \geq (1 - \alpha)^2 \frac{\mathbb{E}[^2 X]}{\mathbb{E}[X^2]}.$$

We can write

$$X = X\mathbb{1}\{X < \theta\mathbb{E}[X]\} + X\mathbb{1}\{X \geq \theta\mathbb{E}[X]\}$$
$$\overset{(i)}{\leq} \theta\mathbb{E}[X] + \sqrt{\mathbb{E}[X^2]\mathbb{P}(X \geq \theta E[X])},$$

where in *(i)* we have used Cauchy-Schwartz inequality to bound the second term. So,

$$\mathbb{E}[X](1-\theta) \leq \sqrt{\mathbb{E}[X^2]\mathbb{P}(X \geq \theta E[X])}.$$

The result follow from taking the square.

6. Let $X_1, \ldots, X_n$ *i.i.d.* univariate random variables with common distribution function F_X. Given $\alpha \in (0,1)$, use the DKW inequality given in class to construct a $1-\alpha$ confidence band for F_X, a pair of random functions (random because dependent on $X_1, \ldots, X_n$), say $\hat{F}_\alpha^{\text{lower}}$ and $\hat{F}_\alpha^{\text{upper}}$, such that

$$\mathbb{P}\left(\hat{F}_\alpha^{\text{lower}}(x) \leq F_X(x) \leq \hat{F}_\alpha^{\text{upper}}(x), \forall x \in \mathbb{R}\right) \geq 1-\alpha.$$

The DKW inequality states that

$$\mathbb{P}\left(\sup_x |F_X(x) - \hat{F}_n(x)| > \epsilon\right) \leq 2\exp\{-2n\epsilon^2\}, \quad \forall \epsilon > 0.$$

Set the right hand side of the above inequality to α and solve for ϵ to conclude that

$$\hat{F}_\alpha^{\text{lower}}(x) = \min\left\{0, \hat{F}_n(x) - \sqrt{\frac{\log 2/\alpha}{2n}}\right\}, \quad x \in \mathbb{R}$$

and

$$\hat{F}_\alpha^{\text{upper}}(x) = \max\left\{1, \hat{F}_n(x) + \sqrt{\frac{\log 2/\alpha}{2n}}\right\}, \quad x \in \mathbb{R}.$$

7. **Joint and marginal convergence.** Below, $\{X_n\}$ is a sequence of random vectors in $\mathbb{R}^d$ and X another random vector in $\mathbb{R}^d$.

   (a) Show that $X_n \xrightarrow{p} X$ if and only if $X_n(j) \xrightarrow{p} X(j)$ for all $j = 1, \ldots, d$. *Note: the same is true about convergence with probability one.*

   By definition, $X_n \xrightarrow{p} X$ if and only if, for each $\epsilon > 0$,

$$\mathbb{P}(\|X_n - X\| \geq \epsilon) \to 0,$$

   or, equivalently,

$$\mathbb{P}(\|X_n - X\| < \epsilon) \to 1,$$

which implies, since $\max_j |X_n(j) - X(j)| \leq \|X_n - X\|$, that $\mathbb{P}(\max_j |X_n(j) - X(j)| < \epsilon) \to 1$. In turn, this implies that, for any j, $X_n(j) \xrightarrow{p} X(j)$. Conversely, if, for any $\epsilon > 0$, $\mathbb{P}(|X_n(j) - X(j)| \geq \epsilon) \to 0$ for all j, then $\mathbb{P}(\|X_n - X\| \geq d\epsilon) \leq \sum_{j=1}^d \mathbb{P}(|X_n(j) - X(j)| \geq \epsilon) \to 0$. Since ϵ is arbitrary, $X_n \xrightarrow{p} X$.

(b) Show that if $X_n \xrightarrow{d} X$, then $X_n(j) \xrightarrow{d} X(j)$ for all $j = 1, \ldots, d$.
There is more than one way to prove this. One could use the definition of convergence in distribution and a limiting argument. A shorter way is to use characteristic functions. Let ϕ_{X_n}, ϕ_X, $\phi_{X_n(j)}$, $\phi_{X(j)}$ be the characteristics functions of X_n, X, $X_n(j)$ and $X(j)$ respectively. Then, if $X_n \xrightarrow{d} X$, $\phi_{X_n}(t) \to \phi_X(t)$ for all $t \in \mathbb{R}^d$. In particular, this is true for the vector $t_x = (t_1, \ldots, t_d)$ such that $t_i = 0$ if $i \neq j$ and $t_j = x$ if $i = j$, where x is any real number. As a result, we get that, for any $x \in \mathbb{R}$,

$$\phi_{X_n(j)}(x) = \phi_{X_n}(t_x) \to \phi_X(t_x) = \phi_{X(j)}(x)$$

(c) In class, we looked at this example in $d = 2$. Set $U \sim \text{Uniform}(0, 1)$ and let $X_n = U$ for all n and

$$Y_n = \begin{cases} U & n \text{ odd}, \\ 1 - U & n \text{ even}. \end{cases}$$

Then, $X_n \xrightarrow{d} U$ and $X_n \xrightarrow{d} U$. In class, I claimed that

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix}$$

does not converge in distribution (in fact, in any meaningful sense). Prove the claim.
One way to prove the claim is to show that the random vector $\begin{bmatrix} X_n \\ Y_n \end{bmatrix}$ take values on the line segment on the plane joining $(0, 0)$ to $(1, 1)$ for all odd n and on the line segment joining $(1, 0)$ to $(0, 1)$ for all odd n. A simpler way is to use the Cramer-Wald device: the characteristic function of the vector evaluated at the point $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is equal to e^i when n is even and to the characteristic function of $2U$ at 1 when n is odd. As a result, it does not converge.

8. Show that the c.d.f. of a random variable can have at most countably many points of discontinuity.
For every point of discontinuity, say x, of the c.d.f. F, the interval $(F(x-), F(x))$ is not empty, by definition. Take any rational number in this interval. Thus, for every point of discontinuity of F, we can find a distinct rational. Since the set of rationals is countable, the set of discontinuity points of F can be put in a one-to-one correspondence with a subset of the set of rational numbers, which is countable.

9. For each n, let X_n a random variable uniformly distributed on $\left\{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right\}$. Show that X_n converges on distribution to $U \sim \text{Uniform}(0,1)$. Let A be the set of all rational numbers in $[0,1]$. Then $\mathbb{P}(X_n \in A) = 1$ for all n but $\mathbb{P}(X \in A) = 0$. Show that this does not violate condition *(v)* of the Portmanteau theorem, as stated in the lecture notes.
   X_n converges on distribution to $U \sim \text{Uniform}(0,1)$ because the c.d.f of X_n is

$$\mathbb{F}_{X_n}(x) = \begin{cases} 0 & x < 0 \\ \frac{\lfloor nx \rfloor}{n} & x \in [0,1] \\ 1 & x > 1 \end{cases} \to F_U(x) = \begin{cases} 0 & x < 0 \\ x & x \in [0,1] \\ 1 & x > 1 \end{cases}$$

   for all $x \in \mathbb{R}$. In this example, condition *(v)* of the Portmanteau theorem is not violated because A is dense in $[0,1]$, so $\partial A = [0,1] \setminus \mathbb{Q}$. Therefore, since $\mathbb{P}(U \in \mathbb{Q}) = 0$,

$$\mathbb{P}(U \in \partial A) = \mathbb{P}(U \in [0,1] \setminus \mathbb{Q}) = \mathbb{P}(U \in [0,1]) = 1 \neq 0.$$