Name: Jenny Petrova

Problem Set 10

Due Date: 11/14 at 11:35 AM

(P.1) *Let $(a_n)_{n\geq 1}$ be a sequence of real numbers, $a \in \mathbb{R}$ a real number, and let $f : \mathbb{R} \mapsto \mathbb{R}$ be continuous. Show that if $a_n \to a$ as $n \to \infty$, then $f(a_n) \to f(a)$ as $n \to \infty$.*

Proof. By definition of a continuous function, since f is continuous at all $a \in \mathbb{R}$, then for any $\epsilon > 0$ there exists a $\delta > 0$ such that for all $x \in \mathbb{R}$,

$$|x - y| < \delta \text{ implies } |f(x) - f(y)| < \epsilon.$$

By definition of convergence, if $a_n \to a$, then for each $\delta > 0$ there exists a number $N \in \mathbb{N}$ such that for all $n \geq N$,

$$|a_n - a| < \delta.$$

This implies there exists a $\epsilon > 0$ such that $|f(a_n) - f(a)| < \epsilon$. Therefore $f(a_n) \to f(a)$ as $n \to \infty$. □

(P.2) *Let $(X_n)_{n\geq 1}$ be a sequence of random variables and X a random variable, all defined on a common measurable space and taking values in $\mathbb{R}$. Show that if $f : \mathbb{R} \mapsto \mathbb{R}$ is continuous, then $X_n \xrightarrow{P} X$ implies $f(X_n) \xrightarrow{P} f(X)$.*

Proof. Let $\epsilon > 0$. For each $k > 0$, let

$$\{|f(X_n) - f(X)| > \epsilon\} \subset \{|f(X_n) - f(X)| > \epsilon, |X| \leq k\} \cup \{|X| > k\}.$$

Since f is continuous, then by definition it is continuous on any bounded interval. Then for any ϵ there exists a $\delta > 0$ such that $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$ for $x, y \in [-k, k]$. Therefore

$$\{|f(X_n) - f(X)| > \epsilon, |X| \leq k\} \subset \{|X_n - X| > \delta, |X| \leq k\} \subset \{|X_n - X| > \delta\},$$

and we combine with the first equation to get

$$\{|f(X_n) - f(X)| > \epsilon\} \subset \{|X_n - X| > \delta\} \cup \{|X| > k\}.$$

By the subadditivity of probability, we get

$$\mathbb{P}(|f(X_n) - f(X)| > \epsilon) \leq \mathbb{P}(|X_n - X| > \delta) + \mathbb{P}(|X| > k).$$

Note that $\mathbb{P}(|X| > k) \to 0$ as $k \to \infty$ since X is a random variable with finite tail probability. Therefore for any $\gamma > 0$ we can pick a fixed k large enough such that $\mathbb{P}(|X| > k) < \gamma$. Therefore

$$\lim_{k\to\infty} \mathbb{P}(|f(X_n) - f(X)| > \epsilon) \leq \lim_{k\to\infty} \mathbb{P}(|X_n - X| > \delta) + \gamma = \gamma.$$

Because $\gamma > 0$ is arbitrary, we can conclude

$$\lim_{k\to\infty} \mathbb{P}(|f(X_n) - f(X)| > \epsilon) = 0.$$

Therefore, by definition of convergence in probability, $f(X_n) \xrightarrow{P} f(X)$.

□

(P.3) *Show that if $f : \mathbb{R} \mapsto \mathbb{R}$ is not a continuous function, then $X_n \xrightarrow{P} X$ does not in general imply $f(X_n) \xrightarrow{P} f(X)$. Hint: Try to find a counter example using jump discontinuities.*

Proof. Let $X \sim \mathcal{N}(0, \frac{1}{n})$ be a normally distributed random variable ($n \in \mathbb{Z}^+$). We know the central moment with order 2 is $\mathbb{E}[X^2] = \sigma^2 = \frac{1}{n}$. Then using Chebyshev's Inequality we show

$$\mathbb{P}(|X_n| > \epsilon) \leq \frac{\mathbb{E}[X^2]}{\epsilon^2} = \frac{1}{n\epsilon^2}.$$

Thus

$$\mathbb{P}(|X_n - 0| > \epsilon) = \mathbb{P}(|X_n| > \epsilon) \to 0 \text{ as } n \to \infty$$

for all $\epsilon > 0$. Therefore $X_n \xrightarrow{P} 0$.
Now suppose $f(x) = \mathbb{1}_{\{0\}}(x)$. Then

$$\mathbb{P}(|f(X_n) - f(0)| > \epsilon) = \mathbb{P}(|\mathbb{1}_{\{0\}}(X_n) - 1| > \epsilon)$$

for $\epsilon \in (0, 1)$. Thus

$$\{|\mathbb{1}_{\{0\}}(X_n) - 1| > \epsilon\}^c = \{X_n = 0\},$$

where

$$\mathbb{P}(X_n = 0) = 0.$$

It follows that

$$\mathbb{P}(|\mathbb{1}_{\{0\}}(X_n) - 1| > \epsilon) = 1 - \mathbb{P}(X_n = 0) = 1.$$

Therefore $f(X_n)$ does not converge in probability to $f(X)$. □

(P.4) *Define a probability space for a* **degenerate** *random variable, by which we mean a random variable X for which $\mathbb{P}_X(\{x\}) = 1$ for some point $x \in \mathbb{R}$.*

Proof. Let $(\Omega, \mathcal{A})$ be a measurable space where $\mathcal{A} = \mathcal{P}(\Omega)$ is the power set of Ω. Let $x \in \Omega$ and define the random variable $X : \Omega \mapsto \mathbb{R}$. We can define the Dirac measure $\delta_X : \mathcal{A} \mapsto [0, 1]$ on X to be

$$\delta_X(A) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A. \end{cases}$$

Notice that $\delta_X(\Omega) = 1$, since $x \in \Omega$. Suppose we have a countable sequence $(A_n)_{n\geq 1}$ of pairwise disjoint elements in $\mathcal{A}$. If $x \in \cup_{n=1}^{\infty} A_n$, then $x \in A_{n_0}$ for exactly one $n_0 \in \mathbb{N}$. Thus $\delta_X(A_{n_0}) = 1$ and $\delta_X(A_n) = 0$ for all $n \neq n_0$. Therefore

$$\delta_X(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \delta_X(A_i) = 1.$$

If $x \notin \cup_{n=1}^{\infty} A_n$, then $\delta_X(A_n) = 0$ for all n. Therefore

$$\delta_X(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \delta_X(A_i) = 0.$$

Thus we have satisfied the axioms of a probability measure. Therefore the Dirac measure is a probability measure on the space $(\Omega, \mathcal{A})$. Hence we have defined the probability space $(\Omega, \mathcal{A}, \delta_X)$ for a degenerate random variable X. □

(P.5) *Let $(X_n)_{n\geq 1}$ be a sequence of independent random variables which are identically distributed and taking values in the support $[a,b]$ $(a<b)$, where $\mu := \mathbb{E}\, X_n \in \mathbb{R}$. Show that $X_n \in \mathcal{L}_2$ and prove that*

$$\overline{X}_n \quad := \quad \frac{1}{n}\sum_{i=1}^{n} X_i \quad \xrightarrow{P} \quad \mu,$$

as $n \to \infty$. Show also that $\overline{X}_n \xrightarrow{a.s.} \mu$ as $n \to \infty$ (Hint: Look up Hoeffding's inequality).

Proof. If $X_n \in [a,b]$, then $a \leq X_n \leq b$ ($\mathbb{P}$-a.s.) and it follows that $|X_n| \leq |b-a|$ ($\mathbb{P}$-a.s.). Since $|b-a| \leq |b|$, then

$$\mathbb{E}|X_n|^2 = \mathbb{E}[X_n^2] \leq \mathbb{E}[b^2] = b^2 < \infty.$$

Then $\mathbb{E}[X_n^2]$ exists and is finite, hence $X_n \in \mathcal{L}_2$. Since each X_i are i.i.d. random variables with mean μ and variance $\mathrm{Var}(X_i) = \sigma^2$, then we apply Definition 5.2 (definition of variance) to define the variance of $\overline{X}_n$ to be

$$\mathbb{E}[|\overline{X}_n - \mu|^2] = \mathbb{E}[(\overline{X}_n - \mu)^2] = \mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}.$$

Thus

$$\lim_{n\to\infty} \mathbb{E}[|\overline{X}_n - \mu|^2] = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0.$$

By Definition 17.2, we say $\overline{X}_n$ converges to μ in $\mathcal{L}_2$. Therefore by Theorem 17.2, we can conclude $\overline{X}_n \xrightarrow{P} \mu$. Further, Hoeffding's inequality states that if we have a sequence of i.i.d. random variables with $\mu = \mathbb{E}[X_i]$ and $a \leq X_i \leq b$, then for any $\epsilon > 0$,

$$\mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon) \leq 2e^{\left(-\frac{2n\epsilon^2}{(b-a)^2}\right)}.$$

Now consider

$$\sum_{n=1}^{\infty} \mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon) \leq \sum_{n=1}^{\infty} 2e^{\left(-\frac{2n\epsilon^2}{(b-a)^2}\right)}.$$

Since $e^{\left(-\frac{2n\epsilon^2}{(b-a)^2}\right)}$ decreases exponentially as $n \to \infty$, then

$$\sum_{n=1}^{\infty} \mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon) < \infty$$

and by the Borel-Cantelli Lemma we can conclude $\mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon \ i.o.) = 0$. Then $\mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon \ f.o.) = 1$, meaning the sequence converges almost surely. Hence there exists an $N > 0$ such that for all $n \geq N$, $|\overline{X}_n - \mu| < \epsilon$ holds. Therefore $\overline{X}_n \xrightarrow{a.s.} \mu$ as $n \to \infty$. $\square$

(P.6) *Consider a sequence of random variable $(X_n)_{n\geq 1}$ and a random variable X with corresponding characteristic functions $(\varphi_{X_n})_{n\geq 1}$ and φ_X, respectively. Show that if $X_n \xrightarrow{P} X$, then $\varphi_{X_n} \to \varphi_X$ (pointwise).*

Proof. Since $X_n \xrightarrow{P} X$, then by definition of convergence in probability we know

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$$

for any $\epsilon > 0$. By Theorem 18.2, we know that if $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{D} X$. Consider the characteristic function $\phi_X(u) = \mathbb{E}[e^{iuX}]$ of X. Since $|e^{iuX}| = 1$ for all $u \in \mathbb{R}$, we know that $|\phi_X(u)| \leq 1$ and similarly $|\phi_{X_n}(u)| \leq 1$ for all n. This implies that the characteristic functions $\phi_{X_n}(u)$ and $\phi_X(u)$ are continuous and bounded for each $u \in \mathbb{R}$. Then by Theorem 18.1 we can say that $X_n \xrightarrow{D} X$ implies

$$\lim_{n\to\infty} \mathbb{E}[\varphi_{X_n}(u)] = \mathbb{E}[\varphi_X(u)],$$

where $\mathbb{E}[\varphi_{X_n}(u)] = \mathbb{E}[\mathbb{E}[e^{iuX_n}]] = \mathbb{E}[e^{iuX_n}] = \varphi_{X_n}(u)$. Then

$$\lim_{n\to\infty} \mathbb{E}[e^{iuX_n}] = \mathbb{E}[e^{iuX}],$$

for all $u \in \mathbb{R}$. Therefore $\varphi_{X_n} \to \varphi_X$ (pointwise). □

(P.7) *Let $(X_n)_{n\geq 1}$ be a sequence of random variables and $S_n \coloneqq \sum_{i=1}^n X_i$ $(n \in \mathbb{Z}^+)$. Show:*

(a) $X_n \xrightarrow{a.s.} 0$ implies $S_n / n \xrightarrow{a.s.} 0$.

(b) $X_n \xrightarrow{\mathcal{L}_p} 0$ implies $S_n / n \xrightarrow{\mathcal{L}_p} 0$ for any $p \geq 1$.

(c) $X_n \xrightarrow{P} 0$ does not imply $S_n / n \xrightarrow{P} 0$.

(d) $S_n / n \xrightarrow{P} 0$ implies $X_n / n \xrightarrow{P} 0$.

(a)

Proof. By definition of almost sure convergence, we know $X_n \xrightarrow{a.s.} 0$ implies

$$\mathbb{P}(\lim_{n\to\infty} X_n \neq 0) = 0.$$

Let $A = \{\omega : \lim_{n\to\infty} X_n \neq 0\}$. Then $A^c = \{\omega : \lim_{n\to\infty} X_n = 0\}$ and $\mathbb{P}(A^c) = 1$. Then for all $\omega \in A^c$ and for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$|X_n - 0| < \epsilon.$$

Then we show

$$\frac{|S_n|}{n} \leq |S_n| = \sum_{i=1}^n |X_i| < n\epsilon$$

for all $n\epsilon > 0$. Therefore there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\mathbb{P}(\{\omega : \lim_{n\to\infty} \frac{|S_n|}{n} = 0\}) = 1.$$

Hence $S_n / n \xrightarrow{a.s.} 0$. □

(b)

Proof. By definition of convergence in $\mathcal{L}_p$, we know $X_n \xrightarrow{\mathcal{L}_p} 0$ implies

$$\lim_{n\to\infty} \mathbb{E}|X_n|^p = 0,$$

for $p \geq 0$. By Theorem 17.2 we also know that if $X_n \xrightarrow{\mathcal{L}_p} 0$, then $X_n \xrightarrow{P} 0$ and the sequence is uniformly integrable. This implies that the average of the sequence $\frac{1}{n}\sum_{i=1}^{n} X_n$ will also converge to 0 in $\mathcal{L}_p$. Then we can apply Jensen's inequality to show

$$\mathbb{E}\frac{|S_n|^p}{n} = \mathbb{E}|\frac{1}{n}\sum_{i=1}^{n} X_n|^p \leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}|X_n|^p.$$

Since each $X_n \xrightarrow{\mathcal{L}_p} 0$, then $\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}|X_n|^p \xrightarrow{\mathcal{L}_p} 0$. Hence

$$\lim_{n\to\infty} \mathbb{E}\frac{|S_n|^p}{n} = 0$$

and therefore $S_n/n \xrightarrow{\mathcal{L}_p} 0$. □

(c)

Proof. Consider the following counterexample. Suppose

$$X_n = \begin{cases} 1 & \text{with prob. } \frac{1}{2} \\ -1 & \text{with prob. } \frac{1}{2}. \end{cases}$$

Then X_n converges in probability to the mean 0, i.e. $X_n \xrightarrow{P} 0$. However, $\sum_{i=1}^{n} X_i = S_n$ behaves like a random walk, meaning S_n diverges. Therefore it is not true that $S_n/n \xrightarrow{P} 0$. □

(d)

Proof. We know that $S_n/n \xrightarrow{P} 0$ by the Weak Law of Large Numbers, meaning

$$\lim_{n\to\infty} \mathbb{P}(|\frac{S_n}{n}| < \epsilon) = 1$$

for any $\epsilon > 0$. Since $|X_n| \leq |S_n| = |\sum_{i=1}^{n} X_i| \leq \sum_{i=1}^{n} |X_i|$, then

$$\lim_{n\to\infty} \mathbb{P}(|\frac{X_n}{n}| < \epsilon) = 1.$$

Therefore $X_n/n \xrightarrow{P} 0$. □