Lab 9

Convergence of RV's

9.1. Convergence in distribution

Definition 9.1.1. A sequence of RV's $X_1, X_2, \cdots$ is said to converge in distribution to X if $\lim_{n\to\infty} F_n(x) = F(x)$ for all $x \in \mathbb{R}$ for which <u>F is continuous</u>. We denote this as $X_n \xrightarrow{\mathcal{D}} X$.

Example 9.1.1 (Only for continuity points)**.** Let $F_n(x) = \mathbb{I}(x \geq \frac{1}{n})$ (cdf of the Unif$(0, \frac{1}{n})$). Then $X_n \xrightarrow{\mathcal{D}} X$, where X is the degenerate RV $X = 0$. But, $F(0) = 1$, while $F_n(x) = 0$ for all n. Hence this cdf fails to converge at $x = 0$, where F is discontinuous.

Example 9.1.2 ($X_n \xrightarrow{\mathcal{D}} X \not\Longrightarrow \lim_{n\to\infty} f_n(x) = f(x)$)**.** We may consider

$$f_n(x) = (1 - \cos(2\pi nx))\mathbb{I}(0 \leq x \leq 1), \qquad F_n(x) = \left(x - \frac{1}{2\pi n}\sin(2\pi nx)\right)\mathbb{I}(0 \leq x \leq 1)$$

We see that $\lim_{n\to\infty} F_n(x) = x\mathbb{I}(0 \leq x \leq 1)$, which is the cdf of Unif$(0,1)$, but $\lim_{n\to\infty} f_n(x)$ does not exist.

9.2. Convergence in probability

Definition 9.2.1. A sequence of RV's $X_1, X_2, \cdots$ is said to converge in probability to X if $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$ for all $\varepsilon > 0$. We denote this as $X_n \xrightarrow{\mathcal{P}} X$.

Example 9.2.1. (Metrization of $\xrightarrow{\mathcal{P}}$) Let

$$d(X, Y) = \mathbb{E}\frac{|X - Y|}{1 + |X - Y|}$$

Then $X_n \xrightarrow{\mathcal{P}} X \iff \lim_{n\to\infty} d(X_n, X) = 0$.

Proof. It suffices to show the inequalities

$$\frac{\varepsilon}{1+\varepsilon}\mathbb{P}(|X_n - X| > \varepsilon) \leq \mathbb{E}\frac{|X_n - X|}{1 + |X_n - X|} \leq \varepsilon + \mathbb{P}(|X_n - X| > \varepsilon)$$

for every $\varepsilon > 0$. Let $W = |X_n - X|$ and consider the function $f(w) = \frac{w}{1+w}$. It is easy to show f is monotone increasing, hence

$$\mathbb{P}(W > \varepsilon) = \mathbb{P}(f(W) > f(\varepsilon)) \leq \frac{\mathbb{E}f(W)}{f(\varepsilon)} = \frac{\mathbb{E}\frac{W}{1+W}}{\frac{\varepsilon}{1+\varepsilon}}$$

by Markov's inequality. Hence, we proved the lower inequality. For the upper inequality,

$$\begin{aligned}
\mathbb{E}f(W) &= \mathbb{E}f(W)\mathbb{I}(f(W) \leq \varepsilon) + \mathbb{E}f(W)\mathbb{I}(f(W) > \varepsilon) \\
&\leq \varepsilon + \mathbb{E}\mathbb{I}(f(W) > \varepsilon) \\
&= \varepsilon + \mathbb{P}(f(W) > \varepsilon) \\
&= \varepsilon + \mathbb{P}\left(W > \frac{\varepsilon}{1-\varepsilon}\right) \\
&\leq \varepsilon + \mathbb{P}(W > \varepsilon)
\end{aligned}$$

where for the first inequality (second line), we used the fact $f(W)\mathbb{I}(f(W) \leq \varepsilon) \leq \varepsilon$ and $f(W) \leq 1$. □

Remark: Notice that d also satisfies the three properties of a *metric* (Lab 1, Additional Problem 6). Not all modes of convergence have a corresponding metric; if such a metric exists, then such a convergence is called *metrizable.* Almost sure convergence is not metrizable.

9.3. Convergence almost surely

Definition 9.3.1. A sequence of RV's $X_1, X_2, \cdots$ is said to converge almost surely to X if $\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1$. We denote this as $X_n \xrightarrow{\text{a.s.}} X$.

An equivalent definition is $\mathbb{P}(\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)) = 1$. That is, the sample points for which $\lim_{n\to\infty} X_n$ and X don't agree has probability 0. Contrast this with *sure convergence*: $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for all $\omega \in \Omega$. That is, X_n converges pointwise to X, even for ω's that have probability 0.

Example 9.3.1. ($X_n \xrightarrow{\text{a.s.}} X \implies X_n \xrightarrow{\mathcal{P}} X$) We didn't prove this in class, so let's prove it now.

Proof. Let $O = \{\omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\}$. By almost-sure convergence, $\mathbb{P}(O) = 0$. Next, for any $\varepsilon > 0$, define

$$\begin{aligned}
A_n &= \bigcup_{m=n}^{\infty} \{|X_m - X| > \varepsilon\} \\
A_\infty &= \bigcap_{n=1}^{\infty} A_n
\end{aligned}$$

Now,

$$\begin{aligned}
\omega \in O^c &\implies \exists N : |X_n(\omega) - X(\omega)| < \varepsilon \quad \forall n \geq N \\
&\implies \omega \notin A_n \quad \forall n \geq N \\
&\implies \omega \notin A_\infty
\end{aligned}$$

which means $O^c \subseteq A_\infty^c \iff O \supseteq A_\infty$. Hence $\mathbb{P}(A_\infty) \le \mathbb{P}(O) = 0$. Finally,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) \le \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A_\infty) = 0$$

where we applied the Completeness Property proved in lecture. □

9.4. Convergence in rth mean

Definition 9.4.1. For $r \ge 1$, a sequence of RV's $X_1, X_2, \cdots$ is said to converge in rth mean (or in the L^r-norm) to X if $\lim_{n\to\infty} \mathbb{E}|X_n - X|^r = 0$. We denote this as $X_n \xrightarrow{L^r} X$. Note that $X_n \xrightarrow{L^r} X \implies X_n \xrightarrow{L^s} X$ for $r \ge s \ge 1$; the proof is a quick application of Lyapunov's inequality.

Example 9.4.1. Let X_i be a sequence of RV's and $S_n = \sum_{i=1}^n X_i$. Show that $X_n \xrightarrow{L^r} 0 \implies \frac{S_n}{n} \xrightarrow{L^r} 0$

Proof. For all $\varepsilon > 0$, there exists an N such that $\mathbb{E}|X_m|^r < \varepsilon$ for all $m \ge N$, by L^r convergence. Then,

$$\begin{aligned}
\mathbb{E}\left|\frac{S_n}{n}\right|^r &\le n^{-r}\left(\sum_{i=1}^n (\mathbb{E}|X_i|^r)^{\frac{1}{r}}\right)^r && \text{Minkowski's inequality}\\
&= n^{-r}\left(\sum_{i=1}^{m-1} (\mathbb{E}|X_i|^r)^{\frac{1}{r}} + \sum_{i=m}^{n} (\mathbb{E}|X_i|^r)^{\frac{1}{r}}\right)^r \\
&< n^{-r}\left(C + (n-m+1)\varepsilon^{\frac{1}{r}}\right)^r && C = \sum_{i=1}^{m-1} (\mathbb{E}|X_i|^r)^{\frac{1}{r}} < \infty\\
&\overset{n\to\infty}{\to} \varepsilon
\end{aligned}$$

Since ε is arbitrary, we have $\lim_{n\to\infty} \mathbb{E}\left|\frac{S_n}{n}\right|^r = 0$. □

9.5. Convergence of functions of RV's

Theorem 9.5.1 (Continuous mapping theorem)**.** Let g be a (possibly many to many) continuous function. Then

1. $\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X} \implies g(\mathbf{X}_n) \xrightarrow{\mathcal{D}} g(\mathbf{X})$
2. $\mathbf{X}_n \xrightarrow{\mathcal{P}} \mathbf{X} \implies g(\mathbf{X}_n) \xrightarrow{\mathcal{P}} g(\mathbf{X})$
3. $\mathbf{X}_n \xrightarrow{\text{a.s.}} \mathbf{X} \implies g(\mathbf{X}_n) \xrightarrow{\text{a.s.}} g(\mathbf{X})$.

Corollary 9.5.1 (Slutsky's theorem)**.** Let X_n, Y_n be sequences of scalar RV's. If $X_n \xrightarrow{\mathcal{D}} X$ and $Y_n \xrightarrow{\mathcal{D}} c$, then

1. $X_n + Y_n \xrightarrow{\mathcal{D}} X + c$
2. $X_n Y_n \xrightarrow{\mathcal{D}} cX$
3. $X_n / Y_n \xrightarrow{\mathcal{D}} X/c$, provided c is invertible

Everything remains true if we replace all $\xrightarrow{\mathcal{D}}$ with $\xrightarrow{\mathcal{P}}$.

9.6. Additional Problems

1. Define the function $\widetilde{d}(X,Y) = \mathbb{E}[|X-Y| \wedge 1]$ (recall $a \wedge b = \min(a,b)$). Prove that $X_n \xrightarrow{\mathcal{P}} X \iff \lim_{n\to\infty} \widetilde{d}(X_n, X) = 0$. **Hint:** Prove that $d(X,Y) \leq \widetilde{d}(X,Y) \leq 2d(X,Y)$, where d is defined in Example 9.2.1.

2. let $X_i \overset{\text{iid}}{\sim} F$ and F satisfies $\lim_{x\to\infty} x^2[1-F(x)] = 0$. Prove that $\max_{1\leq i\leq n} \frac{X_i}{\sqrt{n}} \xrightarrow{\mathcal{P}} 0$.

3. Let $X_i \overset{\text{iid}}{\sim} \text{Exp}(1)$. Show that $\mathbb{P}(X_n > \alpha \log(n) \text{ i.o.}) = \mathbb{I}(\alpha \leq 1)$. **Hint:** Borel-Cantelli + the integral test for convergence.

4. Let $X_1, X_2, \cdots$ be positive, iid RV's with common density $f(x)$, satisfying $\lim_{x\downarrow 0} f(x) = \lambda$. Prove that $n \min(X_1, \cdots, X_n) \xrightarrow{\mathcal{D}} \text{Exp}(\lambda)$.

5. Let $X_i \overset{\text{iid}}{\sim} \text{Unif}(0,1)$ and set $Y_n = \prod_{i=1}^n X_i$. Define the random function $f(z) = \sum_{n=1}^\infty Y_n z^n$, and let R be the radius of convergence of f. Show that $\mathbb{P}(R = e) = 1$, where $e = 2.718\cdots$. **Hint:** Recall radius of convergence is defined as $R = \frac{1}{\limsup |Y_n|^{1/n}}$, and you may use the fact that Y_n has density $f_{Y_n}(y) = \frac{\log^{n-1}(1/y)}{(n-1)!}$.