E2:202 Random Processes **Nov. 27, 2020**

Tutorial 9: Weak Convergence and Some Limit Theorems

Lecturer: Parimal Parag *TA: Arvind* *Scribes: Krishna Chaythanya KV*

As a simple exercise in showing that convergence in probability does not necessarily imply convergence in distribution, can you use the Borel-Cantelli lemmas to construct a random variable (and a probability assignment) which converges to $X \equiv 0$ in probability, but not almost surely?
Hint: Construct a sequence of Bernoulli random variables X_n such that $\lim_{n\to\infty} \mathbb{P}(X_n = 1) = 0$, but $\sum_{n\in\mathbb{N}} \mathbb{P}(X_n = 1) = \infty$.

Now, let us revisit convergence in distribution, and show that convergence in probability implies convergence in distribution.

Theorem. If $X_n \xrightarrow[n\to\infty]{\text{p.}} X$, then $X_n \xrightarrow[n\to\infty]{\text{d.}} X$.

Proof. We first put down an easily verifiable claim:

Claim. For two arbitrary random variables, X and Y, on the same probability space, and for $a \in \mathbb{R}$ and $\epsilon > 0$,

$$\mathbb{P}(Y \le a) \le \mathbb{P}(X \le a + \epsilon) + \mathbb{P}(|Y - X| > \epsilon).$$

Proof of Claim. $\{Y \le a\} \subseteq \{X \le a + \epsilon\} \cup \{|Y - x| > \epsilon\}$. The result then follows from the union bound. □

Now, we have that if a is a point of continuity of the limit random variable X for every $\epsilon > 0$,

$$\mathbb{P}(X_n \le a) \le \mathbb{P}(X \le a + \epsilon) + \mathbb{P}(|X_n - X| > \epsilon). \tag{9.1}$$

Likewise,

$$\mathbb{P}(X \le a - \epsilon) \le \mathbb{P}(X_n \le a) + \mathbb{P}(|X_n - X| > \epsilon), \tag{9.2}$$

where both eqn. (9.1) and eqn. (9.2) follow from the claim. Hence, it is true that

$$\mathbb{P}(X < a + \epsilon) - \mathbb{P}(|X_n - X| > \epsilon) \le \mathbb{P}(X_n \le a) \le \mathbb{P}(X \le a + \epsilon) + \mathbb{P}(|X_n - X| > \epsilon).$$

Taking the limit as $n \to \infty$, we obtain that

$$\lim_{n\to\infty} F_{X_n}(a) = F(a)$$

for every a that is a point of continuity. □

Now, the complete convergence picture looks like this.

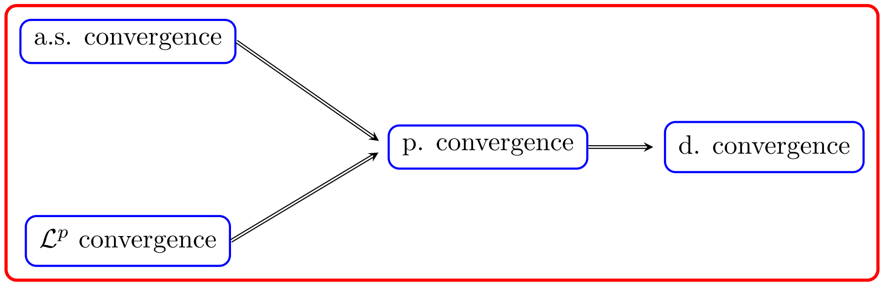

Figure 9.1: Picture of relationship between convergences

To recap the all the lectures of the course briefly, we have learnt what probability spaces are, how random variables are defined on probability spaces, how random variables interact with one another (correlation), and how sequences of random variables behave in the limit.

As an important special case of the limiting behaviour of random variables, we consider the empirical average of i.i.d. random variables, and examine how this new random variable behaves, in the "large-n-limit".

Given i.i.d random variables $X_1, X_2, \ldots, X_n$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, our object(s) of interest are

$$S_n \triangleq \frac{\sum\limits_{i=1}^{n} X_i}{n}, \text{ and,}$$

$$Z_n \triangleq \frac{\sum\limits_{i=1}^{n} (X_i - \mathbb{E}[X_1])}{\sqrt{n}\sqrt{\mathrm{var}(X_1)}}.$$

For convenience, we denote $\mathbb{E}[X_1] \triangleq \mu$ and $\mathrm{var}(X_1) \triangleq \sigma^2$.

Theorem (Weak Law of Large Numbers (WLLN))**.**

$$S_n \xrightarrow[n\to\infty]{\text{p.}} \mu.$$

Proof. By the Chebyshev inequality, for any $\epsilon > 0$,

$$\mathbb{P}\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\mathrm{var}(S_n)}{\epsilon^2}$$
$$= \frac{\sigma^2}{n\epsilon^2} \xrightarrow[n\to\infty]{0} .$$

□

Remark. By turning Chebyshev's inequality on its head, we get that for any $n \geq 1$ and for a fixed $\epsilon > 0$,

$$S_n \in \left[\mu - \frac{\sigma}{\sqrt{n\epsilon}}, \mu + \frac{\sigma}{\sqrt{n\epsilon}}\right] \quad \text{w.p. } \epsilon.$$

Hence, S_n is "concentrated about its mean".

The following theorem strengthens the previous theorem.

Theorem (Strong Law of Large Numbers (SLLN))**.** Given that $\mathbb{E}\left[|X_1|\right] = B < \infty$, it is true that

$$S_n \xrightarrow[n\to\infty]{\text{a.s.}} \mu.$$

We will prove a "weakened" form of this statement, with the additional condition that $\mathbb{E}\left[(X_n-\mu)^4\right] = B < \infty$.

Proof: Borel SLLN. For any $\epsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(|S_n-\mu| \geq \epsilon\right) &= \mathbb{P}\left(\left|\sum_{i=1}^{n}(X_i-\mu)\right| \geq n\epsilon\right)\\ &\leq \frac{\mathbb{E}\left[\left(\sum_{i=1}^{n}(X_i-\mu)\right)^4\right]}{n^4\epsilon^4} \quad \text{(Why?)}\end{aligned}$$

$$\text{Now, } \mathbb{E}\left[\left(\sum_{i=1}^{n}(X_i-\mu)\right)^4\right] = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\mathbb{E}\left[(X_i-\mu)(X_j-\mu)(X_k-\mu)(X_l-\mu)\right].$$

Further, since $(X_i : i \in [n])$ are i.i.d, only terms of the form $\mathbb{E}\left[(X_i-\mu)^2(X_j-\mu)^2\right]$ for $i \neq j$ and $\mathbb{E}\left[(X_i-\mu)^4\right]$ are non-zero. Now,

$$\mathbb{E}\left[(X_i-\mu)^2(X_j-\mu)^2\right] = \mathbb{E}\left[(X_i-\mu)^2\right]\mathbb{E}\left[(X_j-\mu)^2\right] = \sigma^4.$$

We also have $\mathbb{E}\left[(X_i-\mu)^4\right] = B$. Hence,

$$\mathbb{E}\left[\left(\sum_{i=1}^{n}(X_i-\mu)\right)^4\right] = 3n(n-1)\sigma^4 + nB.$$

Show the above result using a counting argument.
Therefore,

$$\mathbb{P}\left(|S_n-\mu| \geq \epsilon\right) \leq \frac{3n\left(n-1\right)\sigma^4 + nB}{n^4\sigma^4},$$

which satisfies $\sum_{n\in\mathbb{N}}\mathbb{P}\left(|S_n-\mu| \geq \epsilon\right) \leq \infty$. By the Borel-Cantelli Lemma,

$$S_n \xrightarrow[n\to\infty]{\text{a.s.}} \mu.$$

□

As an application of the strong law of large numbers, we describe the principle of "Monte-Carlo Integration".

Suppose that we wish to compute the value of a multi-dimensional definite integral

$$I \triangleq \int_{\Omega} f\left(\underline{x}\right) d\underline{x},$$

where $\Omega \subseteq \mathbb{R}^m$ has a volume

$$V \triangleq \int_{\Omega} d\underline{x}.$$

We now write

$$I = V \int_{\Omega} \frac{1}{V} f\left(\underline{x}\right) d\underline{x}, \\ \approx V \frac{1}{n} \sum_{i=1}^{n} f\left(X_i\right)$$

for large n and for i.i.d samples $\underline{X}_1, \ldots, \underline{X}_n$.

Example (Computing the value of π)**.** An interesting example is a method of approximately computing the value of π. Consider

$$H\left(x,y\right) = \begin{cases} 1, & \text{if } x^2 + y^2 \leq 1, \\ 0, & \text{o.w.}, \end{cases}$$

with $\Omega = [-1,1] \times [-1,1]$. Note that here $V = 4$. Now,

$$I_\pi = \int_{\Omega} H\left(x,y\right) dxdy = \pi,$$

which can be approximately computed as

$$I_\pi \approx 4 \frac{1}{N} \sum_{i=1}^{N} H\left(X_i, Y_i\right)$$

for $\left(X_i, Y_i\right) \overset{\text{i.i.d}}{\sim} \text{Unif}\left(\Omega\right)$.

We now turn our attention to yet another limit theorem which involves $\left(Z_n : n \in \mathbb{N}\right)$. Recall that

$$Z_n = \frac{\sum\limits_{i=1}^{n} \left(X_i - \mathbb{E}\left[X_1\right]\right)}{\sqrt{n}\sqrt{\text{var}\left(X_1\right)}}.$$

For simplicity, we will assume that $\mathbb{E}\left[X_1\right] = 0$.

Theorem (Central Limit Theorem (CLT))**.** $Z_n \xrightarrow[n\to\infty]{\text{d.}} \mathcal{N}\left(0,1\right)$.

Proof. To show convergence in distribution, recall that it suffices to show that $\phi_{Z_n}\left(\omega\right) \xrightarrow[n\to\infty]{\text{d.}} \phi_{\mathcal{N}(0,1)}\left(\omega\right)$, for all $\omega \in \mathbb{R}$. Note that

$$\begin{aligned} \phi_{Z_n}\left(\omega\right) &= \mathbb{E}\left[\exp\left(j\omega\left(\frac{\sum_{i=1}^{n} X_i}{\sqrt{n\sigma^2}}\right)\right)\right] \\ &= \mathbb{E}\left[\exp\left(j\frac{\omega}{\sqrt{n\sigma^2}}\sum_{i=1}^{n} X_i\right)\right] \\ &= \left(\mathbb{E}\left[e^{j\frac{\omega}{\sigma\sqrt{n}}X_1}\right]\right)^n \quad \text{because } \left(X_n : n \in \mathbb{N}\right) \text{ are i.i.d.} \end{aligned}$$

Let $\psi\left(\omega\right) \triangleq \mathbb{E}\left[e^{j\frac{\omega}{\sigma\sqrt{n}}X_1}\right]$. By Taylor's theorem,

$$\psi\left(\omega\right) = \psi\left(0\right) + \psi'\left(0\right)\omega + \psi''\left(0\right)\frac{\omega^2}{2} + R\left(\omega\right),$$

where $\frac{R(\omega)}{\omega^2} \xrightarrow[\omega \to 0]{} 0$. Now observe that

$$\begin{aligned} \psi(0) &= 1, \\ \psi'(0) &= 0, \quad \text{Since } \mathbb{E}[X_1] = 0 \\ \psi''(0) &= -\frac{\sigma^2}{n}. \quad \textbf{Calculate this explicitly}. \end{aligned}$$

Hence, we get that

$$\phi_{Z_n}(\omega) = \left(1 - \frac{1}{n}\left(\frac{\omega^2\sigma^2}{2} + \eta_n\right)\right)^n,$$

where $\lim_{n\to\infty} \eta_n = 0$. **Fill in the details as an exercise.**

Using the result that if $(a_n)_{n\in\mathbb{N}}$ is any sequence of real numbers such that $\lim_{n\to\infty} a_n = a$, then

$$\lim_{n\to\infty}\left(1 + \frac{a_n}{n}\right)^n = e^a,$$

we get that

$$\lim_{n\to\infty} \phi_{Z_n}(\omega) = e^{-\frac{\omega^2\sigma^2}{2}}, \quad \text{for any } \omega \in \mathbb{R}.$$

Thus $Z_n \xrightarrow[n\to\infty]{\text{d.}} \mathcal{N}(0,1)$. □

Remark. The key take-away from the CLT is that it provides an asymptotic concentration result. That is, it shows that for "large-enough n" (and for zero mean i.i.d RVs)

$$\begin{aligned} |\mathbb{P}(Z_n \le x) - \Phi(x)| &\le \epsilon \qquad (\Phi(\cdot) \text{ is the CDF of } \mathcal{N}(0,1)) \\ \Rightarrow |\mathbb{P}(Z_n > x) - Q(x)| &\le \epsilon \qquad (Q(\cdot) \text{ is the CCDF of } \mathcal{N}(0,1)) \\ \Rightarrow Q(x) - \epsilon \le \mathbb{P}\left(\sum_{i=1}^n X_i > \sqrt{n}\sigma x\right) &\le Q(x) + \epsilon. \end{aligned}$$

Put x as $Q^{-1}(\delta)$, for some "small δ". Hence,

$$\mathbb{P}\left(\sum_{i=1}^n X_i > \sqrt{n}\sigma Q^{-1}(\delta)\right) \approx \delta + \epsilon.$$

Or in other words, for "large-enough" n,

$$\sum_{i=1}^n X_i \in \left[-\sqrt{n}\sigma Q^{-1}(\delta), \sqrt{n}\sigma Q^{-1}(\delta)\right] \text{ w.p. } 1-\delta.$$

This is a "tighter" concentration bound than that provided by Chebyshev's inequality, but the drawback is that it is asymptotic (i.e., n needs to be *large*).