Lecture 9: Modes of Convergence

If $(X_n)_{n\geq 1}$ & X_n are random variable, how do we quantify statement that X_n's converge to X in some way?

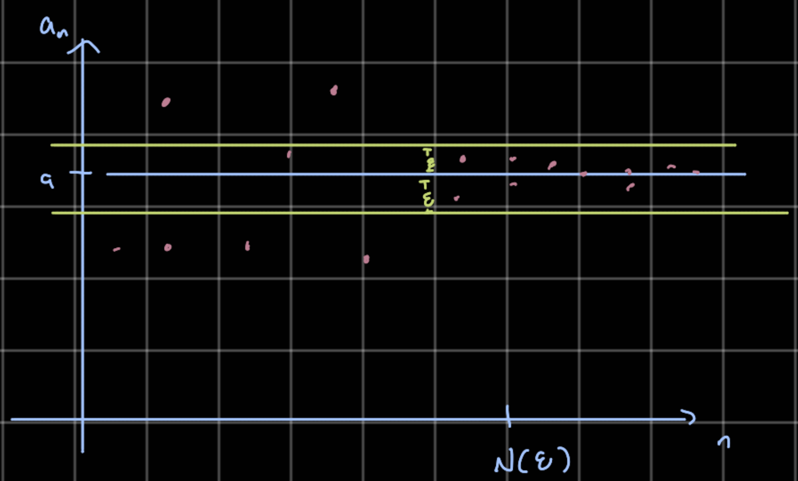
Motivation from last time: WLLN told us that empirical mean converges to true avg as $n \to \infty$.

Convergence of Numbers:

- For a sequence $(a_n)_{n\geq 1} \subset \mathbb{R}$, we say $a_n \xrightarrow{\text{converges}} a$ if $\forall \varepsilon > 0$, $\exists N(\varepsilon)$ s.t.

$$|a_n - a| < \varepsilon \quad \forall \; n \geq N(\varepsilon)$$

Visually:



Convergence of Functions

Q/ If we consider fcns $f_1, f_2, \ldots$ where $f_i : [0,1] \to \mathbb{R}$, what does it mean for $f_n \to f$?

A/ There are many useful and inequivalent (different) ways for $f_n \to f$.

eg:

- $f_n(x) \to f(x) \quad \forall x \in [0,1]$

  "pointwise convergence"

  ↳ once you fix x, this is just a sequence of #s to converge to another number. (it's PW-conv) bc you have to do this $\forall x$)

these are inequivalent notions of convergence

- $\int_0^1 |f_n(x) - f(x)|\,dx \to 0$

  avg diff btwn f_n & f

  "convergence in L1"

  ⇒ just a sequence of #s, once you take the integral

- $\forall\, \varepsilon > 0$

$$\int_0^1 \mathbb{1}_{\{|f_n(x) - f(x)| > \varepsilon\}}\,dx \to 0$$

"convergence in measure to f"

ex: if $f_n(x) = \begin{cases} 0 & x \neq \frac{1}{2} \\ 1 & x = \frac{1}{2} \end{cases}$, $f(x) = 0$

$f_n \not\to f$ pointwise, but $f_n \to f$ in L_1 & in measure

Punchline: Convergence of functions to another fcn is trickier business than convergence of #s. This is because every fcn can be thought of as an (infinitely) long list of numbers.

$$f = \{f(x) : x \in [0,1]\}$$

Q/ How does this connect to probability?

A/ Random variables are functions.

So, when we say $X_n \to X$, we really need to specify "how" convergence takes place. (ie, we need to say the "mode of convergence")

3 examples of modes of convergence:

Def: $X_n \to X$ in probability if $\forall \varepsilon > 0$

motivated by WLLN

$$\lim_{n\to\infty} P\{|X_n - X| > \varepsilon\} = 0$$

1. Ex: WLLN says that empirical means converge to expectation in probability

   ↳ convergence to a constant (the mean)

2. Ex: Let $X_n = X + Z$, where $Z_n \sim \mathcal{N}(0, \frac{1}{n})$

   Z: noise

   noise gets better & better for us as time goes on.

   Claim: $X_n \to X$ in p

$$P\{|X_n - X| > \varepsilon\} = P\{|Z_n| > \varepsilon\}$$

↳ rescale Z so it becomes std normal rv

$$= P\{|z| > \sqrt{n}\,\varepsilon\}$$

Markov's inequality

$$\le \frac{\mathbb{E}[z]}{\varepsilon\sqrt{n}} \to 0 \text{ as } n \to \infty$$

↳ convergence to a RV

↳ shows applications of Markov/Chebyshev in practice

important! the limit is in the probability not the limit of the probabilities, unlike convergence in probability

Def: $X_n \to X$ almost surely (a.s.) if

allows us to state SLLN

$$P\{\underbrace{\lim_{n\to\infty} X_n}_{} = X\} = 1$$

this is a rv, obtained by getting pointwise limit of the rvs. defined by:

$$\omega \longmapsto \lim_{n\to\infty} X_n(\omega)$$

↳ more explicitly:

$$P(\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$$

↳ this is the same as $X_n \to X$ a.s.

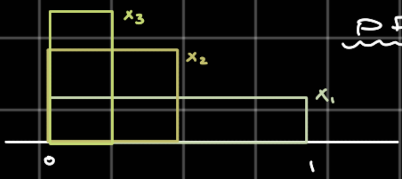
- Note: a.s. convergence is very similar to pointwise convergence in functions, except we don't require $X_n(\omega) \to X(\omega)\ \forall \omega \in \Omega$, just on an event having probability 1.

prob. measure $\quad$ sample sample

ex: $P = \mathrm{Unif}(0,1) \qquad (\Omega = [0,1])$

$$X_n(\omega) = \begin{cases} n^2 & 0 \le \omega \le \frac{1}{n} \\ 0 & \frac{1}{n} < \omega \le 1 \end{cases}$$

↳ seq. of fcns, not independent



Claim: $X_n(\omega) \to 0$ a.s.

pf: $$\lim_{n\to\infty} X_n(\omega) = \begin{cases} +\infty & \omega = 0 \\ 0 & \omega \in (0,1] \end{cases}$$

$$P\{\lim X_n = X\} = P((0,1]) = 1$$

Note: $X_n \to 0$ i.p. (in probability) also in this example:

$$P\{|X_n - 0| > \varepsilon\} = \frac{1}{n}$$

for $\varepsilon < 1$

Caution: Modes of convergence such as a.s. or i.p. do not imply convergence of $\mathbb{E}$.

$$X_n \to X \quad \not\Rightarrow \quad \mathbb{E}[X_n] \to \mathbb{E}[X]$$

indeed, in above example, we have

$$\mathbb{E}[X_n] = \frac{1}{n} \cdot n^2 = n \longrightarrow +\infty$$
$$\ne \mathbb{E}[0] = 0$$

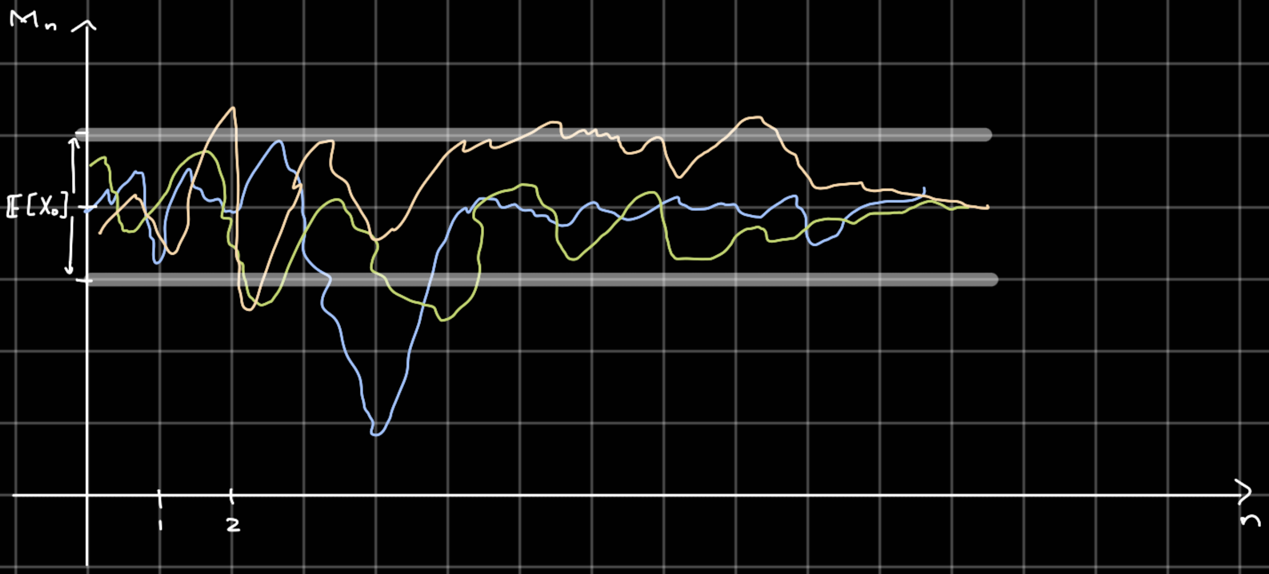
A major result of a.s. convergence is the Strong Law of Large Numbers (SLLN):

Strong Law of Large Numbers:

Let $(X_n)_{n\ge1} \underset{iid}{\sim} X$, $\mathbb{E}[X] < \infty$, then

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i \to \mathbb{E}[X_i] \quad \text{a.s.}$$

note: WLLN is the same but this is replaced with i.p.



tells us about the ensemble of sample paths

WLLN: With overwhelming probability, the majority of paths lie in the narrow strip for n large.

BUT it doesn't tell us long-term behavior of any given sample path $M_n(\omega)$, $n \geq 1$

tells us about individual paths in their limit

SSLN: For all ω in an event, with probability 1, the sample path

$M_n(\omega) \longrightarrow E(X)$ as $n \to \infty$

Note: The "with probability 1" is important

↳ ex: Consider $X_n \underset{\text{iid}}{\sim} \text{Bern}(\frac{1}{2})$ (iid coin flips)

It's possible to flip an infinite sequence of heads (= 1's)

But with probability 1, we won't experience this sequence.

SLLN doesn't say anomolous things (eg, above) but it does say that we'll never experience them (w.p. 1)

We won't prove the SLLN in this course :(

However, you might guess that SLLN ⇒ WLLN. This is true :)

In general $X_n \to X$ a.s. $\Rightarrow X_n \to X$ i.p.

(ie, a.s. convergence is stronger than convergence i.p.

There's an even weaker mode of convergence (implied by convergence i.p): convergence in distribution.

Def: $X_n \to X$ in distribution if

$$F_{X_n}(x) \to F_X(x) \quad \forall \text{ points } x \text{ where } F_X \text{ is continuous}$$

CDF

Informally: The histograms of outcomes for X_n's converge to that of X.

The qualification re: "continuity points of F_X" is important:

ex: $X_n \sim \text{Unif}(-\frac{1}{n}, \frac{1}{n})$, $n \geq 1$

Claim: $X_n \to 0$ in distribution (& i.p)

$$F_{X_n}(x) \to \begin{cases} 0 & x < 0 \\ \frac{1}{2} & x = 0 \\ 1 & x > 0 \end{cases}$$

compare with

$$F_0(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

In general:

$$X_n \to X \text{ a.s.} \Rightarrow X_n \to X \text{ i.p} \Rightarrow X_n \to X \text{ in dist.}$$

Perhaps most important example of convergence in distribution:

Central Limit Thm: Let $(X_n)_{n \geq 1} \overset{\sim}{\text{iid}} X$ with $\text{var}(X) = \sigma^2$, $\mathbb{E}[X] = \mu$. If $Z \sim \mathcal{N}(0,1)$ then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{(X_i - \mu)}{\sigma} \to Z \text{ in distribution}$$