Convergence of random variables

lec 42

Convergence of r.v's is Probably the most challenging Part of this course.

if we are given the seq^n of real number's $(a_n)_{n \in \mathbb{N}}$ we have Convergence of Number's.

* we are going to talk about convergence of r.v's. (seq^n of r.v's)

$(\Omega, \mathcal{F}, \mathbb{P})$ Probability space

& $X_1, X_2, X_3, \ldots$. seq^n of r.v's

* random variable are function's

$$X_i : \Omega \longrightarrow \mathbb{R}$$

so we are essentially taking about convergence of function's.

Definition ① :-

The seq^n $(X_n)_{n \in \mathbb{N}}$ converge pointwise or surely to X if $X_n(\omega) \longrightarrow X(\omega)$ $\forall \omega \in \Omega$ as $n \longrightarrow \infty$

Pointwise convergence of function's

from Real Analysis: Pointwise convergence

seq of functions: $(f_1, f_2, f_3, f_4, \ldots)$

$$f_n : I \longrightarrow \mathbb{R}$$

Pointwise convergence:

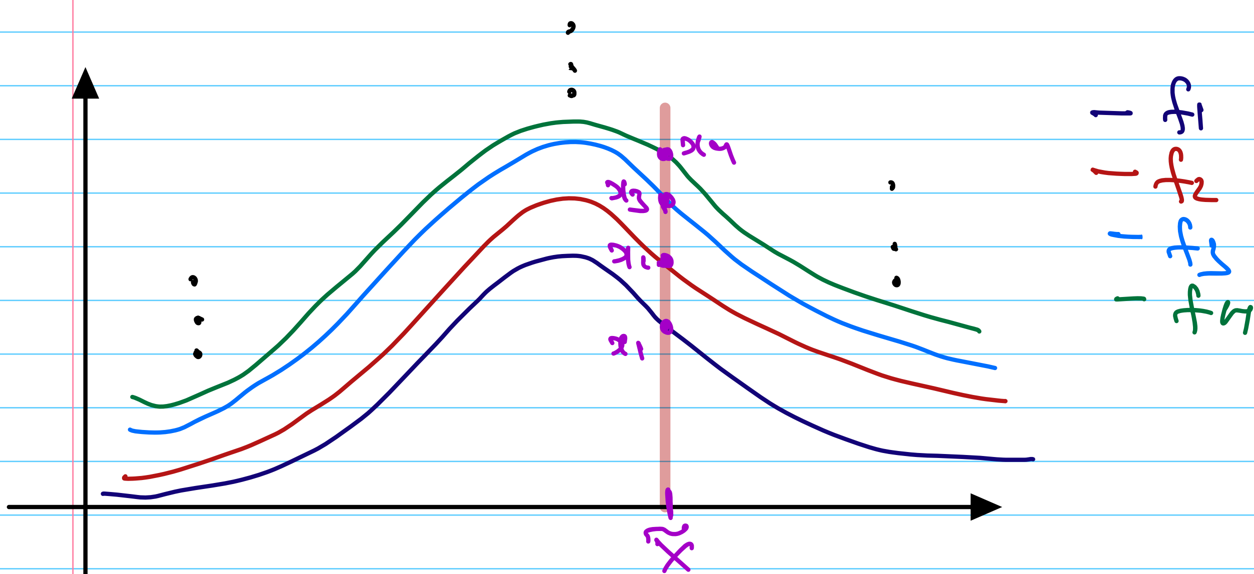
$$(f_1, f_2, f_3, f_4, \dots)$$ is

pointwisely convergent to a function

$f: I \longrightarrow \mathbb{R}$ if for all $\tilde{x} \in I$:

$$(f_1(\tilde{x}), f_2(\tilde{x}), f_3(\tilde{x}), \dots)$$

is convergent to $f(\tilde{x})$



$$\left[\forall \tilde{x} \in I, \ \forall \varepsilon > 0 \ \exists N \in \mathbb{N}, \ \forall n \geq N: \ |f_n(\tilde{x}) - f(\tilde{x})| < \varepsilon \right]$$

Back to convergence of r.v's

$X_n(\omega) \xrightarrow{n\to\infty} X(\omega)$ (Pointwise convergence of r.v)

we pick any $\omega \in \Omega$ in sample space, for that ω we have a seq^n of real number's

$$\left(X_n(\omega)\right)_{n\in\mathbb{N}}$$ (seq^n of real number's)

As soon as ω realizes (picked by mother god)
The seq^n of random variables becomes seq^n of real number's.

$\Rightarrow$ $\left(X_n\right)_{n\in\mathbb{N}}$ (seq^n of real number's)

converges to $X(\omega)$

$\Rightarrow$ $\forall \omega \in \Omega$, $\exists \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that
$\forall n > N$ we have $|X_n(\omega) - X(\omega)| < \varepsilon$

$$\Rightarrow \lim_{n\to\infty} X_n(\omega) = X(\omega) \quad \forall \omega \in \Omega$$

$$\Rightarrow \lim_{n\to\infty} X_n(\omega) - X(\omega) = 0$$

This is too stricter notion of Probability.
because in Probability, we are willing to sacrifice Prob measure 0 sets. we don't really care about Probability measure 0 sets.

we will slightly weaken the convergence

Def 1:- Almost sure convergence

we say X_n converges to X almost surely or with Prob 1 if $X_n(\omega) \xrightarrow{n\to\infty} X(\omega)$ on a set of Prob 1.

That is

$$\mathbb{P}\left(\{\omega \mid X_n(\omega) \longrightarrow X(\omega)\}\right) = 1$$

The set of all ω's $\in \Omega$.

* This convergence happen's not necessarily for all $\omega \in \Omega$, But on a set of Prob 1.

$\Rightarrow$ There may be some $\omega \in \Omega$ that left out (convergence does not happen) But they have Prob 0. (almost sure convergence, or strong convergence)

we can also Prove that The limit of a seq^n of measurable function's is always a measurable function's.

Definition 2 Convergence in Probability

we say X_n Converges to X in Probability if

$$\forall \varepsilon > 0, \quad \lim_{n \to \infty} \mathbb{P}\left(|X_n - X| > \varepsilon\right) = 0$$

* Convergence in Prob is very different from Convergence with Probability 1.

* in sure Convergence and almost sure Convergence, in these 2 cases, we have seq of r.v themselves Converging

* in convergence in Probability, it is bit of misnomer to say $X_n \longrightarrow X_n$ in Probability.
 This is Just that some seq^n of Prob is Converging to 0.

(The seq^n of Probabilities converges)

let $\mathbb{P}(|X_n - X| > \varepsilon) = P_n$

then the Seq $(P_n)_{n \in \mathbb{N}}$ goes to 0.

1. for any fixed value $\varepsilon > 0$

(i) we look at the event $|X_n - X| > \varepsilon$, which represents the sets of outcomes $\omega \in \Omega$, where the difference b/w $X_n(\omega)$ and $X(\omega)$ is at least ε

(ii) The Prob $P(|X_n - X| > \varepsilon)$ measures how "Large" this event is in term's of the Probability measure P.

(iii) Convergence in Prob mean's that this Prob goes to 0 as $n \to \infty$

2. Breakdown of the definition

Let $A_n^{\varepsilon} = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \varepsilon\}$

Then

$$\mathbb{P}(|X_n - X| \geq \varepsilon) = \mathbb{P}(A_n^{\varepsilon})$$

Convergence in Prob mean

$$\forall \varepsilon > 0 \quad \lim_{n \to \infty} P(A_n^{\varepsilon}) = 0$$

Example's:-

let $(\Omega, \mathcal{F}, \mathbb{P})$ be a Probability space.

EACH r.v $X_n : \Omega \longrightarrow \mathbb{R}$ is a measurable function on the same sample space.

we study how the seq $(X_n)_{n \in \mathbb{N}}$ converges to a r.v X

① almost sure convergence (a.s)

Def: we say X_n converges almost surely (a.s) to X if

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

equivalently, $\forall \varepsilon > 0$

$$\mathbb{P}\left(\left\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon \text{ infinitely often in } n\right\}\right) = 0$$

Connection to Real analysis (Pointwise convergence)

in Real Analysis $f_n \to f$ pointwise on X means:

$\forall x_0 \in X$, $\forall \varepsilon > 0$, $\exists N(x_0, \varepsilon) \in \mathbb{N}$ s.t $\forall n > N(x_0, \varepsilon)$

we have $|f_n(x_0) - f(x_0)| < \varepsilon$

Here $X_n \xrightarrow{a.s} X$ mean's the same thing, but we allow this connection to fail on a set of ω's of Probability 0.

That is

$$\mathbb{P}\left(\{\omega : \exists N(\omega, \varepsilon) \text{ s.t } |X_n(\omega) - X(\omega)| < \varepsilon, \forall n > N(\omega, \varepsilon)\}\right) = 1$$

So, it is literally Pointwise Convergence on the sample space, up to a negligible subset.

② Convergence in Probability

<u>Definition:</u>

we say X_n Converges in Probability to X if

$$\forall \varepsilon > 0 \,, \quad \mathbb{P}(|X_n - X| > \varepsilon) \longrightarrow 0$$

Connection to real analysis (uniform in measure)

in real analysis, "uniform convergence" says:

$\forall \varepsilon > 0, \ \exists N(\varepsilon) \in \mathbb{N}$, s.t $\forall n > N(\varepsilon)$

we have $\quad |f_n(x) - f(x)| < \varepsilon, \ \forall x \in X$

$$\Rightarrow \lim_{n \to \infty} \forall x \ |f_n(x) - f(x)| = 0$$

Convergence in Probability relaxes with, $\forall \varepsilon > 0$

$\forall \delta > 0, \ \exists N(\varepsilon, \delta) \in \mathbb{N}$, s.t $\forall n > N(\varepsilon, \delta)$

$$\mathbb{P}(\{x : |X_n(x) - X(\omega)| \geq \varepsilon\}) < \delta$$

for arbitarily small $\delta > 0$.

So we don't require closeness at every point, just for most point's (probability mass approaching 1).

Example 1: (Convergence both a.s & i.p)

$\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$ Borel set's

$\mathbb{P}$ = Lebesgue measure.

define

$$X_n(\omega) = \omega^n, \quad X(\omega) = \begin{cases} 0, & 0 \leq \omega < 1 \\ 1, & \omega = 1 \end{cases}$$

(a) almost sure convergence

for a fixed $\omega_0 \in [0,1)$:

$$\lim_{n\to\infty} X_n(\omega) = \lim_{n\to\infty} \omega^n = 0$$

for $\omega = 1$

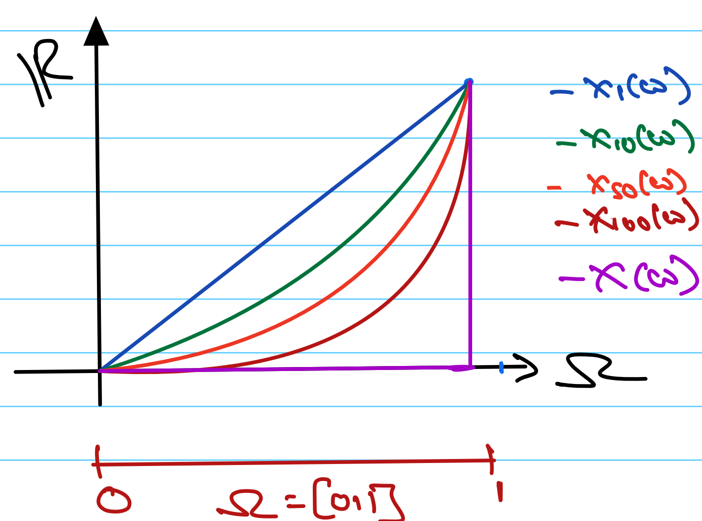
$$X_n(\omega) = 1 = X(\omega)$$

Hence

$$\forall \omega \in [0,1],\ \forall \varepsilon > 0,\ \exists N(\omega,\varepsilon) \text{ s.t } \forall n > N$$

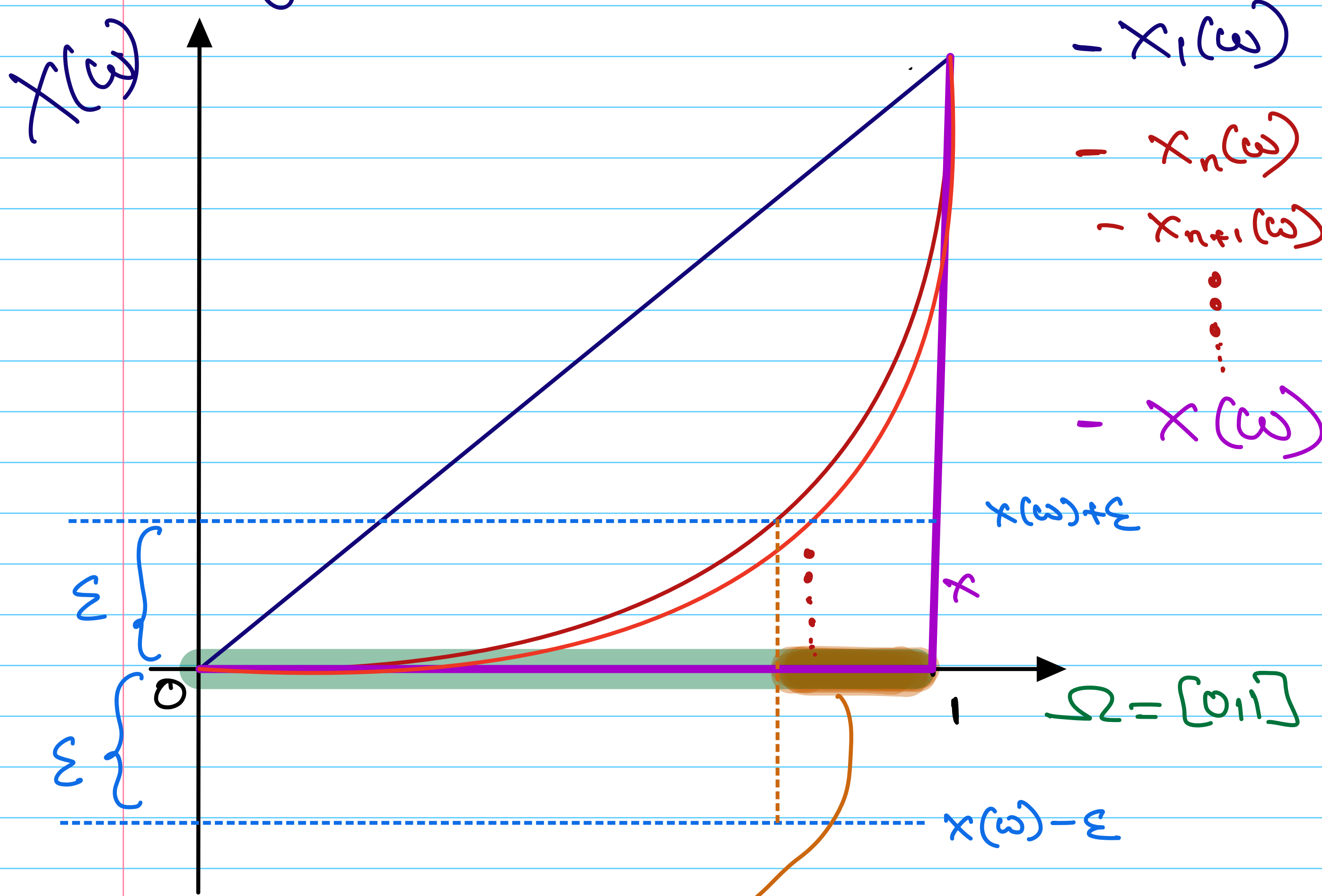
$$|X_n(\omega) - X(\omega)| < \varepsilon.$$

$X_n \longrightarrow X$ pointwise everywhere.

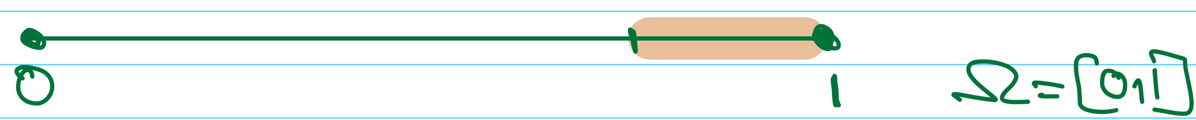


(b) Convergence in Probability.

for any $\varepsilon > 0$:



set $\{\omega \in \Omega \mid |X_n(\omega) - X(\omega)| > \varepsilon\}$



so this $\mathbb{P}(|X_n(\omega) - X(\omega)| > \varepsilon) < \delta$

this $\mathbb{P}$ Shrink's

$$\mathbb{P}(|X_n - X| > \varepsilon)$$

$$= \mathbb{P}(|\omega^n - 0| > \varepsilon)$$

$$= \mathbb{P}(\omega^n > \varepsilon)$$

$$= \mathbb{P}(\omega > \varepsilon^{1/n})$$

$$= 1 - \varepsilon^{1/n}$$

So $\forall \varepsilon > 0$, and $\forall \delta > 0$, $\exists N \in \mathbb{N}$

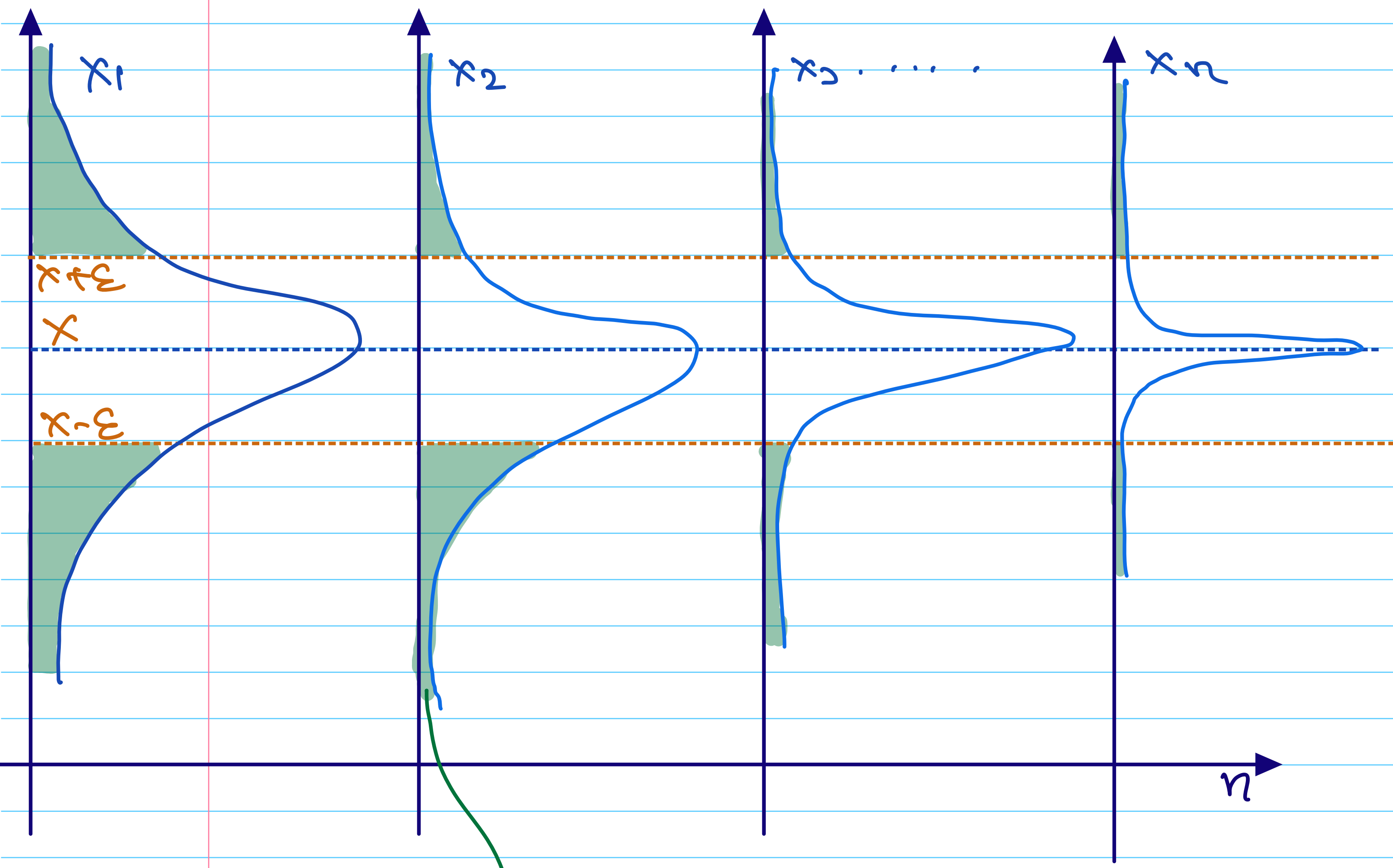
s.t $\forall n > N$ we have

$$|1 - \varepsilon^{1/n} - 0| < \delta$$

$$\Rightarrow \lim_{n \to 0} 1 - \varepsilon^{1/n} = 0$$

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

Convergence in Probability



$\mathbb{P}(|X_n - X| \geq \varepsilon)$, this value

Shrink's as $n \to \infty$

$$\Rightarrow \lim_{n \to \infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0$$

Lec 43 : Convergence of RANDOM variables - I

Definition 3: Convergence in the r^{th} mean.

We say X_n converges to X in the r^{th} mean, if

$$\lim_{n \to \infty} \mathbb{E}\left[|X_n - X|^r \right] = 0$$

For $r = 2$, X_n is said to converge to X in the mean-square sense.

Definition 4: Convergence in distribution (weak convergence)

We say X_n converges to X in distribution if

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

$\forall x$ where $F_X(\cdot)$ is continuous.

"Convergence in distribution" is not really a convergence of random variable, it's only a convergence of distribution function's.

* It's not as though X_n and X are getting closer anyway.

⇒ Convergence in almost sure, X_n and X are getting close for almost all $\omega \in \Omega$.

⇒ In this case X_n and X are very different. for any given $\omega \in \Omega$, It's not necessarily 2 at all $X(\omega)$, and $X_n(\omega)$ are very close. All that is demanded here is the CDF's Converge. The actual value of $X(\omega)$ & $X_n(\omega)$ can be very different.

for Convergence in distribution, the r.v's need not be defined on a single Probability space!

Notation

⓪ Pointwise convergence:

$$X_n \xrightarrow{P.W} X$$

① Almost sure Convergence:

$$X_n \xrightarrow{a.s} X \text{ or }$$
$$X_n \xrightarrow{w.p.1} X$$

② Convergence in Probability:

$$X_n \xrightarrow{i.P} X$$

③ Convergence in r^{th} mean:

$$X_n \xrightarrow{r} X \text{ , when } r=2 \quad X_n \xrightarrow{m.s} X$$

④ Convergence in distribution

$$X_n \xrightarrow{D} X \text{ or } X_n \Longrightarrow X$$

Theorem: (Hierarchy of convergence)

The following Implication's hold

$$
\begin{array}{ccccccc}
\text{p.w} & \Longrightarrow & \text{a.s} & \searrow & & & \\
 & & & & \text{i.p} & \Longrightarrow & \text{D} \\
 & & r^{th}\text{ mean } (r \geq 1) & \nearrow & & & \\
 & & r^{th}\text{ mean } (r \geq 1) & \Longrightarrow & s^{th}\text{ mean } (r > s \geq 1) & &
\end{array}
$$

No other Implication's hold in general.

Theorem:

$$X_n \xrightarrow{r^{th}} X \implies X_n \xrightarrow{i.p} X, \quad \forall r \geq 1$$

Proof: use markov inequality

$$\mathbb{P}\left(|X_n - X|^r > \varepsilon^r \right) \leq \frac{\mathbb{E}\left[|X_n - X|^r \right]}{\varepsilon^r}$$

$$\Rightarrow \lim_{n\to\infty} \mathbb{P}(|X_n - X|^r > \varepsilon^r) \leq \lim_{n\to\infty} \underbrace{\frac{\mathbb{E}[|X_n - X|^r]}{\varepsilon^r}}_{=0}$$

$\forall \varepsilon > 0$

$$\Rightarrow \forall \varepsilon > 0 \quad \lim_{n\to\infty} \mathbb{P}(|X_n - X|^r > \varepsilon^r) = 0$$

$$\Longrightarrow \forall \varepsilon > 0 \quad \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$$

Theorem: $X_n \xrightarrow{i.p} X \Longrightarrow X_n \xrightarrow{D} X$

Proof: fix $\varepsilon > 0$

we have $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon)$

$$F_n(x) = \mathbb{P}(X_n \leq x)$$

$$= \mathbb{P}(X_n \leq x;\ X \leq x+\varepsilon)$$

$$+ \mathbb{P}(X_n \leq x;\ X > x+\varepsilon)$$

we have

$$\mathbb{P}(X_n \le x;\ X \le x+\varepsilon) \le P_X(X \le x+\varepsilon)$$
$$\le F_X(x+\varepsilon)$$

& $\mathbb{P}(X_n \le x,\ X > x+\varepsilon)$
$$\le \mathbb{P}(|X_n - X| > \varepsilon)$$

$\Rightarrow$

$$F_n(x) \le F_X(x+\varepsilon) + P(|X_n - X| > \varepsilon)$$

Simillarly:

$$F_X(x-\varepsilon) = P(X \le x-\varepsilon)$$
$$\le F_{X_n}(x) + P(|X_n - X| > \varepsilon)$$

Thus

$$F_X(x-\varepsilon) - \mathbb{P}(|X_n - X| > \varepsilon) \le F_{X_n}(x) \le F_X(x+\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon)$$

as $n \to \infty$ since $X_n \xrightarrow{i.p} X$

$$\Rightarrow P(|X_n - X| > \varepsilon) \to 0$$

$$\Rightarrow \lim_{n\to\infty} F_X(x-\varepsilon) \leq \liminf_{n\to\infty} F_{X_n}(x) \leq \limsup_{n\to\infty} F_{X_n}(x)$$

$$\leq F_X(x+\varepsilon)$$

$$\forall \varepsilon > 0$$

Example:

Convergence in Distribution does not mean convergence in Probability

Sol^n

Let $X_1, X_2, \ldots$ be such that $X_i = X$ $\forall i \geq 1$, $X \sim Ber(1/2)$

$$\Rightarrow \mathbb{P}_X(X=0) = 1 - 1/2 = 1/2$$

$$\mathbb{P}_X(X=1) = 1/2$$

Let $Y = 1 - X$

$$P_Y(Y=0) = P_X(X=1) = P$$

$$P_Y(Y=1) = P_X(X=0) = 1-p$$

So clearly $X_n \xrightarrow{D} Y$

But $|X_n - Y| = 1$

But there is partial converse that is true. The converse hold's in a special case.

* Suppose X_n converges to a constant o.r.v C, then i.p $\Longleftrightarrow$ D

$X_n \xrightarrow{D} C$ where C is constant

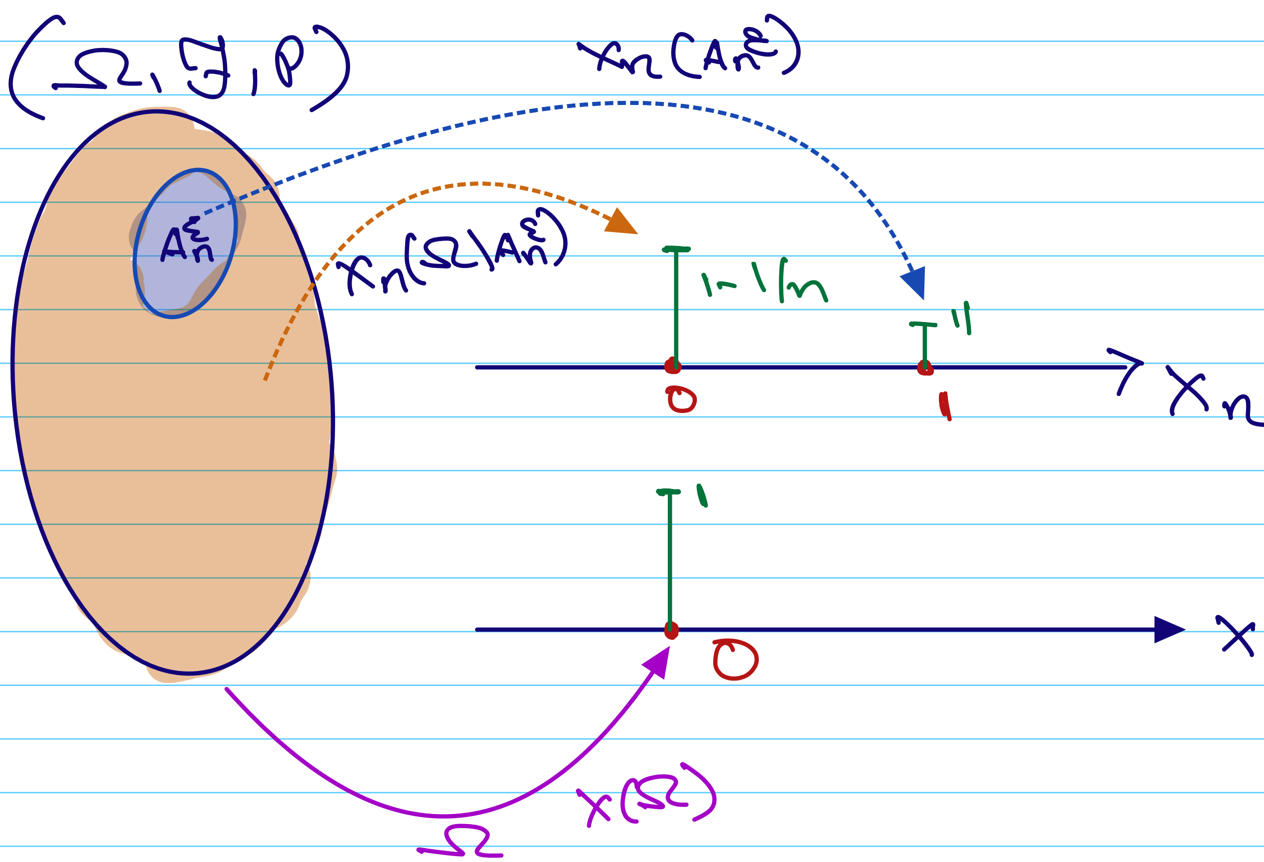
then $X_n \xrightarrow{i.p} C$

Theorem : $X_n \xrightarrow{i.p} X \not\Longrightarrow X_n \xrightarrow{a.s} X$

Proof: $(X_n)_{n \in \mathbb{N}}$ seq^n of r.v's.

$$X_n = \begin{cases} 1 & \text{w.p } \frac{1}{n} \\ 0 & \text{w.p } 1-\frac{1}{n} \end{cases}$$

$$X = \begin{cases} 1 & \text{w.p } 0 \\ 0 & \text{w.p } 1 \end{cases}$$



Clearly

$$\mathbb{P}(|X_n - 0| > \varepsilon)$$

$$= \mathbb{P}(|X_n| > \varepsilon)$$

$$\Rightarrow \mathbb{P}(X_n = 1)$$

$$\Rightarrow \lim_{n \to \infty} \mathbb{P}(X_n = 1) = \lim_{n \to \infty} \frac{1}{n}$$

$$= 0$$

$\Rightarrow$ Convergence in probability