

Convergence of random variables

LEC 42

Convergence of r.v.'s is probably the most challenging part of the course.

if we are given the seqⁿ of real numbers

$(a_n)_{n \in \mathbb{N}}$ we have convergence of

numbers.

- * we are going to talk about convergence of r.v.'s. (seqⁿ of r.v.'s)

(Ω, \mathcal{F}, P) Probability space

x_1, x_2, x_3, \dots . seqⁿ of r.v's

If random variable are function's

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

so, we are essentially talking about convergence
of function's.

Definition 0 :-

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

Pointwise convergence of function's

from Real Analysis: Pointwise Convergence

Seq. of functions: $(f_1, f_2, f_3, f_4, \dots)$

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

Pointwise Convergence:

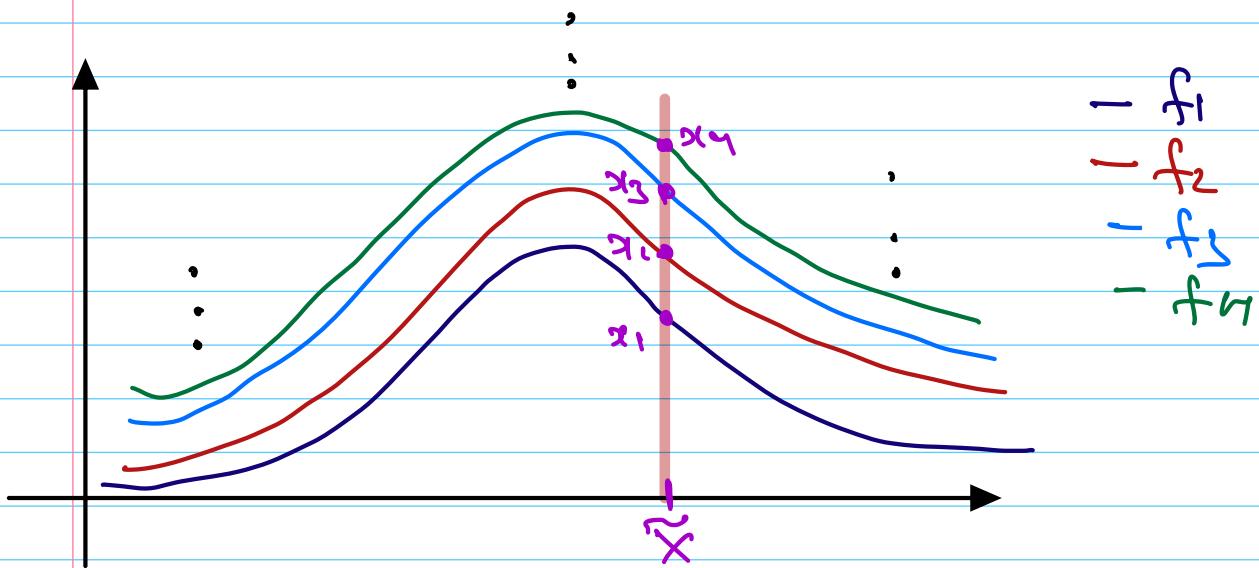
$(f_1, f_2, f_3, f_4, \dots)$ in

Pointwisely convergent to a function

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

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

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



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

Back to Convergence of RV's

$$X_n(\omega) \xrightarrow{n \rightarrow \infty} X(\omega) \quad (\text{Pointwise convergence})$$

of RV's

We pick any $\omega \in \Omega$ in sample space, for that ω we have a seqⁿ of real numbers

$$(X_n(\omega))_{n \in \mathbb{N}} \quad (\text{Seq}^n \text{ of real numbers})$$

As soon as ω realized (picked by mother god)

The seqⁿ of random variables becomes

seqⁿ of real numbers.

$$\Rightarrow (X_n)_{n \in \mathbb{N}} \quad (\text{Seq}^n \text{ of real numbers})$$

Converges to $X(\omega)$

$\Rightarrow \forall \omega \in \Omega, \exists \varepsilon > 0, \exists N \in \mathbb{N}$ such that

$N > N$

we have $|X_n(\omega) - X(\omega)| < \varepsilon$

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

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

This is two stronger notion of Probability.

because in Probability, we are willing to sacrifice Prob measure O sets. we don't really care about Probability measure O 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

$X_n(\omega) \xrightarrow{n \rightarrow \infty} X(\omega)$ on a set of

Prob 1.

That is

$$P(\{\omega \mid X_n(\omega) \rightarrow x(\omega)\}) = 1$$

The set of all ω 's $\in \Omega$.

* This convergence happens 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 seqⁿ 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

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - x| > \varepsilon) = 0$$

* Convergence in Prob in very different form
Convergence with Probability 1.

* in Sure Convergence and almost sure
Convergence, in these 2 cases, we have
Seq of rv themselves converging

* in convergence in Probability, it is
bit of misnomer to say $X_n \rightarrow X$
in Probability.

This is just that some seqⁿ of
Prob is converging to 0.

(The seqⁿ of Probabilities
converges)

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

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 $|X_n - x| \geq \varepsilon$,
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 $P(|X_n - x| \geq \varepsilon)$ measures
how "large" this event is in terms
of the probability measure P .

(iii) Convergence in Prob means that
this prob goes to 0 as $n \rightarrow \infty$

② Breakdown of the definition

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

Then

$$P(|X_n - x| \geq \varepsilon) = P(A_n^\varepsilon)$$

Convergence in Prob mean

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

Example's:-

Let (Ω, \mathcal{F}, P) be a Probability space.

EACH R.V $X_n: \Omega \rightarrow \mathbb{R}$ 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

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

equivalently, $\forall \varepsilon > 0$

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

Connection to Real analysis (Pointwise Convergence)

In Real Analysis $f_n \rightarrow f$ pointwise on X
means:

$$\forall x \in X, \forall \epsilon > 0, \exists N(x, \epsilon) \in \mathbb{N} \text{ s.t } \forall n > N(x, \epsilon)$$

we have

$$|f_n(x_0) - f(x_0)| < \epsilon$$

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

That is

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

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

② Convergence in Probability

Definition:

We say X_n converges in Probability to X if

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0$$

Connection to Real Analysis (Uniform in measure)

In Real Analysis, "Uniform Convergence" says:

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

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

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

Convergence in Probability relates with

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

$$P\left(\{x : |X_n(x) - X(\omega)| > \varepsilon\}\right) < \delta$$

for arbitrarily small $\delta > 0$.

so we don't require closeness at every point, just for most points (Probability mass approaching 1).

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

$$\Omega = [0,1], \mathcal{F} = \mathcal{B}([0,1]) \text{ Borel sets}$$

P = Lebesgue measure.

define

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

Ⓐ almost sure convergence

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

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

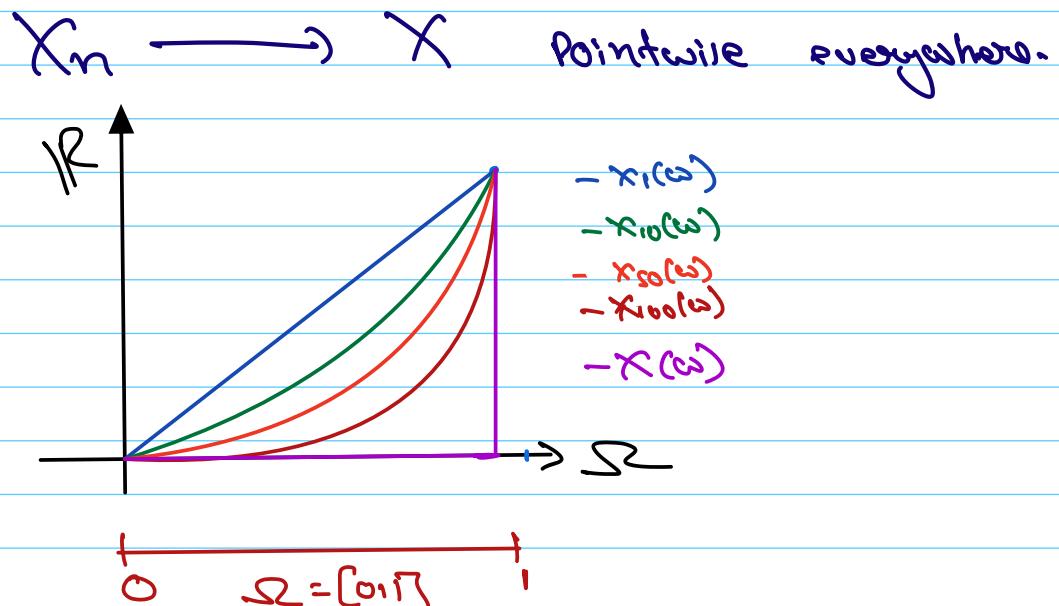
for $\omega = 1$

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

Hence

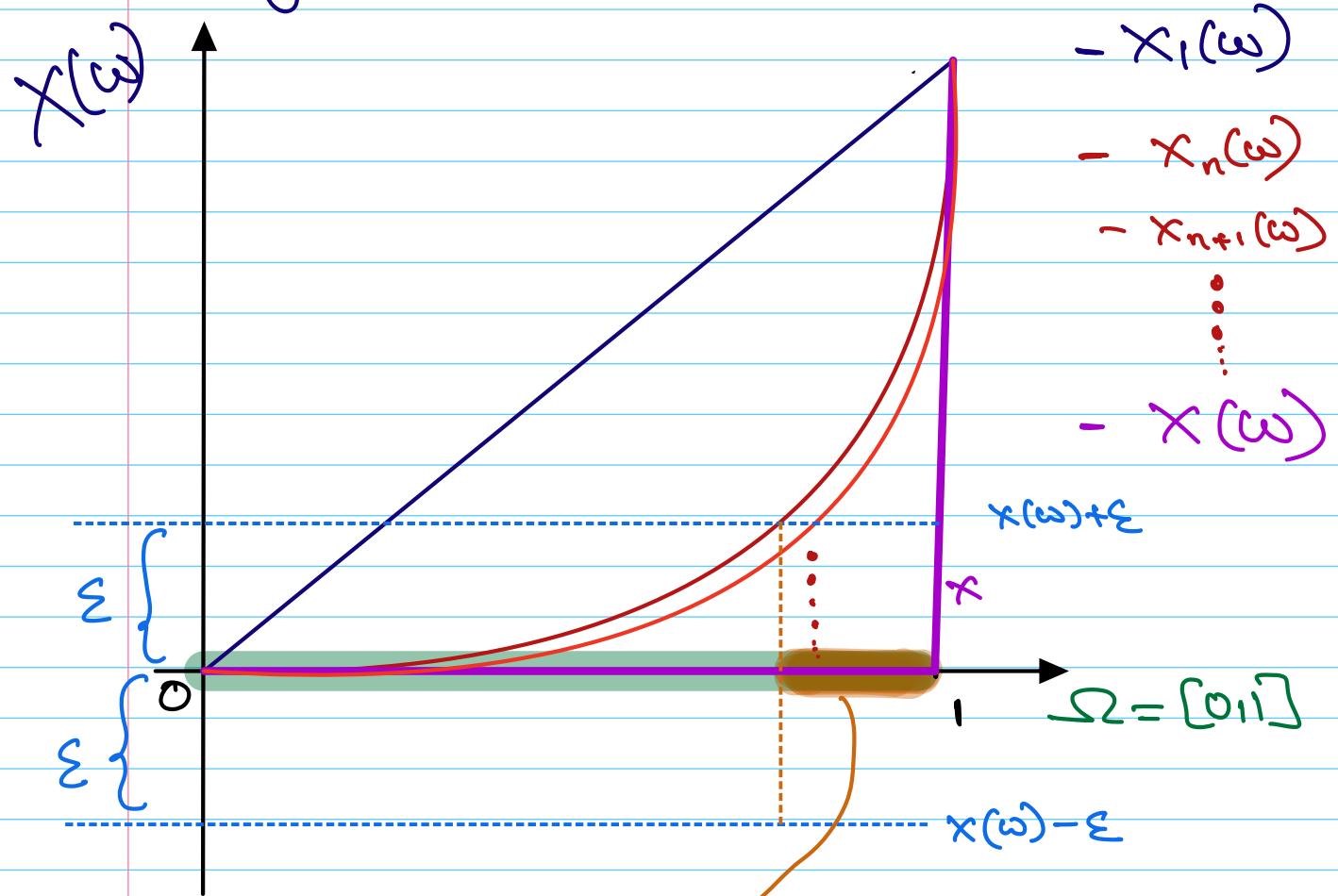
$\forall \omega \in [0, 1], \forall \varepsilon > 0, \exists N(\omega, \varepsilon)$ s.t. $\forall n > N$

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

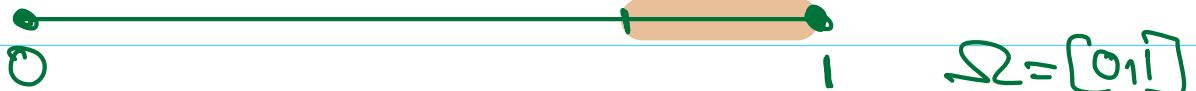


(b) Convergence in Probability.

for any $\varepsilon > 0$:



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



so this $IP(|X_n(\omega) - X(\omega)| > \varepsilon) < \delta$

this IP Shows it's

$$\text{IP}(|X_n - X| > \varepsilon)$$

$$= \text{IP}(|\omega^n - 0| > \varepsilon)$$

$$= \text{IP}(\omega^n > \varepsilon)$$

$$= \text{IP}(\omega > \varepsilon^{1/n})$$

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

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

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

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

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

$$\Rightarrow \lim_{n \rightarrow \infty} \text{IP}(|X_n - X| > \varepsilon) = 0$$

LEC 43 : Convergence of RANDOM variables - I

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

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

$$\lim_{n \rightarrow \infty} E[|X_n - X|^\alpha] = 0$$

for $\alpha \geq 2$ 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

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

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

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

○ 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 z 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

① Pointwise Convergence :

$$X_n \xrightarrow{\text{P.w}} X$$

② Almost Sure Convergence :

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

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

③ Convergence in Probability :

$$X_n \xrightarrow{\text{i.p.}} X$$

④ Convergence in σ^2 mean:

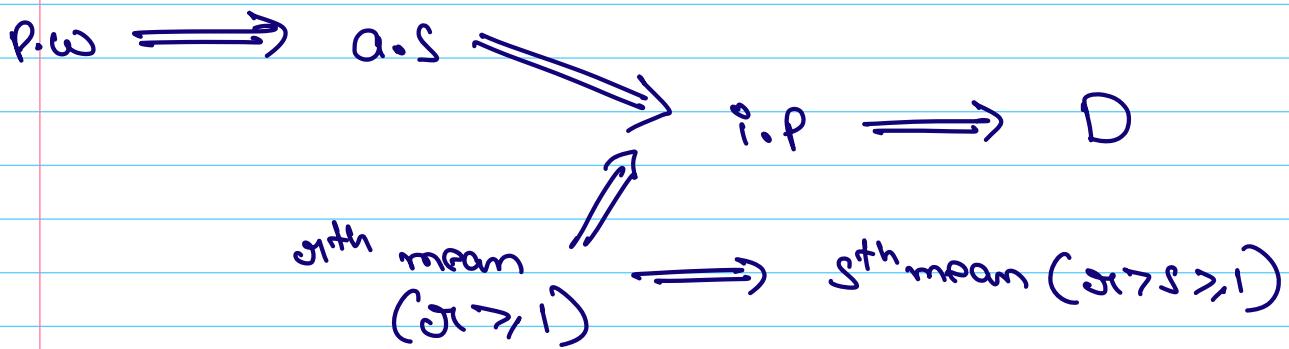
$$X_n \xrightarrow{\sigma} X, \text{ when } \sigma = 2 \quad X_n \xrightarrow{\text{m.s.}} X$$

⑤ Convergence in distribution

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

Theorem: (Hierarchy of convergence)

The following Implications hold



No other Implications hold in general.

Theorem:

$$X_n \xrightarrow{\alpha^{1/n}} X \Rightarrow X_n \xrightarrow{\alpha \cdot p} X, \forall \alpha > 1$$

Proof: USE markov inequality

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

$$\Rightarrow \lim_{n \rightarrow \infty} \text{IP}(|X_n - x|^\alpha > \varepsilon^\alpha) \leq \lim_{n \rightarrow \infty} \frac{\mathbb{E}[|X_n - x|^\alpha]}{\varepsilon^\alpha} \xrightarrow{\varepsilon \downarrow 0} 0$$

$$\Rightarrow \forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} \text{IP}(|X_n - x|^\alpha > \varepsilon^\alpha) = 0$$

$$\Rightarrow \forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} \text{IP}(|X_n - x| > \varepsilon) = 0$$

Theorem:

$$X_n \xrightarrow{i.p} x \implies X_n \xrightarrow{D} x$$

Proof: fix $\varepsilon > 0$

we have

$$\lim_{n \rightarrow \infty} \text{IP}(|X_n - x| > \varepsilon)$$

$$F_n(x) = \text{IP}(X_n \leq x)$$

$$= \text{IP}(X_n \leq x; x \leq x + \varepsilon)$$

$$+ \text{IP}(X_n \leq x; x > x + \varepsilon)$$

we have

$$\begin{aligned} \text{IP}(x_n \leq x; X \leq x+\varepsilon) &\leq P_x(X \leq x+\varepsilon) \\ &\leq F_x(x+\varepsilon) \end{aligned}$$

$$\begin{aligned} \text{IP}(x_n \leq x, X > x+\varepsilon) &\leq \text{IP}(|x_n - x| > \varepsilon) \end{aligned}$$

→

$$F_n(x) \leq F_x(x+\varepsilon) + P(|x_n - x| > \varepsilon)$$

Similarly:

$$\begin{aligned} F_x(x-\varepsilon) &= P(X \leq x-\varepsilon) \\ &\leq F_{x_n}(x) + P(|x_n - x| > \varepsilon) \end{aligned}$$

Thus

$$\begin{aligned} F_x(x-\varepsilon) - \text{IP}(|x_n - x| > \varepsilon) &\leq F_{x_n}(x) \leq F_x(x+\varepsilon) \\ &\quad + \text{IP}(|x_n - x| > \varepsilon) \end{aligned}$$

Q) $n \rightarrow \infty$ since $x_n \xrightarrow{P} x$

$$\Rightarrow P(|x_n - x| > \varepsilon) \rightarrow 0$$

$$\Rightarrow \lim_{n \rightarrow \infty} F_x(x-\varepsilon) \leq \liminf_{n \rightarrow \infty} F_{x_n}(x) \leq \limsup_{n \rightarrow \infty} F_{x_n}(x) \\ \leq F_x(x+\varepsilon)$$

Q.E.D

Example:

Convergence in Distribution does not mean Convergence in Probability

Solⁿ

Let x_1, x_2, \dots be such that $x_i = x$
 $\forall i \geq 1$, $x \sim \text{Ber}(1/2)$

$$\Rightarrow P_x(x=0) = 1/2 = 1/2$$

$$P_x(x=1) = 1/2$$

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 holds in a special case.

* Suppose X_n converges to a constant $c \in C$, then $i.p \iff D$

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

then

$X_n \xrightarrow{i.p} c$

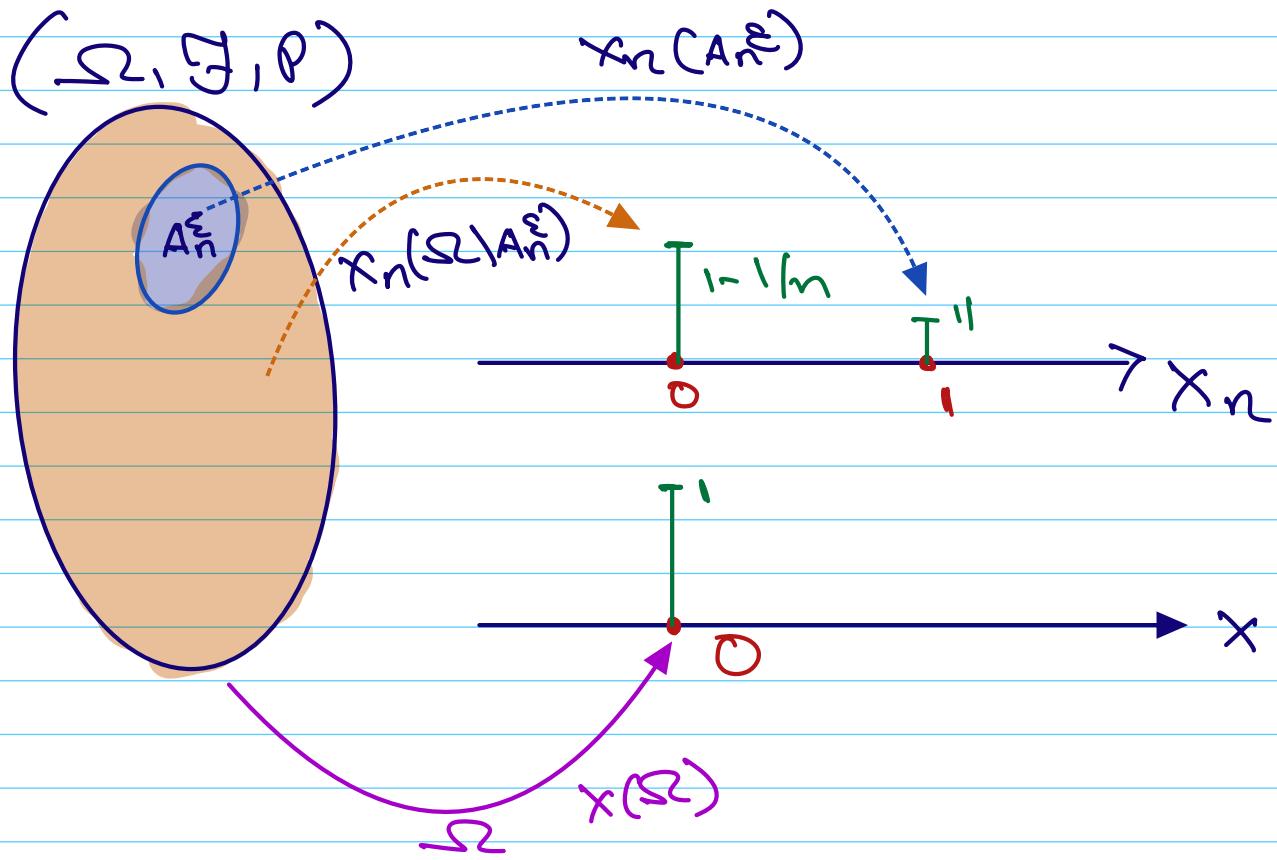
Theorem : $x_n \xrightarrow{i.p} x \Leftrightarrow x_n \xrightarrow{a.s} x$

Proof:

$(x_n)_{n \in \mathbb{N}}$ seqⁿ of g.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

$$\text{IP}(\lim_{n \rightarrow \infty} X_n = 1) > \varepsilon$$

$$= \text{IP}(|X_n| > \varepsilon)$$

$$\Rightarrow \text{IP}(X_n = 1)$$

$$\Rightarrow \lim_{n \rightarrow \infty} \text{IP}(X_n = 1) = \lim_{n \rightarrow \infty} \frac{1}{n}$$
$$= 0$$

\Rightarrow Convergence in probability