Probability

Random variable: $X : (\Omega, \mathcal{F}, P) \to (E, \xi)$

$$X^{-1}(A) = \langle \omega : X(\omega) \in A \} \in \mathcal{F} \quad \forall A \in \xi.$$

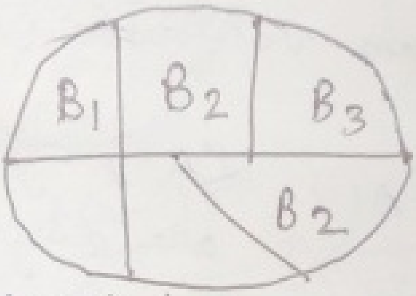
$\sigma(X)$ = smallest sub-σ-field w.r.t. which X is measurable ~~which~~

= intersection of all sub-σ-fields of $\mathcal{F}$ w.r.t which X is measurable

$= \langle X^{-1}(A) : A \in \xi \rangle$.

Suppose X takes values $\langle a_1, a_2, \ldots a_n \rangle$.

Let $B_i = \langle \omega : X(\omega) = a \rangle$



Suppose Y is a real random variable, Y is $\sigma(X)$-measurable:

$\langle \omega : Y(\omega) \in A \rangle \in \sigma(X)$ for all Borel sets $A \subset R$.

$\Leftrightarrow \sigma(Y) \subset \sigma(X)$

(what we can figure out from Y, we can do that from X also)

$\langle \omega : Y(\omega) \in A \rangle$ is of the form $\langle \omega : X(\omega) \in C \rangle$

Fact: $\exists$ a measurable $h : E \to R$ s.t. $Y = h(X)$.

Examples of random variables:

(i) Indicator function: $X(\omega) = I_A(\omega)$, $A \in \mathcal{F}$.

$$\triangleq \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A. \end{cases}$$

(Characteristic function) in measure theory, but NOT in probability.

χ_A is the characteristic function in case of probability.

(ii) Simple function: $X(\omega) = \sum_{i=1}^{N} a_i I_{A_i}$, $A_i \in \mathcal{F}$ disjoint.

(values of X are a_i's)

(iii) Elementary random variable:

$$X(\omega) = \sum_{i=1}^{\infty} a_i I_{A_i}, \quad \{A_i\} \subset \mathcal{F} \text{ are disjoint.}$$

Fact: Any r.v. X can be approximated pointwise by simple random variables and uniformly by elementary random variables.

Def^n: (i) $f_n \to f$ pointwise $\iff f_n(x) \to f(x)\ \forall x$.

(ii) $f_n \to f$ uniformly $\iff \sup_x |f_n(x) - f(x)| \to 0$

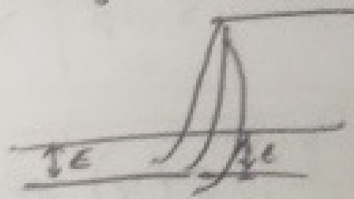
(i) $\iff$ Given $\epsilon > 0$, for any x, $\exists\, n_0 \geq 0$ s.t.

$|f_n(x) - f(x)| < \epsilon$ for $n \geq n_0$ (n_0 depends on x here)

(ii) $\iff$ Given $\epsilon > 0$, $\exists\, n \geq 0$ s.t.

$|f_n(x) - f(x)| < \epsilon$ for all x, $n \geq n_0$

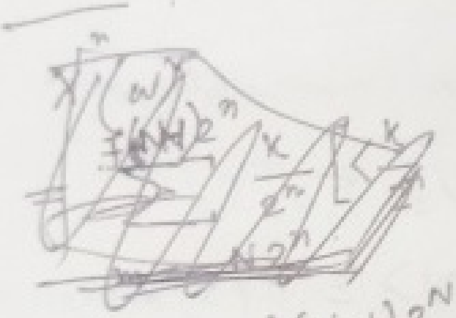
(n_0 does not depend on x, convergence is at similar rate at all points)



Convergence pointwise but not uniformly

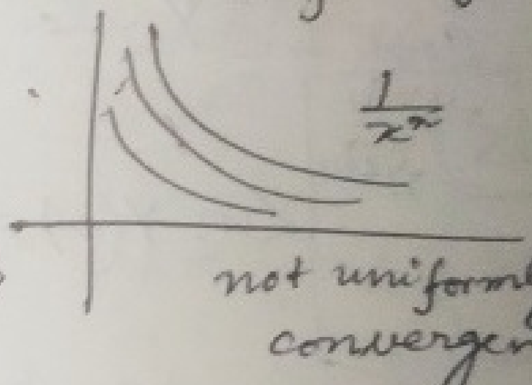
Real r.v.s.

$X(\omega)$ real-valued r.v.



$I\langle\,\rangle$

$I\langle\cdot\rangle$



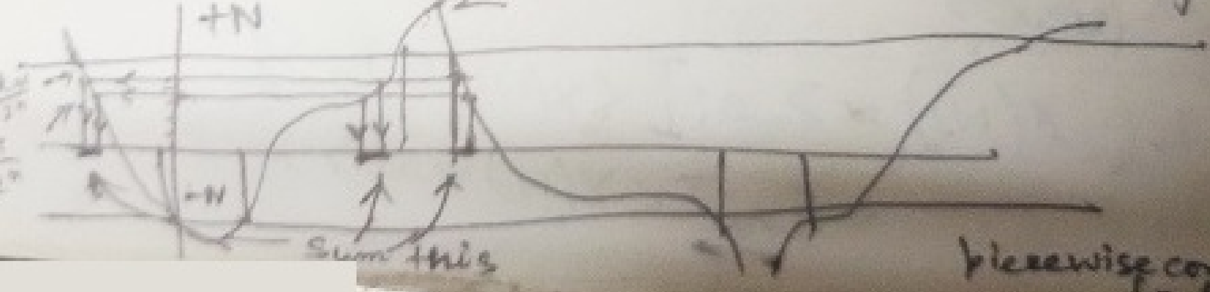
$$X^n(\omega) = \sum_{k=-N2^n}^{+(N-1)2^n} \frac{k}{2^n}\, I\left\langle \frac{k}{2^n} \leq X(\omega) < \frac{k+1}{2^n} \right\rangle$$

not uniformly convergent

→ Simple.

takes values $\frac{k}{2^n}$, $-N2^n \leq k < N2^n$

In bounded function no truncation will be required



Piecewise const approx f^n from

It will not in general be uniform convergence

By increasing N we are admitting more & more pts to come

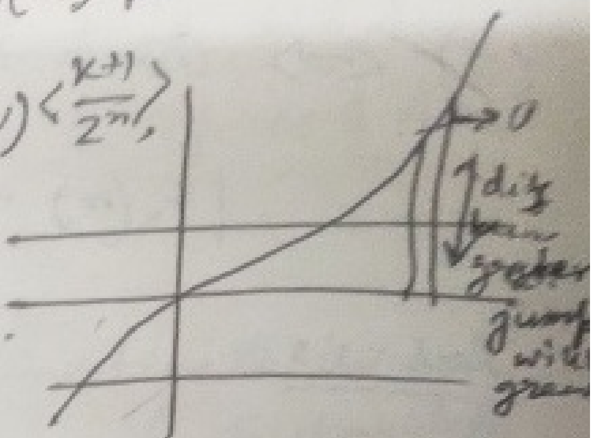
n we are decreasing the intervals.

$X^n(\omega) \uparrow X(\omega)$ pointwise.

$$\tilde{X}^n(\omega) = \sum_{k=-\infty}^{\infty} \frac{k}{2^n} I\left\langle \frac{k}{2^n} \le X(\omega) < \frac{k+1}{2^n} \right\rangle$$

$\tilde{X}^n$ elementary

$$\sup_{\omega} |\tilde{X}^n(\omega) - X(\omega)| < \frac{1}{n}.$$



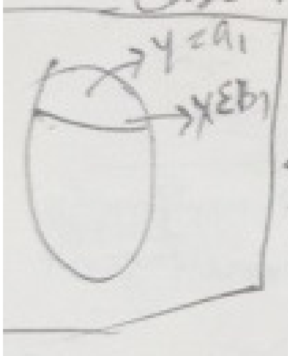
Cannot approximate uniformly

$\boxed{Y \text{ real}, \ \sigma(Y) \subset \sigma(X)}$

$\Rightarrow Y = h(X)$ for some measurable ~~X~~.

Case 1: Y elementary, $Y = \sum a_i I_{A_i}$,

$\langle A_i \rangle \in \sigma(X)$ disjoint.



$\langle \omega : Y(\omega) = a_i \rangle = A_i = \langle \omega : X(\omega) \in B \rangle$,

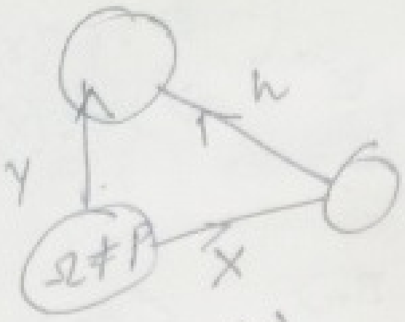
$\langle B_i \rangle$ disjoint.

Let $h = \sum a_i I_{B_i}$

$h(\cdot) = \sum a_i I_{B_i}$, $h(x) = \sum_i a_i I\langle x \in B_i \rangle$

$h(X(\omega)) = \sum a_i I\langle X(\omega_i) \in B_i \rangle$

$= \sum a_i I\langle Y(\omega) = a_i \rangle = Y$.



If $Y = h(X)$

then $\{w : Y(w) \in A\} = \{w : h(X(w)) \in A\}$

$= \{w : X(w) \in h^{-1}(A)\} \in \sigma(X)$.

General Y: take Y_n elementary s.t.

$Y_n(w) \uparrow Y(w)$ as above.

$\sigma(Y_n) \subset \sigma(X) \ \forall n$.

$\{w : Y_n = \frac{k}{2^n}\} = \{w : \frac{k}{2^n} \le Y \le \frac{k+1}{2^n}\}$

$\in \sigma(X)$.

$\therefore Y_n = h_n(X)$ Define $h = \lim_n h_n$ where the limit exists.

$= 0$ otherwise.

$\lim_{n \to \infty} h_n$ exists on $\{x : X(w) = x$ for some $w\}$

$Y = h(X)$.

$(\because Y \in \sigma(Y) \Rightarrow Y \in \sigma(X))$

~~$h_n(X) = \frac{k}{2^n}$~~

$$\langle w : \frac{k}{2^n} \le Y < \frac{k+1}{2^n} \rangle$$

$$= \langle w : X(w) \in B_n \rangle$$

$\langle B_n \rangle$ disjoint.

Fix w, look at $X(w), Y(w)$.

$Y_n(w) = h_n(X(w)) = \frac{k}{2^n}$, if $Y(w) \in [\frac{k}{2^n}, \frac{k+1}{2^n}]$.

as $n \to \infty$, $Y_n(w) \to Y(w)$

$\therefore h_n(X(w)) \to Y(w)$ (the limit converges).

The limit is measurable

$$h = \lim_n h_n \; I \underbrace{\langle \lim_n h_n \text{ exists} \rangle}_{H}.$$

$$\tilde{h}_n = h_n I_H$$

Product of two measurable fns will be measurable.

The set is measurable, that we have to show!

$\langle x : \lim_{n} h_n \text{ exists} \rangle$

$= \bigcap_{k \geq 1} \bigcup_{n \geq 0} \bigcap_{j \geq n} \bigcap_{m \geq 0} \langle |h_j(x) - h_{j+m}(x)| < \frac{1}{k} \rangle$

$\bigcap$ for all $\quad \bigcup$ there exists

$|h_n(x) - h_m(x)| \to 0$

Cauchy convergence

$\langle A_n \rangle$ infinitely often

$\iff \bigcap_{n=1}^{\infty} \bigcup_{m \geq n} A_m$, i.o

$\bigcup \to$ or

$\to$ infinitely often

$\Downarrow$

$1 \quad -1 \quad 1 \quad -1$

↳ 1 infinitely often

↳ not 1 eventually

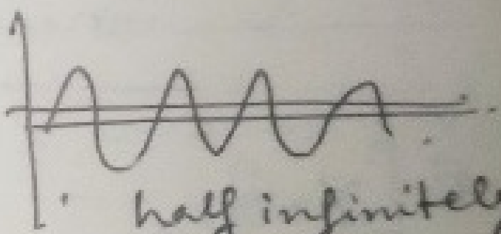
$\langle A_n \rangle$ eventually $\iff$

$\bigcup_{n=1}^{\infty} \bigcap_{m \geq n} A_m$,

(after some time all will contain)

infinitely often

$\equiv \sim$ eventually



half infinitely often

not eventually

$\sigma(X_1, \ldots X_n)$

Expectation:

$E[X]$, $\int X(\omega)\, dP(\omega)$,

$\int X(\omega)\, P(d\omega)$

$M[X]$.

(Average is a more general word)

Mean

$$X=\sum_{i=1}^{n} a_i I_{A_i}, \quad \langle A_i \rangle \text{ disjoint.}$$

$$E[X]=\sum_{i=1}^{n} a_i P(A_i)$$

$$=\sum_{i=1}^{n} a_i P(X=a_i)$$

$$E[I_A]=0\cdot P(A^c)+1\cdot P(A)=P(A),$$

so probability is a special case of expectation

$X_1, X_2, \dots X_\tau$

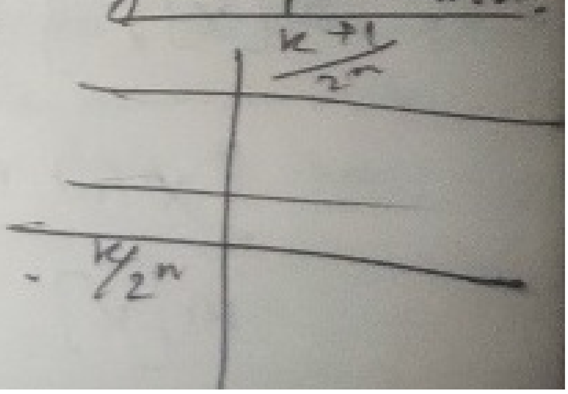
$\tau=n$ should be measurable

maximize $E[X_\tau]$ → optimal stopping problem

$$X\geq 0, \text{ let } X_n = \sum_{m=0}^{n2^n} \frac{k}{2^n} I\left\langle \frac{k}{2^n}\leq X<\frac{k+1}{2^n}\right\rangle.$$

Now X_n converge increasing to X (X_n is an increasing sequence).

New partition is subdivision of old partition.



$$E[X_n]$$

$$=\sum_{k=0}^{n2^n} \frac{k}{2^n} P\left(\frac{k}{2^n}\leq X<\frac{k+1}{2^n}\right)$$

X_n is increasing for every ω

$\frac{k}{2^n} P\left(\frac{k}{2^n} \le X < \frac{k+1}{2^n}\right) \to$

$\frac{2k}{2^{n+1}} P\left(\frac{2k}{2^{n+1}} \le X < \frac{2k+1}{2^{n+1}}\right)$

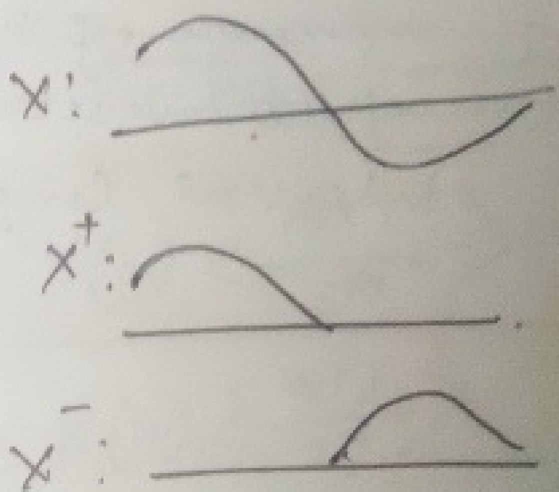
$\to \frac{2k+1}{2^{n+1}} P\left(\frac{2k+1}{2^{n+1}} \le X \le \frac{2k+2}{2^{n+1}}\right)$

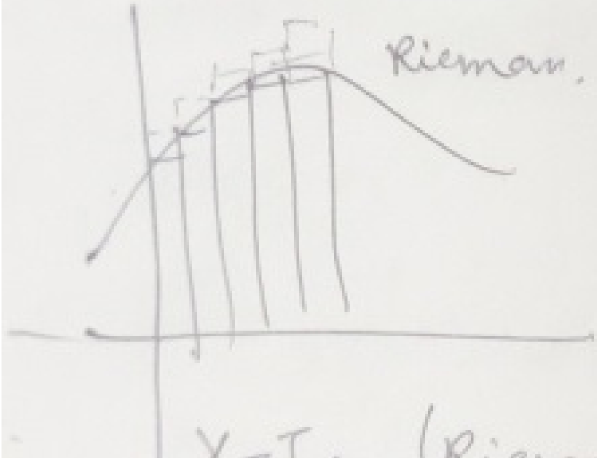
$E[X] = \lim_{n \to \infty} E[X_n]$ (can be $+\infty$]

General X,

$X^+ = \max(X, 0)$.

$X^- = -\min(X, 0)$





Rieman.

$X = X^+ - X^-$, $X^+, X^- \geq$

$E[X] = E[X^+] - E[X^-]$

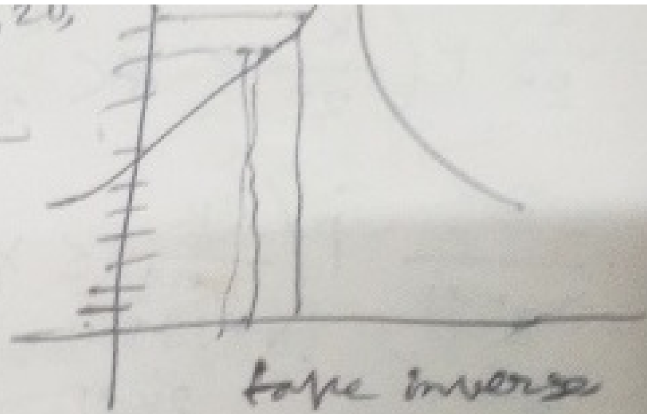
when not both $E[X^+], E[X^-]$ are ∞

(both are finite, o w $E[X]$ is not defined).

$X = I_{\mathbb{Q}}$ (Rieman's integrals fail, because sup is always 1, inf " " 0)

5, 10, 5, 20, 10, 20, 5, 10, 20, 20,

5, 20, 10



take inverse images

This is known as lebesgue integral

we partition the range & take inverse image. take the lower bound and multiply it with the measure of summation of the inverse image interval lengths.

$X = I_{\mathbb{Q} \cap [0,1]}$.

$E[X]$

$= 1 \cdot P(\mathbb{Q} \in [0,1])$

$+ 0 \cdot P(\mathbb{Q}^c \in [0,1])$

$= 0$