Probability 24.08.05.

$(\Omega, \mathcal{F}, P) \xrightarrow{X} \mathbb{R}$

Indicator: $X = I_A(\omega) = 1$, if $\omega \in A$,
$= 0$, ow

Simple: $X = \sum_{i=1}^{n} a_i I_{A_i}$ $\langle A_i \rangle$ disjoint.

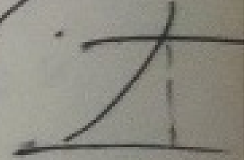
Elementary: $X = \sum_{i=1}^{\infty} a_i I_{A_i}$ $\langle A_i \rangle$ disjoint

(approx. take by elementary function)

$X \geq 0$

$$X^n = \sum_{m=0}^{n2^n} \frac{m}{2^n} I \left\langle \frac{m}{2^n} \leq X(\omega) < \frac{m+1}{2^n} \right\rangle$$

$$\tilde{X}^n = \sum_{m=0}^{\infty} \frac{h}{2^n} I \left\langle \frac{m}{2^n} \leq X(\omega) < \frac{m+1}{2^n} \right\rangle.$$

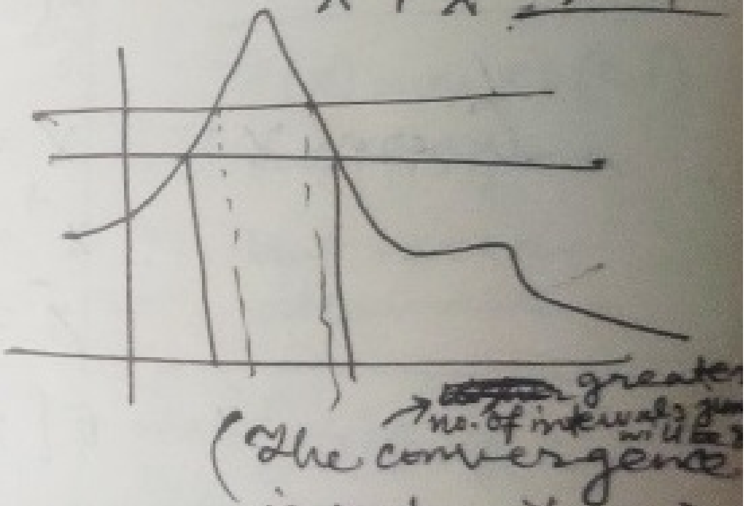


$X \uparrow X$

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

$\Leftrightarrow \langle \omega : Y(\omega) \in A \rangle$
$= \langle \omega : X(\omega) \in A' \rangle$

$\Rightarrow Y = h(X)$



greater no. of intervals will

(The convergence is not uniform)

$\langle \omega : Y(\omega) = a \rangle$
$= \langle \omega : X(\omega) \in A' \rangle$
$h(\omega) = a, \ \omega \in A'$

(since the fn is bounded) After some n on it will be uniform convergence

$Y^n \uparrow Y$
$Y^n = h_n(X)$ } whatever the limit is call that h_n.

$h = \lim_n h_n$, when possible
arbitrary, outside the set.

Expectation

$X \geq 0$ $\quad X = \sum_{i=1}^{n} a_i I_{A_i}$ $\langle A_i \rangle$ disjoint.

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

$$E[I_A] = P(A).$$

$X \geq 0$, $X^n \geq 0$ simple $\uparrow X$,

$$E(X) = \lim_n E[X^n]. \text{ (can be } +\infty)$$

General X: $\quad X = X^+ - X^-$ where

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

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

when $\min(E[X^+], E[X^-]) < \infty$

$X \geq 0$

$E[X] = \sup E[\psi]$.

$0 \leq \psi \leq X$
simple

(take approx. from below)

Properties: (1) $E[\alpha X] = \alpha E[X]$
$\forall \alpha \in R$

(2) $E[X+Y] = E[X] + E[Y]$.

Pf ψ, ϕ simple, $\psi \leq X$,
$\phi \leq Y \Rightarrow$
$\psi + \phi \leq X + Y$.

~~$E[X+Y]$~~
$\sup E[\psi + \phi]$

$\leq \sup\limits_{\zeta \text{ simple} \leq X+Y}$

$E[\zeta] \leq E[X+Y]$.

$X, Y \geq 0$
(~~expectation~~ ~~from~~ ~~to~~ by linearity)

$$\sup_{\psi, \phi} E[\psi + \phi] = \sup_{\phi} E[\phi] + \sup_{\psi} E[\psi]$$

$$= E[X] + E[Y]$$

$\max\limits_{x \in A} f(x) = y$.

There is $x^* \in A$ s.t.
$f(x^*) \geq f(x)$
$\forall x \in A$

$\sup\limits_{x \in A} f(x) = y$
$\Rightarrow y \geq f(x)\ \forall x \in A$ &
$x \in A$.
$y' \geq f(x)\ \forall x \in A$
$\Rightarrow y' \geq y$.

$f(x) = x,\ x \in (0,1)$
$\sup(f(x)) = 1$
but never attainable

$\max\limits_{x,y} (f(x) + g[y])$
$= \max\limits_{x,y} f(x) + \max\limits_{x,y} (g(y)$

$\therefore E[X+Y] \geq E[X] + E[Y]$.

Let $\zeta \leq X+Y$

Let $\psi = \min(X, \zeta)$, $\phi = \zeta - \psi$.

$\psi \leq X$.

If $X \leq \zeta$, then $\psi = X$ &

$$\phi = \zeta - \psi = \zeta - X$$
$$\leq X + Y - X$$
$$= Y$$

$\zeta < X$, then $\psi = \zeta$, & $\phi = \zeta - \psi$
$= 0 \leq Y$.

$\therefore \zeta = \psi + \phi$

$\psi \leq X$

$\phi \leq Y$.

$$\therefore E[\zeta] = E[\psi_\zeta] + E[\phi_\zeta]$$

subscript denotes dependence on ζ.

$$\therefore E[X+Y] = \sup_{\zeta} E[\zeta] = \sup_{\zeta} \left[E[\psi_{\zeta}] + E[\varphi_{\zeta}] \right] \quad \text{(by linearity)}$$

$$\le \sup_{\zeta} E[\psi_{\zeta}] + \sup_{\zeta} [\varphi_{\zeta}]$$

$$\le \sup_{\substack{\varphi \le X \\ \text{simple}}} E[\psi] + \sup_{\substack{\psi \le Y \\ \text{simple}}} E[\varphi].$$

$$= E[X] + E[Y].$$

$$\Rightarrow E[X+Y] \le E[X] + E[Y].$$

$$\therefore E[X+Y] = E[X] + E[Y].$$

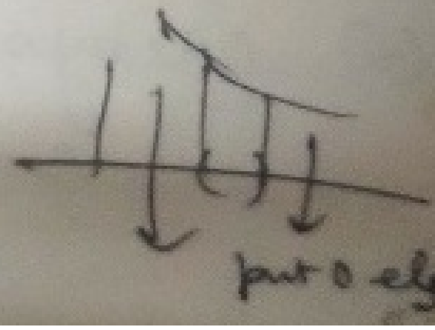
(iii) $X \ge 0 \Rightarrow E[X] \ge 0.$ apply this on $\underline{X-Y}$

$X \ge Y \Rightarrow E[X] \ge E[Y].$

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

use this

$$\int_A X \, dP = E[X I_A].$$



put 0 elsewhere

$$\int_{A \cup B} X \, dP = \int_A X \, dP + \int_B X \, dP, \text{ if } A, B \text{ disjoint.}$$

(iv) Let μ be the law of X.

$$\mu(A) = P(X \in A)$$

$$\int f(X) \, dP = \int f(x) \, d\mu(x), \text{ whenever defined.}$$

Pf $\cancel{X = I_A}$ $f = I_A$.

$$\text{L.H.S.} = \int I\langle X \in A \rangle \, dP$$

$$= P(X \in A) = \mu(A) = \int_A I(x) \, d\mu(x)$$

Extend by linearity to simple f, by approximation argument to general f.

(v) Chebyshev inequality:

$$X \geq 0, \ \varepsilon > 0, \ P(X \geq \varepsilon) \leq \frac{E[X]}{\varepsilon}.$$

$P(|X-E(X)|\geq\epsilon) \leq \dfrac{E[(X-E(X))^2]}{\epsilon^2}$ (Markov)

$X \geq 0$ $\quad E[X] = E[X[I\langle X\geq\epsilon\rangle + I\langle X<\epsilon\rangle]]$

$= E[XI\langle X\geq\epsilon\rangle] + E[XI\langle X<\epsilon\rangle]$

$\downarrow$ X being positive this quantity ≥ 0

$\geq E[XI\langle X\geq\epsilon\rangle]$

$\geq \epsilon E[I\langle X\geq\epsilon\rangle]$.

$\because XI\langle X\geq\epsilon\rangle \geq \epsilon I\langle X\geq\epsilon\rangle$

$\downarrow$

[when $X\geq\epsilon$ then both are 1, or both are 0]

$= \epsilon P(X\geq\epsilon)$

$\Rightarrow P(X\geq\epsilon) \leq \dfrac{E[X]}{\epsilon}$

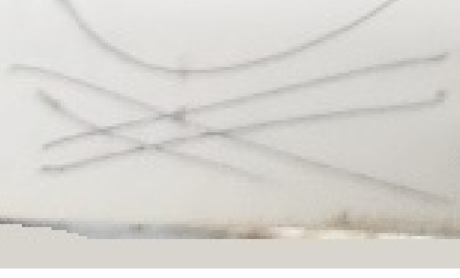
(vi) Jensen's inequality: f convex $\Rightarrow E[f(X)] \geq f(E[X])$

f concave $\Rightarrow E[f(X)] \leq f(E[X])$

$X = [x_1, \ldots, x_d]$

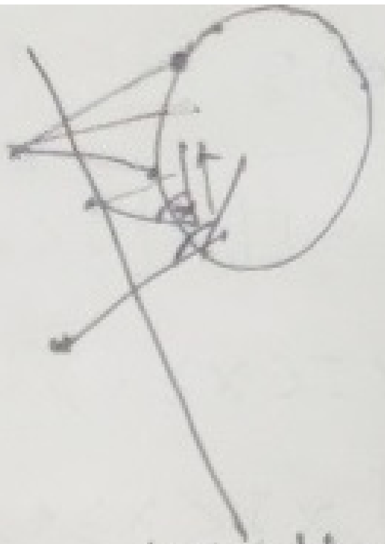
f convex $\Rightarrow f(X) = \sup_g \langle g(X) = \sum_{i=1}^{d} a_i x_i + b,$

$g(x) \leq f(x)\ \forall x \rangle$



$f(X)$ will be supremum of all affine functions

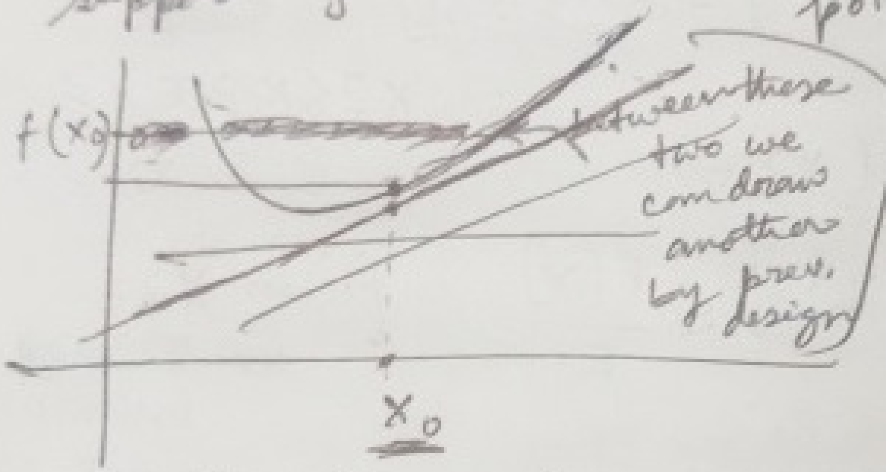
Fact: [illegible]



There must be ~~[illegible]~~ an unique closest point. If we join two pts will be $\geq 90°$

hyperplane

suppose false at x_0



pointwise supremum of families of linear functions

$$E[f(X)] = E\left[\sup_g g(X)\right] \geq E[g(X)]$$

$$= g\left[E[X]\right] \quad \because g \text{ is affine}$$

i.e. linear ~~[illegible]~~ in

$$E[f(X)] \geq g(E[X]) \quad \forall g \text{ affine} \leq f.$$

$$\therefore E[f(X)] \geq \sup_{\substack{g \leq f \\ \text{affine}}} E[g(X)]$$

$$= f(E[X])$$

$\delta_x(dw)$ $\quad$ $\delta_x(A) = 1$ if $x \in A$

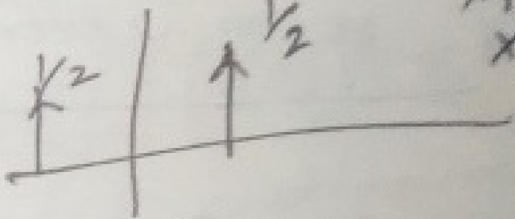
$= 0$ ow.

Dirac

$P = \sum_{i=1}^{n} a_i \delta_{x_i}$ $\quad$ $a_i \geq 0$, $\sum a_i = 1$.

any convex combination of probability measure is also a probability

$\int f \, dP$



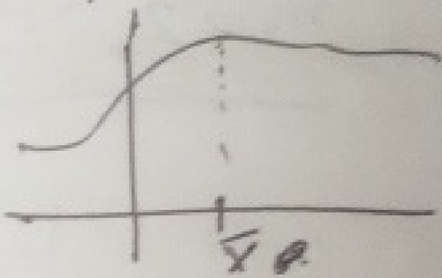
$x_1 = 1$

$x_2 = -1$

$a_1 = a_2 =$

$= \sum a_i f(x_i)$

$\geq f(\sum a_i x_i)$



? $= f\left(\int x \, dP\right)$ [for discrete valued random variables it's a re-statement of convexity]

$\int f \, d\delta_{\bar{x}} = f(\bar{x})$

it just samples the function at that point.

$\int f(x) \, dP = \int f(x) \, d\mu(x)$.

Def. $E[X]$ mean, expectation.

$E[X^m]$, $m \geq 1$, m-th moment.

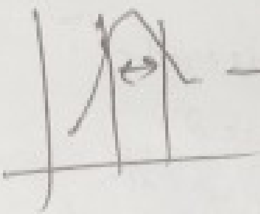
$E[(X-E[X])^m]$, $m \geq 1$, $\rightarrow$ m-th centred moment

$E[|X|^m]$, $m \geq 1$, with absolute moment.

Moment & absolute ~~value~~ Moment will be same if m is even

$m=2$, $\therefore E[(X-E(X))^2] \rightarrow$ variance

$\sqrt{\text{variance}}$ = standard deviation



$E[X,Y] = \langle X, Y \rangle$

Standard deviation becomes its norm

Convergence concepts:

(0) $X_n \longrightarrow X$ pointwise if $X_n(w) \rightarrow X(w)$ $\forall w$

(1) $X_n(w) \longrightarrow X(w)$ almost sure or with probability one if $P(X_n(w) \rightarrow X(w)) = 1$

Probability of two r.v.s are equal if ~~[illegible]~~ they are equal with almost sureity.

(2) $X_n(w) \to X(w)$ in probability if $\varepsilon > 0$,
$$P(|X_n - X| \geq \varepsilon) \to 0.$$

(3) $X_n \to X$ in m-th mean if
$$E[|X_n - X|^m] \to 0.$$

(4) $X_n \to X$ in law if $E[f(X_n)] \to E[f(X)]$ for bounded continuous functions f.

$$\begin{array}{l} (0) \to (1) \searrow \\ \qquad\qquad\quad (2) \to (4) \\ \quad\;\; (3) \nearrow \end{array}$$

(a) (2) depends only on pair laws (X_n, X) only on the joint distr

(b) (4) depends only on laws.

$\boxed{\text{Convergence of laws (distr's, not convergence of random variables}}$

Suppose $X_n \to X$ a.s. let $\varepsilon, \delta > 0$.

claim $\exists$ $N \geq K$ s.t.

$P(|X_n - X| \geq \varepsilon) > \delta$

[In measure theory: almost everywhere

[almost certainly is also the rarely

$A_K = \langle w : |X_n(w) - X(w)| \geq \varepsilon$ for some $n \geq K \rangle$, $K \geq 1$.

$\therefore P\left(\bigcap_{K=1}^{\infty} A_K\right)$

~~...~~ no matter how big n is chosen $|X_n(w) - X(w)| \geq \varepsilon$ occurs infinitely often.

$\leq P(X_n \not\to X) = 0$

$A_{K+1} \subset A_K$

(if it happens after $K+1$, it will happen after K also)

$$\lim_{N \uparrow \infty} P\left(\bigcap_{K=1}^{N} A_K\right) = P\left(\bigcap_{K=1}^{\infty} A_K\right) = 0,$$

$\therefore$ Can find N s.t. $P\left(\bigcap_{K=1}^{N} A_K\right) < \delta$.

$$\begin{array}{c} P\left(\bigcap_{K=1}^{N} A_K\right) \\ \shortparallel \\ P(A_N) \end{array}$$

$P(|X_n - X| \geq \epsilon$ for some $n \geq N) < \delta$. —— (1)

$\bigcup$ $\forall\, n \geq N$ ↑ This set is contained on that

$P(|X_n - X| \geq \epsilon) < \delta$

$\lim_{n \to \infty} P(|X_n - X| \geq \epsilon) \to 0.$

That is, (1) → (2).

$m \geq 1$, $E[|X_n - X|^m] \to 0.$

$$P(|X_n - X| \geq \epsilon)$$
$$= P(|X_n - X|^m \geq \epsilon^m) \leq \frac{E(|X_n - X_1|^m]}{\epsilon^m} \to 0$$

$\langle w : |X_n - X| \geq \epsilon$ for some $n \geq N \rangle$

$$= \bigcup_{n \geq N} \langle w : |X_n(w) - X(w)| \geq \epsilon \}$$