Probability Theory

- Preliminaries - σ-algebras, Dynkins π-λ Theorem, independence, Borel-Cantelli lemmas, Kolmogorov's 0-1 Law, Kolmogorov's maximal inequality, strong and weak laws of large numbers

- Central Limit Theorems - weak convergence, characteristic functions, tightness, iid central limit theorem, Lindeberg-Feller central limit theorem

- Conditioning - conditional probability and expectation, regular conditional probabilities

- Martingales - stopping times, upcrossing inequality, uniform integrability, A.S. convergence, Doob's decomposition, Doob's inequality, L^p convergence, L^1 convergence, reverse martingale convergence, optional stopping time, Wald's Identity

- Markov Chains - countable state space, stationary measures, convergence theorems, recurrence and transience, asymptotic behavior

Theorems & Proof Ideas

Chebyshev's Inequality

$i_A P(X \in A) \leq E\varphi(X)$

$a^2 P(|X| \geq a) \leq E|X|^2$

Proof Idea:

$E\left(i_A 1_{X\in A} \leq \varphi(X) 1_{X \in A} \leq \varphi(X) \right)$

$i_A = \inf\{\varphi(x) : x \in A\}$.

iid Weak LLN

$X_1, X_2, \ldots$ iid

$EX_i = \mu$ $\quad var(X_i) < \infty$

$\frac{X_1 + \cdots + X_n}{n} \to \mu$ in P

Proof Idea:

Chebyshev $P(|\frac{S_n}{n} - \mu| > \varepsilon) \leq \varepsilon^{-2} var(\frac{S_n}{n}) = \frac{\sigma^2}{n\varepsilon^2} \to 0$

Alt: characteristic functions $\to e^{i\mu t} = \varphi_\mu(t)$.

and $\Rightarrow \mu$ implies in P μ.

(If $E|X_i| < \infty$ instead of variance use truncation and triangular arrays)

Borel-Cantelli Lemma

$\sum_{n=0}^{\infty} P(A_n) < \infty \Rightarrow P(A_n \text{ i.o.}) = 0$

Proof Idea:

$N = \sum_k 1_{A_k}$ $\quad EN = \sum P(A_n) < \infty$

so $N < \infty$ a.s. $\rightsquigarrow P(A_n \text{ i.o.}) = P(N = \infty) = 0$.

Second Borel-Cantelli Lemma

$\sum_{n=0}^{\infty} P(A_n) = \infty$ & A_n independent

$\Rightarrow P(A_n \text{ i.o.}) = 1$

Proof Idea:

$P(\bigcup_M^\infty A_n) \longrightarrow P(A_n \text{ i.o.})$ as $M \to \infty$.

$1 - P(\bigcup_M^N A_n) = P(\bigcap_M^N A_n^c) = \prod_M^N (1 - P(A_n)) \leq \prod_M^N e^{-P(A_n)} = e^{-\sum P(A_n)} \to 0$.

so $P(\bigcup_M^N A_n) \to 1$ $\forall M$ as $N \to \infty$. So $P(\bigcup_M^\infty A_n) = 1 \to P(A_n \text{ i.o.})$

iid Strong LLN

$X_1, X_2, \ldots$ iid

$EX_i = \mu$ $\quad EX_i^4 < \infty$

$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{a.s.} \mu$

Proof Idea:

STP $P(|\frac{S_n}{n}| > \varepsilon \text{ i.o.}) = 0$ $\forall \varepsilon > 0$ (assume $EX_i = \mu = 0$).

By Cheb. $P(|\frac{S_n}{n}| > \varepsilon) \leq \underline{E|S_n|^4}/n^4\varepsilon^4 < \frac{C}{\varepsilon^4 n^2}$

$ES^4 = nEX_i^4 + \alpha n^2 EX_i^2 EX_j^2 \leq \beta n^2 EX_i^4 \leq Cn^2$

So $\sum P(\quad) < \infty$ and BC says $P(|\frac{S_n}{n}| > \varepsilon \text{ i.o.}) = 0$.

If $EX = \infty$ $X \geq 0$ take $Y_{n,B} = 1_{X_n \leq B} X_n$. apply.

Kolmogorov 0-1 Law

$X_1, X_2, \ldots$ independent

$A \in \mathcal{T}(X_1, X_2, \ldots)$

$\Rightarrow P(A) \in \{0, 1\}$

Proof Idea:

show A independent from itself so

$P(A) = P(\cap A) = P(A)^n \in \{0, 1\}$ as $n \to \infty$.

Theorems & Proof Ideas

Kolmogorov Max Inequality

$X_1, X_2, \dots$ independent

$EX_i = 0 \quad var(X_i) < \infty$

$S_n = X_1 + \dots + X_n$

$P(\max_{1 \le m \le n} |S_m| \ge x) \le x^{-2} var(S_n)$

Proof Idea:

Break up by $A_k = \{|S_k| \ge x \text{ first time}\}$.

$var(S_n) = ES_n^2 \ge \sum_k \int_{A_k} S_n^2 dP \ge \sum \int_{A_k} S_k^2 dP$

$\ge \sum x^2 P(A_k) \left(= x^2 P(\max_{1 \le m \le n} |S_m| \ge x)\right)$

clever quadratic rewriting trick.

Inversion Formula for φ

$\int |\varphi(t)| dt < \infty \quad \mu$ has bdd density

$f(y) = \frac{1}{2\pi} \int e^{-ity} \varphi(t) dt$

Proof Idea:

use general inversion formula and $\frac{e^{-itx} - e^{-it(x+h)}}{it} = \int_x^{x+h} e^{-ity} dy$ and apply Fubini's.

Continuity Theorem

$\varphi_n(t) \to \varphi_\infty(t)$ pt wise &

φ_∞ cts @ $t=0 \implies \mu_n$ tight

and $\mu_n \Rightarrow \mu_\infty$ (w/ char fn φ_∞)

Proof Idea:

decay of measure near ∞ bounded by integral of φ near 0. Continuity sends this to 0 so no mass loss $\to$ tightness.

iid Central Limit Theorem

$X_1, X_2, \dots$ iid $EX_i = \mu$

$Var(X_i) = \sigma^2 \in (0, \infty)$

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow \mathcal{N}(0,1)$$

Proof Idea:

$\varphi_X(t) = 1 + itEX - \frac{t^2 EX^2}{2} + o(t^2) \quad EX = 0$

$= 1 + 0 - \frac{t^2\sigma^2}{2} + o(t^2)$

$\varphi_{\frac{S_n}{\sigma\sqrt{n}}}(t) = \left(1 - \frac{t^2}{2n} + o(\frac{t^2}{n})\right)^n \to e^{-t^2/2} = \varphi_{\mathcal{N}}(t)$

Continuity thm says $\frac{S_n}{\sigma\sqrt{n}} \Rightarrow N(0,1)$

Lindeberg-Feller CLT

$X_{n,m}$ independent $1 \le m \le n$ $EX_{n,m} = 0$

(i) $\sum_{m=1}^n E(X_{n,m}^2) \to \sigma^2 > 0$

(ii) $\forall \varepsilon > 0, \sum_{m=1}^n E(|X_{n,m}|^2; |X_{n,m}| > \varepsilon) \to 0$

$X_1 + \dots + X_n \Rightarrow \sigma \mathcal{N}(0,1)$

Proof Idea:

$\varphi_{S_n}(t) = \prod^n \varphi_{n,m}(t) \to \prod^n \left(1 - \frac{t^2 \sigma_{n,m}^2}{2}\right) \to \exp(-t^2\sigma^2/2)$

L-F $\Rightarrow$ IID: $X_{n,m} = \frac{X_m - \mu}{\sqrt{n}} \quad X_{n,1} + \dots + X_{n,n} = \frac{S_n - n\mu}{\sqrt{n}}$

Theorems & Proof Ideas

Upcrossing Inequality

X_m submartingale, $a < b$

U_n = # of upcrossings by time n

$(b-a)EU_n \leq E(X_n - a)^+ - E(X_0 - a)^+$

Proof Idea:

$Y_m = a + (X_m - a)^+$ H upcrossing betting

$(b-a)U_n \leq (H \cdot Y)_n$

$(1 - H \cdot Y)_n$ submartingale

$(b-a)EU_n \leq (H \cdot Y)_n \leq (H \cdot Y)_n + (1 - H \cdot Y)_n = EY_n - EY_0.$

A.S. Martingale Convergence

X_n submartingale, $\sup EX_n^+ < \infty$

$\Rightarrow X_n \xrightarrow{a.s.} X$ and $E|X| < \infty$

Proof Idea:

upcrossing $EU_n \leq (b-a)^{-1} E(X_n - a)^+ \leq \frac{|a| + EX_n^+}{b-a}$

Bdd $\sup EX_n^+$ shows $EU_n \uparrow EU < \infty$. ($U < \infty$ a.s.)

Holds for all a, b so always ends up inside a narrow range $\Rightarrow$ X_n conv a.s.

Doob's Decomposition

X_n submartingale, has unique decomposition $X_n = M_n + A_n$

M_n martingale, A_n inc. pred. seq.

Proof Idea:

$A_n - A_{n-1} = E(X_n | \mathcal{F}_{n-1}) - X_{n-1}$ $(A_0 = 0)$

set $M_n = X_n - A_n$ and check conditions.

Bounded T Optional Stopping

X_n submartingale, T stopping time

$P(T \leq k) = 1$ for some k

$\Rightarrow EX_0 \leq EX_T \leq EX_k$

Proof Idea:

$X_{T \wedge n}$ submartingale $EX_0 \leq EX_{T \wedge n} \leq EX_{T \wedge k} = EX_T$

$K_n = 1_{N < n} \rightarrow (K \cdot X)_n = X_n - X_{N \wedge n}$ submart.

$EX_K - EX_N = E(K \cdot X)_K \geq E(K \cdot X)_0 = 0.$

Doob's Inequality:

X_n submartingale, $\lambda > 0$

$\lambda P(\max_{0 \leq m \leq n} X_m^+ \geq \lambda) \leq EX_n^+$

Proof Idea:

$N = \inf\{m : X_m \geq \lambda \text{ or } m = n\}$

$\lambda P(\underbrace{\max X_m^+ \geq \lambda}_{A}) \leq \underbrace{EX_N 1_A \leq EX_n 1_A}_{P(N \leq n) = 1} \leq EX_n \leq EX_n^+$

Theorems & Proof Ideas

L^p Maximal Inequality

X_n submartingale, $1<p<\infty$

$$E\left(\max_{0\le m\le n} (X_m^+)^p\right) \le \left(\frac{p}{p-1}\right)^p E(X_n^+)^p$$

Proof Idea:

Express $E(\max{}^p)$ as integral, apply Doob's inequality and some clever calculus.

L^p Convergence Theorem:

X_n submartingale $1<p<\infty$

$\sup E|X_n|^p<\infty$ then

$X_n \to X$ a.s. and in L^p

Proof Idea:

$(EX_n^+)^p \le (E|X_n|)^p \le E|X_n|^p$ so get $\sup EX_n^+<\infty$ and $X_n\to X$ a.s. convergence.

$E(\max|X_m|^p)<\infty$ by $\sup E|X_n|^p<\infty$,

$|X_n-X|^p \le (2\sup|X_n|)^p$ + dominated conv

$E|X_n-X|^p\to 0$ so $X_n\to X$ in L^p.

L^1 convergence Theorem:

X_n submartingale TFAE

(i) X_n uniformly integrable

(ii) X_n converges in L^1 and a.s.

Martingale $\Rightarrow X_n=E(X|\mathcal{F}_n)\ \forall n$.

Proof Idea:

U.I. $\Rightarrow \sup E|X_n|\le M+1<\infty$ so $\sup EX_n^+<\infty$. convergence (& in P).

gives $X_n\to X$ a.s. w/ $E|X|<\infty$.

$$\varphi_M(x)=\begin{cases}\pm M & |x|\ge M\\ x & |x|<M\end{cases}$$ cut off after band of M

$$E|X_n-X|\le \underbrace{E|X_n-\varphi_M(X_n)|}_{\to 0 \text{ by U.I.}}+\underbrace{E|\varphi_M(X_n)-\varphi_M(X)|}_{\to 0 \text{ since in P}}+\underbrace{E|X-\varphi_M(X)|}_{\to 0 \text{ by UI and } E|X|<\infty}$$

$\to 0$.

Reverse Martingale Convergence

X_n reverse martingale

$X_n\to X_{\infty}$ in L^1 and a.s.

Proof Idea:

Same upcrossing inequality gives $X_n\to X$ a.s. Martingale gives $X_n=E(X_0|\mathcal{F}_n)$ (by reverse) is U.I collection so converges in L^1 too.

U.I. $\Rightarrow$ Optional Stopping

X_n U.I. submartingale

$\Longrightarrow EX_0\le EX_N\le EX_\infty$

N stopping time

Proof Idea:

$\underbrace{X_{N\wedge n}}_{\text{submartingale}}$ U.I.: $E(|X_{N\wedge n}| : |X_{N\wedge n}|>M) = E(|X_N|; |X_N|>M, N\le n) + \underbrace{E(|X_n|; |X_n|>M, N>n)}_{\to 0 \text{ by U.I.}}$

$EX_{N\wedge n}\le EX_n$

$\sup EX_{N\wedge n}^+\le \sup EX_n^+<\infty$ by U.I. so $X_{N\wedge n}\to X_N$ a.s. and $E|X_N|<\infty$ so $P(|X_N|>M)\to 0$. UI $\Rightarrow X_{N\wedge n}\to X_N$ $\underline{\underline{L^1}}$.

$$EX_0-X_N=\underbrace{EX_0-X_{N\wedge n}}_{\le 0 \text{ bdd stop time}}+\underbrace{X_{N\wedge n}-X_N}_{\to 0\ L^1 \text{ conv.}}$$

"Increments" Optional Stopping

X_n submartingale

$E(|X_{n+1}-X_n||\mathcal{F}_n)\le B$ a.s.

$EN<\infty$, N stop time

$\Bigg\}$ $X_{N\wedge n}$ U.I. $\Rightarrow EX_0\le EX_N$

Proof Idea:

Show $X_{N\wedge n}$ dominated by int. r.v so is UI.

$$|X_{N\wedge n}|\le |X_0|+\sum|X_{m+1}-X_m|1_{N>m}$$

$E(\Sigma -)\le BE(N)<\infty$.

so $EX_0\le EX_N$.

Theorems & Proof Ideas:

$\boxed{0.5}$

Wald's Equation:

$\xi_1, \xi_2, \ldots$ iid $E\xi_i = \mu$

N st time $EN < \infty$

$\Rightarrow ES_N = \mu EN$

Proof Idea:

$X_n = S_n - \mu n$ martingale $\rightsquigarrow ES_{N \wedge n} = \mu E(N \wedge n)$.

$0 \le N \wedge n \uparrow N$ so monotone convergence of RHS.

$S_{N \wedge n} \to S_N$ so $ES_{N \wedge n} \to ES_N$ also.

Alt: "increments" optional stopping

Existence of Stat. Meas.

$\exists$ recurrent x

$\mu_x(y) = E_x\left(\sum_{n=0}^{T_x - 1} 1_{X_n = y}\right)$

stationary measure

Proof Idea:

$T_x = \inf\{n \ge 1 : X_n = x\}$

$E_x\left(\sum_{n=0}^{T-1} 1_{X_n = y}\right) =$ Expected # of visits to y in $\{0, \ldots, T-1\}$ $= \sum_{n=0}^{\infty} P_x(X_n = y, T > n)$

$\sum_z \underbrace{\mu_x(z)}_{\text{expand}} p(z, y) = \sum_z \sum_{n=0}^{\infty} (\quad) \underset{\text{Fubini}}{=} \sum_{n=1}^{\infty} P_x(X_n = y, T \ge n)$

$= E_x\left(\sum_{n=1}^{T} 1_{X_n = y}\right) =$ expected visits to y in $\{1, \ldots, T\}$

$X_0 = X_T = x \ne y$ $\left(\begin{array}{l}\text{or } x = y \text{ \&} \\ \mu_x(x) = 1\end{array}\right)$

Uniqueness of Stat. Meas.

irreducible & $\exists$ recurrent x

$\Rightarrow$ stat meas unique (up to scaling)

Proof Idea:

ν stat., a recurrent $\nu(z) = \nu(a)p(a,z) + \sum_{y \ne a} \nu(y)p(y,z)$

$\rightsquigarrow \nu(z) = \nu(a)\mu_a(z) + P(-) \ge \nu(a)\mu_a(z)$

$\nu(a) = \sum_x \nu(x)p(x,a) \ge \sum_x \nu(a)\mu_a(x)p(x,a) = \nu(a)\mu_a(a) = \nu(a)$

gives termwise equality since all terms ≥ 0 so

$\nu(z) = \underbrace{\nu(a)}_{\text{scaling factor}}\mu_a(z)$.

Existence of Stat. Dist.

irreducible, TFAE

(i) $\exists$ positive recurrent state x

(ii) $\exists$ stat distribution $\pi(x)$

Proof Idea:

(i) $\Rightarrow$ (ii)

μ_x stat. meas.

$\sum_y \mu_x(y) = \sum_y \sum_n P_x(X_n = y, T_x > n) = \sum_n P_x(T_x > n) = E_x T_x < \infty$

stat meas. unique up to scaling so divide by $E_x T_x$ gives stat dist.

(ii) $\Rightarrow$ (i)

Every state recurrent so $\mu_y(z)$ stat meas, but unique up to scaling

$\frac{\mu_y(z)}{E_y T_y} = \pi(z) \rightsquigarrow \pi(y) = \frac{1}{E_y T_y}$

$\pi(y) \ne 0 \ \forall y$ so $E_y T_y < \infty \ \forall y$.

Markov Convergence

p irreducible, aperiodic, and stat. dist π then

$p^n(x,y) \longrightarrow \pi(y)$

Proof Idea:

$X_n \times Y_n$ copies of chain, Y starts @ π dist.

p irred + aperiodic $\Rightarrow \bar{p}(\bar{X}, \bar{Y})$ irreducible too.

π stationary $\to \bar{\pi}$ stationary $\to \bar{p}$ all recurrent states.

(x,x) recurrent, $T_{(x,x)} < \infty$ a.s. $\to T_{x=y} < T_{(x,x)} < \infty$ a.s.

Bound $\sum_n |p^n(x,y) - \pi(y)| \le 2P(T > n) \to 0$.

Partition of Recurrent States:

$R = \{\text{recurrent states}\}$

$R = \bigcup_i R_i \leftarrow$ closed & irreducible

Proof Idea:

$C_x = \{y : \rho_{xy} > 0\}$ show this satisfies equivalence relation on R.

Each C_x is closed & irreducible by construction.

σ- algebras

Defns

- σ-algebra
  (i) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
  (ii) $\{A_i\} \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$
  (countable unions)
- algebra
  (i) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
  (ii) $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$
  (finite unions)
- semi-algebra
  (i) $A \in \mathcal{F} \Rightarrow A^c = \bigsqcup_{i=1}^{n} B_i, B_i \in \mathcal{F}$
  (ii) $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$
  (finite intersection)

$$\boxed{\sigma\text{-algebras} \subsetneq \text{algebras} \subsetneq \text{semi-algebras}}$$

Example

$\Omega = \mathbb{Z}$

$\mathcal{F}$ = finite or cofinite subsets of $\mathbb{Z}$

$\mathcal{F}$ is an algebra but not a σ-algebra

$2\mathbb{Z} = \bigcup_n \{2n, -2n\} \notin \mathcal{F}$

Example

$\Omega = \mathbb{R}$

$\mathcal{F} = \{\phi\} \cup \{(a,b] : -\infty \leq a < b \leq \infty\}$

$\mathcal{F}$ is a semi-algebra but $(a,b]^c = (-\infty, a] \sqcup (b, \infty] \notin \mathcal{F}$.

Random Variables

Defn

- $X:(\Omega,\mathcal{F},P)\to\mathbb{R}$ is $\mathcal{F}$-measurable if $\forall B\in\mathcal{R}$, $X^{-1}(B)\in\mathcal{F}$. Then X is a random variable.
- $X:(\Omega,\mathcal{F},P)\to\mathbb{R}^d$ is $\mathcal{F}$-measurable then it is a random vector. $(X=(X_1,\dots,X_d))$ random variables
- $\sigma(X)=\{X^{-1}(B): B\in\mathcal{R}\}$ is the σ-field generated by X and is smallest σ-field in which X is meas.

Combinations

- compositions of measurable maps are measurable
- $X_1,\dots,X_n$ rand. var then $X_1+\dots+X_n$ is too (finite sums)

  Pf: $\{X_1+X_2<r\}=\bigcup_{q\in\mathbb{Q}}\{X_1<a-q\}\cap\{X_2<q\}\in\mathcal{F}$ and induction.
- $\inf X_n$, $\sup X_n$, $\liminf X_n$, $\limsup X_n$ random variables (possibly on extended real line $\mathbb{R}^*$)

Random Variables

$\boxed{X}$

$X(\omega)=\sup\{y: F(y)<\omega\}$, $\omega\in(0,1)$

$X(\omega)=\sup\{y:\mu((-\infty,y])<\omega\}$, $\omega\in(0,1)$

Distribution Functions

$F(y):=P(X\le y)$

$\boxed{F}$

$F(y):=\mu((-\infty,y])$

Probability Measures

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

extend $\mu((-\infty,y]):=F(y)$

$\boxed{\mu}$

Measures and Distribution Functions

Defns

- μ is a <u>measure</u> on $(\Omega, \mathcal{F})$ ($\leftarrow$ σ-algebra) if $\mu : \mathcal{F} \to \mathbb{R}$
  (i) $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathcal{F}$
  (ii) $A_i \in \mathcal{F}$ countable and disjoint
  $$\mu\left(\bigsqcup_i A_i\right) = \sum_i \mu(A_i)$$
- μ is a <u>probability measure</u> if $\mu(\Omega) = 1$.
- X a random variable $(\Omega, \mathcal{F}, P) \to \mathbb{R}$ defines
  its <u>distribution function</u> $F(x) = P(X \leq x) = P(X^{-1}((-\infty, x]))$
  its <u>probability measure</u> (on $\mathbb{R}$) $\mu(A) = P(X \in A) = P(X^{-1}(A))$.
- If $F(x) = \int_{-\infty}^{x} f(y)\,dy$ then $f(y)$ is the <u>density function</u> of X and makes X <u>absolutely continuous</u>
- μ is <u>σ-finite</u> if $\exists \{A_n\}$ w/ $\mu(A_n) < \infty$ and $\cup_n A_n = \Omega$.

Properties

- Distribution functions
  (i) F is nondecreasing
  (ii) F is right continuous
  (iii) $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$.

  } characterize dist fun so if F satisfies then $X(w) = \sup\{y : F(y) < w\}$ $X : (0,1) \to \mathbb{R}$ is R.V. w/ dist fun F.

- measures
  $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$
  $A \subseteq \cup_i A_i \Rightarrow \mu(A) \leq \sum_i \mu(A_i)$
  $A_i \uparrow A \Longrightarrow \mu(A_i) \uparrow \mu(A)$
  $A_i \downarrow A$, $\mu(A_1) < \infty \Longrightarrow \mu(A_i) \downarrow \mu(A)$

Integration

$\longrightarrow \mu$ is σ-finite $\longleftarrow$

Simple Functions

$\varphi = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$ for disjoint A_i, with $\mu(A_i) < \infty$

$\int \varphi \, d\mu := \sum a_i \mu(A_i)$ [invariant of φ representation]

Bounded Functions

Take simple $\varphi \leq f$ and $\psi \geq f$ then

$\int f \, d\mu := \sup_{\varphi \leq f} \int \varphi \, d\mu = \inf_{\psi \geq f} \int \psi \, d\mu$

Non-negative Functions

$\int f \, d\mu := \sup_h \left\{ \int h \, d\mu : 0 \leq h \leq f,\ h \text{ bounded},\ \mu(\{x : h(x) \neq 0\}) < \infty \right\}$

Integrable Functions

f is integrable if $\int |f| \, d\mu < \infty$

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

$\int f \, d\mu := \int f^+ \, d\mu - \int f^- \, d\mu$

Properties

- $f \geq 0$ a.e. $\Rightarrow \int f \, d\mu \geq 0$
- $\forall a \in \mathbb{R}$ $\int af \, d\mu = a \int f \, d\mu$
- $\int f+g \, d\mu = \int f \, d\mu + \int g \, d\mu$
- $g \leq f$ a.e. $(g = f$ a.e.$) \Rightarrow \int g \, d\mu \leq \int f \, d\mu$ $(\int g \, d\mu = \int f \, d\mu)$
- $\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$.

Convergence Theorems

Monotone Convergence

$f_n \geq 0 \quad f_n \uparrow f$

$$\Rightarrow \int f_n \, d\mu \uparrow \int f \, d\mu$$

Dominated Convergence

$$\left.\begin{array}{l} f_n \to f \text{ a.e.} \\ |f_n| \leq g \quad \forall n \\ g \text{ integrable} \\ \left(\int |g| \, d\mu < \infty\right) \end{array}\right\} \int f_n \, d\mu \longrightarrow \int f \, d\mu$$

Bounded Convergence

$$\left.\begin{array}{l} f_n \to f \text{ a.e.} \\ |f_n| \leq M \end{array}\right\} \int f_n \, d\mu \longrightarrow \int f \, d\mu$$

Fatou's Lemma

$$\left. f_n \geq 0 \quad \right\} \liminf_{n \to \infty} \int f_n \, d\mu \geq \int \liminf_{n \to \infty} f_n \, d\mu$$

Expected Values

~~Defn~~ Expected Value

$$EX = \int_\Omega X\,dP \quad \left(\begin{array}{l}\text{Expected Value} \\ \text{or mean of } X\end{array}\right)$$

X has measure μ

$$E\varphi(X) = \int \varphi(x)\,\mu(dx)$$

Note:
$EX = EX^+ - EX^-$
so if $EX < \infty$ then $EX^{\pm} < \infty$
and $E|X| < \infty$ too.

X has density $f(x)$

$$E\varphi(X) = \int_{-\infty}^{\infty} \varphi(x) f(x)\,dx$$

$$E|X| = \int P(|X| > x)\,dx \quad \text{or} \quad \sum_n P(|X| \geq n)$$

$$E(X+Y) = E(X) + E(Y)$$

$$E(aX+b) = aE(X) + b$$

<u>k^{th} moment</u>

$$EX^k$$

if k even or $X \geq 0$ $\Rightarrow EX^k = \int_0^\infty k y^{k-1} P(|X| > y)\,dy$

<u>variance</u>

$$E(X-\mu)^2 = EX^2 - (EX)^2 = EX^2 - \mu^2$$

Chebyshev's Inequality

General form

$\varphi \geq 0 \qquad i_A = \inf\{\varphi(y) : y \in A\}$

$$i_A P(X \in A) \leq \underbrace{E(\varphi(X); X \in A)}_{\int_A \varphi(x)\,dP} \leq E\varphi(X)$$

Common Form

$$P(|X| \geq a) \leq a^{-2} \operatorname{var}(X)$$

$$\| \quad P(|X|^2 = X^2 \geq a^2)$$

$\varphi(x) = x^2 \quad A = \{x : |x| \geq a\} \quad i_A = a^2 \quad$ when $EX = 0$ $\; E\varphi(x) = \operatorname{var}(x)$

Jensen's Inequality

φ convex (e.g. x^2, $|x|$)

$$\varphi\left(\int f\,d\mu\right) \leq \int \varphi(f)\,d\mu$$

$$\varphi(EX) \leq E\varphi(X)$$

Example: $(EX^2)^2 \leq EX^4$

Holder's Inequality

$p, q \in [1, \infty]$ $\quad \frac{1}{p} + \frac{1}{q} = 1$

$$E|XY| \le (E|X|^p)^{1/p} (E|Y|^q)^{1/q}$$

$$\int |fg| d\mu \le \left(\int |f|^p d\mu\right)^{1/p} \left(\int |g|^q d\mu\right)^{1/q}$$

Cauchy-Schwarz Inequality

$p = q = 1/2$ Holder's.

$$E|XY| \le \sqrt{EX^2} \sqrt{EY^2}$$

Note: since $XY \le |XY|$, we have $(EXY)^2 \le EX^2 EY^2$.

Fubini's Theorem

Fubini's Theorem

μ_1, μ_2 σ-finite $\quad f \geq 0$ OR $\int |f| \, d\mu < \infty$

then $$\int_X \int_Y f \, d\mu_1 \, d\mu_2 = \int_{X \times Y} f \, d\mu_1 \times \mu_2 = \int_Y \int_X f \, d\mu_2 \, d\mu_1$$

Counterexample $(f \not\geq 0)$

- $f : \mathbb{N} \times \mathbb{N} \to \{0, \pm 1\}$

$$\begin{array}{c|ccccc}
 & \vdots & \vdots & \vdots & \vdots & \iddots \\
\uparrow & 0 & 0 & 1 & -1 & \cdots \\
m & 0 & 1 & -1 & 0 & \cdots \\
 & 1 & -1 & 0 & 0 & \cdots \\
\hline
 & & n \longrightarrow & & &
\end{array}$$

$$\sum_{n=0}^{\infty} f(n,m) = \int f \, d\mu^{\text{counting measure}}$$

$$\sum_m \sum_n f(n,m) = \sum_m 0 = 0 \quad \times$$

$$\sum_n \sum_m f(n,m) = \left(\sum_{n>0} 0\right) + 1 = 1$$

Counterexample (μ not σ-finite)

- $X = Y = (0,1)$ $\quad \mu_1$ Lebesgue, μ_2 counting meas.

$$f(x,y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$$

$$\int_X \int_Y f \, d\mu_2 \, d\mu_1 = \int_X 1 \, d\mu_1 = 1$$

$$\int_Y \int_X f \, d\mu_1 \, d\mu_2 = \int_Y 0 \, d\mu_2 = 0 \quad \#$$

Application

$$E|X| = \int_\Omega |X| \, d\mu = \int_\Omega \int_0^\infty 1_{(x < |X|)} \, dx \, d\mu \overset{\text{Fubini}}{=} \int_0^\infty \int_\Omega 1_{(x < |X|)} \, d\mu \, dx$$

$$= \int_0^\infty P(|X| > x) \, dx$$

Dynkin's π-λ Theorem

Dynkin's π-λ Theorem

$$\left.\begin{array}{l} \mathcal{P} \text{ a } \underline{\pi\text{-system}} \\ (A,B\in\mathcal{P} \Rightarrow A\cap B\in\mathcal{P}) \\ \mathcal{L} \text{ a } \underline{\lambda\text{-system}} \\ (\Omega\in\mathcal{L},\ A,B\in\mathcal{L}\Rightarrow A-B\in\mathcal{L} \\ \quad A_i\in\mathcal{L} \Rightarrow \cup_i A_i\in\mathcal{L}) \end{array}\right\} \quad \mathcal{P}\subseteq\mathcal{L} \Rightarrow \sigma(\mathcal{P})\subseteq\mathcal{L}$$

S a σ-algebra $\Longleftrightarrow$ S π, λ system

Significance

Lifting properties from π-system to its σ-alg

Ex: $\mathcal{L} = \{A: \mu_1(A)=\mu_2(A)\}$

show $\quad \mu_1 = \mu_2$ on $\mathcal{P}\subseteq\mathcal{L}$ then

$\quad \mu_1 = \mu_2$ on $\sigma(\mathcal{P})$.

Types of Convergence

$X_n \xrightarrow{L^r} X$ (in mean, r-mean)

$$\lim_{n\to\infty} E(|X_n - X|^r) = 0$$

$X_n \xrightarrow{P} X$ (in probability, in P)

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

$X_n \xrightarrow{a.s.} X$ (almost sure)

$$Pr(\lim_{n\to\infty} X_n = X) = 1. \quad \left(\begin{array}{l}\forall\varepsilon>0\\ Pr(|X_n - X| > \varepsilon \text{ i.o.}) = 0\end{array}\right)$$

$X_n \xrightarrow{D} X$ (in distribution) $(X_n \Rightarrow X)$

Distribution functions F_n and F

$$\lim_{n\to\infty} F_n(y) = F(y) \text{ for all } y \text{ where } F \text{ cts @ } y.$$

Relationships

$$\begin{tikzcd}
 & \xrightarrow{L^r} \arrow[d, Rightarrow] & \\
\xrightarrow{a.s.} \arrow[r, Rightarrow] & \xrightarrow{P} \arrow[r, Rightarrow] & \xrightarrow{D} \arrow[l, dashed, bend left, "X_n \xrightarrow{D} c"] \arrow[ll, dashed, bend left=60, "{\exists Y_n \overset{d}{=} X_n \text{ s.t. } Y_n \xrightarrow{a.s.} Y}"]
\end{tikzcd}$$

Examples

- $X_n \xrightarrow{P} X$, $X_n \overset{a.e.}{\not\to} X$

  $X_n = 1_{A_n}$, A_n shrinking rotating intervals on $(0,1)$.

- $X_n \xrightarrow{P} X$, $X_n \overset{L^r}{\not\to} X$

  $X_n = n 1_{[0,1/n]} \xrightarrow{P} 0$

  but $E|X_n|^r = n^{r-1} \not\to 0$. $(r \geq 1)$

- $X_n \Rightarrow X$, $X_n \overset{P}{\not\to} X$

  $X_n(w) = \begin{cases} w & n \text{ even} \\ 1-w & n \text{ odd}\end{cases}$ $\quad F_n(y) = y$ on $(0,1)\ \forall n$.

Types of Convergence (cont.)

a.s. $\Rightarrow$ in P

$P(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} P(\lim X_n \neq X) \underset{\text{a.s.}}{=} 0$

$L^p \Rightarrow$ in P

$P(|X_n - X| > \varepsilon) \underset{\text{chebyshev}}{\leq} \varepsilon^{-p} E|X_n - X|^p \xrightarrow[L^p \text{ conv}]{} 0$

in P $\Rightarrow$ weak conv.

$P(X \leq a) \leq P(B \leq a+\varepsilon) + P(|X-B| > \varepsilon)$

$F_n(a) \leq F(a+\varepsilon) + P(|X_n - X| > \varepsilon)$

$F(a-\varepsilon) \leq F_n(a) + P(|X_n - X| > \varepsilon)$

$F(a-\varepsilon) \leq \lim F_n(a) \leq F(a+\varepsilon)$

Let $\varepsilon \to 0$, at cts pts $\lim F_n(a) = F(a)$ ✓

$X_n \Rightarrow C \Rightarrow X_n \to C$ in P

$F_c(y) = \begin{cases} 1 & y \geq c \\ 0 & y < c \end{cases}$ cts @ $\mathbb{R} - \{c\}$.

$P(|X_n - C| > \varepsilon) = F_n(c-\varepsilon) + 1 - F_n(c+\varepsilon)$
$\to F(c-\varepsilon) + 1 - F(c+\varepsilon) = 0.$

$X_n \xrightarrow{\text{in P}} X \iff$ every X_m has subseq $X_{m_k} \xrightarrow{\text{a.s.}} X$.

Choose $\varepsilon_k \to 0$ and m_k s.t.
$P(|X_{m_k} - X| > \varepsilon_k) \leq 2^{-k}$. By BC
$\sum P(|X_{m_k} - X| > \varepsilon_k) < \infty \to P(|X_{m_k} - X| > \varepsilon_k \text{ i.o.}) = 0$
so $X_{m_k} \xrightarrow{\text{a.s.}} X$.

$X_n \xrightarrow{\text{in P}} X$, f cts $\Rightarrow f(X_n) \xrightarrow{\text{in P}} f(X)$

$X_n \to X$ every seq X_m has subseq. X_{m_k}
$X_{m_k} \xrightarrow{\text{a.s.}} X$ and so $f(X_{m_k}) \xrightarrow{\text{a.s.}} f(X)$.
but then hold for all seq so $f(X_n) \xrightarrow[\text{in P}]{} f(X)$.

$F_n \Rightarrow F_\infty$ implies $\exists\, Y_n \sim F_n, Y_\infty \sim F_\infty$
s.t. $Y_n \longrightarrow Y_\infty$ a.s.

$X_n \Rightarrow X_\infty \iff \forall$ bdd cts function g
$Eg(X_n) \to Eg(X_\infty)$.

Independence

$$P(A \cap B) = P(A)P(B)$$

Independent σ-fields

$\mathcal{F}_1, \mathcal{F}_2, \ldots$
any choice $A_i \in \mathcal{F}_i$ $\Bigg\}$ $P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$

Independent variables

$\sigma(X_1), \sigma(X_2), \ldots \sigma(X_n)$ independent

$\iff \forall A_1, A_2, \ldots, A_n \quad P\left(\bigcap_{i=1}^{n} \{X_i \in A_i\}\right) = \prod_{i=1}^{n} P(X_i \in A_i)$

$\iff \forall x_1, x_2, \ldots \quad P(X_i < x_i \ \forall i=1,\ldots,n) = \prod_{i=1}^{n} P(X_i < x_i)$

Dynkin π-λ

independence $\xRightarrow{X_i \geq 0 \text{ or } E|X_i| < \infty}$ uncorrelated $\Rightarrow$ variance adds

$E(XY) = EXEY$ $\qquad$ $var(X_1 + \cdots + X_n) = \sum_{i=1}^{n} var(X_i)$

$$var(cX) = c^2 var(X).$$

Weak Law of Large Numbers

iid version WLLN

$X_1, X_2, \ldots$ iid
finite variance
(or just $E|X_i| < \infty$)
$\mu = EX_1$
$S_n = X_1 + \ldots + X_n$

$\Bigg\}$ $\frac{S_n}{n} \xrightarrow{P} \mu$ (in probability)

Pf Sketch (finite variance)

$\forall \varepsilon > 0$

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \underset{\text{Chebyshev}}{\le} \varepsilon^{-2} E\left(\frac{S_n}{n} - \mu\right)^2 = \varepsilon^{-2} \operatorname{var}\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n \varepsilon^2} \to 0 \text{ as } n \to \infty$$

Extension to $E|X_i| < \infty$ uses truncation & triangle arrays.

Alt Pf (characteristic functions)

$\varphi_X(t)$ ch. f. for X.

Taylor Series $\varphi_X(t) = E e^{itX} = 1 + itEX + \underbrace{o(t)}_{\text{goes to 0 faster than } t \text{ as } t \to 0}$

$$\varphi_{\frac{S_n}{n}}(t) = \left(\varphi_{\frac{X}{n}}(t)\right)^n = \varphi_X\left(\frac{t}{n}\right)^n = \left(1 + i\mu\frac{t}{n} + o\left(\frac{t}{n}\right)\right)^n$$

as $n \to \infty$ $o\left(\frac{t}{n}\right) \to 0$ fast so $\varphi_{\frac{S_n}{n}}(t) \to e^{i\mu t} = \varphi_\mu(t)$

so $\frac{S_n}{n} \Rightarrow \mu$ but μ constant so $\frac{S_n}{n} \to \mu$ in probability

Borel–Cantelli Lemmas

Borel–Cantelli lemma

$\sum_{n=1}^{\infty} P(A_n) < \infty$ then

"infinitely often"

$$P(\limsup_{n\to\infty} A_n) = P(\lim_{n\to\infty} \cup_{m=n}^{\infty} A_m) = P(A_n \text{ i.o.}) = 0$$

Second Borel–Cantelli lemma

$\sum_{n=1}^{\infty} P(A_n) = \infty$ and A_n independent

$$P(\limsup_{n\to\infty} A_n) = P(\lim_{n\to\infty} \cup_{n=m}^{\infty} A_m) = P(A_n \text{ i.o.}) = 1$$

Note: BC2 is a partial converse of BC1.
If A_n not independent $A_n = (0, 1/n)$
$\sum_1 P(A_n) = \sum_1 1/n = \infty$ but $P(A_n \text{ i.o.}) = P(\emptyset) = 0 \neq 1$.

Applications

Thm $X_n \xrightarrow{P} X$ if and only if every subsequence X_m has subsequence $X_{m_k} \xrightarrow{a.s.} X$.

Pf choose m_k s.t. $P(|X_{m_k} - X| > \varepsilon_k) \leq 2^{-k}$ then
$\sum_1 P(|X_{m_k} - X| > \varepsilon_k) \leq \sum_1 2^{-k} < \infty$ so $P(|X_{m_k} - X| > \varepsilon_k \text{ i.o.}) = 0$
so $X_{m_k} \xrightarrow{a.s.} X$.

Thm $X_n \xrightarrow{P} X$, f cts $\Longrightarrow$ $f(X_n) \xrightarrow{P} f(X)$
also f bounded $\Longrightarrow$ $Ef(X_n) \to Ef(X)$

Pf use equivalent characterization of $\xrightarrow{P}$ above.

Strong Law of Large Numbers

iid version SLLN

$$\left.\begin{array}{l} X_1, X_2, \ldots \quad \text{iid} \left(\begin{array}{l}\text{pairwise}\\ \text{ind suff}\end{array}\right) \\ EX_i = \mu \left(\begin{array}{l}\text{or } EX_i^+ = \infty \\ EX_i^- < \infty\end{array}\right) \\ EX_i^4 < \infty \left(\begin{array}{l}\text{sufficient} \\ E|X_i| < \infty\end{array}\right) \end{array}\right\} \quad \frac{S_n}{n} \xrightarrow{a.s.} \mu$$

Pf Sketch $(EX_i^4 < \infty)$

Take $\mu = 0$ $(X \mapsto X - \mu)$

$$ES_n^4 = \sum_{i,j,l,k} E X_i X_j X_k X_l \underset{\text{independence}}{=} \sum_i E X_i^4 + \sum{}' E X_i^2 X_j^2$$

$$= n E X_i^4 + \alpha n^2 (E X_i^2)^2$$

Jensen's Inequality

$$\leq C n^2 E X_i^4 \leq C' n^2$$

$$P(\underbrace{|S_n| > n\varepsilon}_{A_n}) \leq (n\varepsilon)^{-4} E|S_n|^4 \leq C'/n^2\varepsilon^4$$

Borel-Cantelli $\Rightarrow$ $P(A_n \text{ i.o.}) = 0$, let $\varepsilon \to 0$

so $\frac{S_n}{n} \xrightarrow{a.s.} 0 = \mu$.

Extension

$X_i \geq 0$ and $EX_i = \infty$ then $\frac{S_n}{n} \xrightarrow{a.s.} \infty$

Pf: $Y_n = X_n 1(X \leq B)$ $\qquad \frac{Y_1 + \cdots + Y_n}{n} \to EY_1$

$Y_n \leq X_n$ so $\frac{Y_1 + \cdots + Y_n}{n} \leq \frac{S_n}{n}$ so $\frac{S_n}{n} \xrightarrow{a.s.} \infty$

$\downarrow$ a.s.

$EY_1 \to EX_1 = \infty$

Kolmogorov 0-1 Law

tail σ-field

$\mathcal{T}$ depends on $X_1, X_2, \ldots$

where $A \in \mathcal{T} \iff A$ immune to finite changes to X_i.

$$\mathcal{T} = \bigcap_n \sigma(X_n, X_{n+1}, \ldots)$$

Examples

$\{\lim_{n\to\infty} S_n \text{ exists}\} \in \mathcal{T}$ [Think $X_2 = X_3 = \ldots = 0$

$\{\limsup S_n > 0\} \notin \mathcal{T}$ then $S_n = X_1 = \begin{cases} 0 \\ 1 \end{cases}$]

$\{A_n \text{ i.o.}\} \in \mathcal{T}$

Kolmogorov's 0-1 law

$\left.\begin{array}{l} \text{If } X_1, X_2, \ldots \text{ independent} \\ A \in \mathcal{T} \end{array}\right\} \quad P(A) = 0 \text{ or } 1$

Pf Idea: Show A independent from itself, so then

$P(A) = P(\bigcap_n A) = P(A)^n \longrightarrow 0$ or 1.

Kolmogorov Maximal Inequality

Kolmogorov Max Inequality

$X_1, X_2, \ldots$ independent
$EX_i = 0 \quad var(X_i) < \infty$
$S_n = X_1 + \cdots + X_n$

$\Bigg\}\; P\left(\max_{1 \le k \le n} |S_k| \ge x\right) \le x^{-2}\, var(S_n)$

Note: Chebyshev's says only $P(|S_n| \ge x) \le x^{-2} var(S_n)$.

Pf Idea: Break space into first time $|S_k| \geq x$

$A_k = \{ |S_k| \ge x$ but $|S_j| < x \;\forall\; j < k \}$

Split ES_n^2 integral by A_K (disjoint)

clever rewriting of quadratic & simplification

Characteristic Functions

Defn

X a random variable

$\varphi(t) = E(e^{itX}) = E(\cos(tX)) + iE(\sin(tX))$

$\hookrightarrow = \int e^{itx} f(x)\,dx$ if X has density f

Properties

- $\varphi(0) = 1$ • $\varphi(-t) = \overline{\varphi(t)}$ • $\varphi_{aX+b}(t) = Ee^{it(aX+b)} = e^{itb}\varphi(at)$
- X_1, X_2 independent $\Longrightarrow \varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t)$

Inversion Formula

- If $\varphi(t) = \int e^{itx}\,d\mu$ (μ is measure for X for example)

$$\mu(a,b) + \frac{1}{2}\mu(\{a,b\}) = \lim_{T\to\infty} (2\pi)^{-1} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)\,dt$$

"can express distribution in terms of characteristic functions"

- If $\int |\varphi(t)|\,dt < \infty$ then μ has bdd cts density $f(y) = \frac{1}{2\pi}\int e^{-ity}\varphi(t)\,dt$

Tightness

Weak convergence

$F_n \Rightarrow F_\infty$, $\mu_n \Rightarrow \mu_\infty$ or $X_n \Rightarrow X_\infty$ means

$\lim_{n\to\infty} F_n(y) = F(y)$ whenever F cts @ y

$X_n \Rightarrow X_\infty \Longleftrightarrow$ $\forall$ bdd cts g, $Eg(X_n) \to Eg(X_\infty)$

Pf: by a.s. char. of weak converg.

Helly's Selection Thm (vague convergence)

F_n sequence of distribution functions

$\exists F_{n_k}$ subsequence s.t.

$F_{n_k} \Rightarrow_v G \leftarrow$
- right continuous
- nondecreasing

(may not satisfy $G(x) \to 0$ as $x \to -\infty$, $G(x) \to 1$ as $x \to \infty$)

Tight

F_n are tight if $\forall \varepsilon > 0$ $\exists M_\varepsilon$ s.t. $\limsup_{n\to\infty} \underbrace{1 - F_n(M_\varepsilon) + F_n(-M_\varepsilon)}_{\mu([-M_\varepsilon, M_\varepsilon]^c)} \le \varepsilon$

Thm: F_n tight $\Longleftrightarrow$ Helly's G is a distribution function

Continuity Theorem

$\mu_1, \mu_2, \ldots$ probability measures

$\varphi_1, \varphi_2, \ldots$ corresponding characteristic functions

(i) $\mu_n \Rightarrow \mu_\infty$ implies $\varphi_n(t) \to \varphi_\infty(t)$ $\forall t$

(ii) $\varphi_n(t) \to \varphi_\infty(t)$ pointwise and $\underline{\varphi_\infty \text{ continuous at } 0}$ then μ_n are tight and $\mu_n \Rightarrow \mu_\infty$ for μ_∞ w/ char f. φ_∞.

Pf Idea: Decay of measure at ∞ bounded by integral with φ near 0, so continuity sends integral to 0 and mass loss too, meaning the μ_n's are tight. And (i) b/c $\varphi(t) = Eg(X)$ for $g(t) = e^{itx}$ bdd and continuous.

IID Central Limit Theorem

iid CLT

$$\left.\begin{array}{l} X_1, X_2, \ldots \text{ iid} \\ EX_i = \mu \\ \text{var}(X_i) = \sigma^2 \in (0,\infty) \end{array}\right\} \quad \frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow \mathcal{X} = \mathcal{N}(0,1)$$

Pf Sketch:

Take $\mu = 0$ and φ char. fun. for X.

Taylor Series $\varphi(t) = 1 + itEX - \frac{t^2 EX^2}{2} + \underbrace{o(t^2)}_{\text{goes to 0 faster than } t^2 \text{ as } t \to 0}$

so $\varphi_{\frac{X}{\sigma\sqrt{n}}}(t) = 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right) \rightsquigarrow \varphi_{\frac{S_n}{\sigma\sqrt{n}}}(t) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n$

as $n \to \infty$ $o(t^2/n) \to 0$ so

$\varphi_{\frac{S_n}{\sigma\sqrt{n}}}(t) \longrightarrow \lim_{n\to\infty}\left(1 + \frac{-t^2/2}{n}\right)^n = e^{-t^2/2} = \varphi_{\mathcal{X}}(t)$

so by continuity theorem $\frac{S_n}{\sigma\sqrt{n}} \Rightarrow \mathcal{X}$ as desired.

convergence type

only weak convergence holds because $\frac{S_n - n\mu}{\sqrt{n}}$ does not converge in probability (and hence not a.s.)

Assume $EX = 0$

$Y_n = \frac{S_{2n}}{\sqrt{2n}} - \frac{S_n}{\sqrt{n}}$

If $\frac{S_n}{\sqrt{n}}$ conv in P then $Y_n \to 0$ in probability (and $Y_n \Rightarrow 0$)

$Y_n = Y_n' + Y_n'' \approx c\frac{S_n}{\sqrt{n}} \Rightarrow c\mathcal{X}$ by CLT so $Y_n \Rightarrow \mathcal{X}$

but $\mathcal{X} \overset{d}{\neq} 0$ contradiction.

Lindeberg-Feller CLT

Lindeberg-Feller Central Limit Theorem

$X_{n,m}$ independent for $1 \le m \le n$

$EX_{n,m} = 0$

(i) $\sum_{m=1}^{n} E(X_{n,m}^2) \to \sigma^2 > 0$

(ii) $\forall \varepsilon > 0 \quad \sum_{m=1}^{n} E(|X_{n,m}|^2 ; |X_{n,m}| > \varepsilon) \to 0$

If $S_n = X_{n,1} + \dots + X_{n,m}$ $\quad S_n \Rightarrow \sigma \chi$ as $n \to \infty$.

Pf Idea $\varphi_{S_n}(t) = \prod_{m=1}^{n} \varphi_{n,m}(t) \to \prod_{m=1}^{n} (1 - \frac{t^2 \sigma_{n,m}^2}{2}) \to \exp(-t^2\sigma^2/2)$

Lindeberg-Feller ⇒ iid

$X_1, X_2, \dots$ iid $EX_i = \mu$ $\quad var(X_i) = \sigma^2 \in (0,\infty)$ (iid CLT set up)

$X_{n,m} = \frac{X_m - \mu}{\sqrt{n}}$ so that (i) $\sum E X_{n,m}^2 = \sigma^2$ and

(ii) $\sum E(|X_{n,m}|^2 ; |X_{n,m}| > \varepsilon) = \sum E(|X_m|^2 ; |X_m| > \varepsilon\sqrt{n}) \to 0$

by dominated convergence and Chebyshev's Inequality

$(\int |X_m|^2 \cdot 1_{A_n} \to \int |X_m|^2 \cdot 1_A) \quad P(A) = 0$ by $(P(|X_m| > \varepsilon\sqrt{n}) < \varepsilon^{-2} n^{-1} var(X) \to 0)$

so $X_{n,1} + \dots + X_{n,n} = \frac{S_n - n\mu}{\sqrt{n}} \Rightarrow \sigma \chi$ as desired.

Poisson Convergence

Thm

$X_{n,m}$ independent Bernoulli Events for $1 \le m \le n$
with $P(X_{n,m}=1) = p_{n,m} = 1 - P(X_{n,m}=0)$.

(i) $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0,\infty)$ as $n \to \infty$

(ii) $\max_{1 \le m \le n} p_{n,m} \to 0$ as $n \to \infty$ ("Law of Rare Events")

Then $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow \text{Poisson}(\lambda) \to P(=k) = \frac{e^{-\lambda}\lambda^k}{k!}$

Pf Idea $\varphi_{S_n}(t) = Ee^{itS_n} \to \exp(\lambda(e^{it}-1)) = \varphi_{P(\lambda)}(t)$

Intuition: Divide interval into n subintervals with at most 1 event (likely) per interval. Probability for each mini interval represented by Bernoulli trial.

Generalization (Poisson Processes)

$P(X_{n,m}=1) = p_{n,m}$ but $P(X_{n,m} \ge 2) = \varepsilon_{n,m}$ $\quad X_{n,m} \in \mathbb{Z}^+$

and $\sum_{m=1}^{n} \varepsilon_{n,m} \to 0$ then $S_n \Rightarrow \text{Poisson}(\lambda)$ too.

Example

$S_n = \#$ of babies born of a fixed day

for small enough time interval, at most 1 baby born.
equally distributed births gives

$P(X_{n,m}=1) = P\left(\begin{matrix}\text{a baby born in}\\ m^{th}\ 1/n \text{ interval}\\ \text{of the particular day}\end{matrix}\right) = \frac{1}{n \cdot 365} = p_{n,m}$

$\sum p_{n,m} = 1/365 = \lambda$ so $S_n \Rightarrow \text{Poisson}(1/365)$.

$\max p_{n,m} = \frac{1}{n365} \to 0$

Random Vector CLT

Defns

- $\vec{X}_n \Rightarrow \vec{X}$ weak convergence when $Ef(\vec{X}_n) \to Ef(\vec{X})$ $\forall$ bdd cts f
- Distribution functions $F(\vec{x} \le \vec{y}) = P(X_1 \le y_1, \ldots, X_d \le y_d)$. and $\vec{X}_n \Rightarrow \vec{X}$ implies $F_n(y) \to F(y)$ for cts pts of F.
- F_n are tight if $\forall \varepsilon > 0$ $\exists M_\varepsilon$ s.t. $\lim_{n\to\infty} \mu_n([-M,M]^d) \ge 1-\varepsilon$
- characteristic functions $\varphi(\vec{t}) = Ee^{i\vec{t}\cdot\vec{x}} = Ee^{i(t_1X_1 + \ldots + t_dX_d)}$
  - still have an inversion formula
  - $\vec{X}_n \Rightarrow \vec{X}$ if and only if $\varphi_n(\vec{t}) \to \varphi(\vec{t})$

Central Limit Theorem in $\mathbb{R}^d$

$X_1, X_2, \ldots$ iid random vectors, $EX_i = \vec{\mu} \in \mathbb{R}^d$ and finite covariance, $\Gamma_{ij} = E(X_i - \mu_i)(X_j - \mu_j) < \infty$.

Then
$$\frac{S_n - n\vec{\mu}}{\sqrt{n}} \Rightarrow N_d(0, \Gamma)$$
multivariate Gaussian w/ mean 0 covariance Γ.

Conditional Expectation

Defns

- If $E|X| < \infty$, $E(X|\mathcal{F})$ is a random variable such that
  (i) $E(X|\mathcal{F}) \in \mathcal{F}$ (is $\mathcal{F}$ measurable)
  (ii) $\forall A \in \mathcal{F}$ $\int_A X\,dP = \int_A E(X|\mathcal{F})\,dP$

  $\}$ is the <u>conditional expectation of X given $\mathcal{F}$</u>

  This exists (by Radon-Nikodym derivatives) and is uniquely defined up to a.e.

  (idea: $\mathcal{F}$ is some potential information, $E(X|\mathcal{F})$ is best guess)
- $P(A|\mathcal{F}) = E(\mathbb{1}_A|\mathcal{F})$

  $P(A|B) = P(A \cap B)/P(B)$

  $E(X|Y) = E(X|\sigma(Y))$

Properties

- $E(aX + Y|\mathcal{F}) \overset{a.s.}{=} aE(X|\mathcal{F}) + E(Y|\mathcal{F})$ $\forall X, Y$ where $E(\cdot|\mathcal{F})$ exists (i.e. $E|X|, E|Y| < \infty$)
- (monotonicity) $X \le Y \Rightarrow E(X|\mathcal{F}) \overset{a.s.}{\le} E(Y|\mathcal{F})$
- $X_n \ge 0$, $X_n \uparrow X$, $EX < \infty \Rightarrow E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$ (monotone convergence)

  $Y_n \downarrow Y$ $E|Y_1|, E|Y| < \infty \Rightarrow E(Y_n|\mathcal{F}) \downarrow E(Y|\mathcal{F})$
- (Jensen's) φ convex, $E|X|, E|\varphi(X)| < \infty \Rightarrow \varphi(E(X|\mathcal{F})) \le E(\varphi(X)|\mathcal{F})$
- "smaller field wins" $\mathcal{F}_1 \subset \mathcal{F}_2 \Rightarrow E(E(X|\mathcal{F}_i)|\mathcal{F}_j) = E(X|\mathcal{F}_1)$ $\forall i,j \in \{1,2\} (i \ne j)$
- If $X \in \mathcal{F}$, $E|X|, E|XY| < \infty$ then $E(XY|\mathcal{F}) = XE(Y|\mathcal{F})$.
- $E(E(X|\mathcal{F})) = EX$

Examples

- $X \in \mathcal{F} \to E(X|\mathcal{F}) = X$, specifically $E(C|\mathcal{F}) = C$ for any constant.
- $X, \mathcal{F}$ independent $\to E(X|\mathcal{F}) = EX$
- $\Omega_1, \Omega_2, \ldots$ disjoint partition of Ω

  $E(X|\sigma(\Omega_1, \Omega_2, \ldots)) = \dfrac{E(X; \Omega_i)}{P(\Omega_i)}$ on each Ω_i.

Regular Conditional Probabilities

Defns

$(\Omega, \mathcal{F}, P)$ probability space, $\mathcal{G} \subset \mathcal{F}$

$X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ measurable

- $\mu : \Omega \times \mathcal{S} \to [0,1]$ is a regular conditional ~~probability~~ distribution if
  (i) $\forall A \in \mathcal{S}$ $\quad \omega \longmapsto \mu(\omega, A)$ is a version of $P(X \in A \mid \mathcal{G})$
  (ii) a.e. ω, $\quad A \longmapsto \mu(\omega, A)$ is a probability measure on $(S, \mathcal{S})$
- $\mu : \mathcal{G} \times \mathcal{S} \to [0,1]$ ~~[illegible]~~ is a regular conditional probability if
  (i) $\forall A \in \mathcal{S}$ $\quad \omega \longmapsto \mu(\omega, A)$ version of $P(A \mid \mathcal{G})$
  (ii) a.e. ω $\quad A \longmapsto \mu(\omega, A)$ is a probability measure on $(S, \mathcal{G})$

Motivation: Tool for computing $E(f(X) \mid \mathcal{F})$

$\mu(\omega, A)$ r.c.d. for X given $\mathcal{F}$, $f : (S, \mathcal{S}) \to (\mathbb{R}, \mathcal{R})$

$E|f(X)| < \infty$ then

$$E(f(X) \mid \mathcal{F}) = \int \mu(\omega, dx)\, f(x) \quad \text{a.s.}$$

Martingales

Defns

- $F_0 \subset F_1 \subset \cdots \subset F_n \subset \cdots$ of σ-fields is a <u>filtration</u> F_n
- X_n sequence of rand. var. with $X_n \in F_n$ is <u>adapted to F_n</u>
- X_n is a <u>martingale</u> if (submartingale) (supermartingale)
  - (i) $E|X_n| < \infty \;\forall n$
  - (ii) $X_n \in F_n \;\forall n$
  - (iii) $E(X_{n+1} | F_n) = X_n \;\forall n$ — submartingale: $E \geq X_n$; supermartingale: $E \leq X_n$

Properties

- $\forall\, n > m \quad E(X_n | F_m) \begin{smallmatrix} \geq \\ = \\ \leq \end{smallmatrix} X_m$ if X_n is submartingale / martingale / supermartingale

- X_n submartingale $\Longleftrightarrow -X_n$ supermartingale
- X_n (sub)martingale (w.r.t F_n), φ (increasing) convex function, $E|\varphi(X_n)| < \infty \;\forall n$ $\Big\}$ $\varphi(X_n)$ is submartingale w.r.t. F_n

  Examples: X_n subM $\to (X_n - a)^+$ submart.

  X_n mart, $E|X_n|^p < \infty$ $\to |X_n|^p$ submart.

Examples:

- $\xi_1, \xi_2, \ldots$ iid
  $S_n = c + \xi_1 + \cdots + \xi_n$
  $F_n = \sigma(\xi_1, \ldots, \xi_n)$
  (Random Walk)

  $\mu = E\xi_i = 0 \longrightarrow S_n$ martingale
  $\mu = E\xi_i \leq 0 \longrightarrow S_n$ supermartingale
  $\mu = E\xi_i \geq 0 \longrightarrow S_n$ submartingale
- Polya's Urn – r red, g green. Each time, pick 1, add c of picked color
  X_n = fraction of greens @ time n is martingale

Predictable Sequences

Defn: H_n is a predictable sequence (w.r.t $\mathcal{F}_n$) if $H_n \in \mathcal{F}_{n-1}$ $\forall n$.
(idea: H_n is a betting scheme, bets can only be decided based on information before the betting round)

$$(H \cdot X)_n = \sum_{m=1}^{n} H_m (X_m - X_{m-1})$$

← if X_n = net money betting a dollar each round then this is net earnings at time n w/ H_n betting

Examples:

- $H_n = 1_{A_n}$ ← only bet when some A_n condition is met
- Classic Martingale Betting
$$H_n = \begin{cases} 2H_{n-1} & X_{n-1} - X_{n-2} = -1 \text{ (lost last bet)} \\ 1 & X_{n-1} - X_{n-2} = 1 \text{ (won last bet)} \end{cases}$$
- Double or Nothing
$$H_n = \begin{cases} -X_{n-1} & X_{n-1} < 0 \\ 0 & X_{n-1} \geq 0 \end{cases}$$
(Either doubles losses or recovers debt to "nothing".)

Facts

- If X_n is (sub/super)martingale, $H_n \geq 0$ and each H_n bounded, and H_n is a predictable sequence, then $(H \cdot X)_n$ is a (sub/super) martingale.

Stopping Times

Defn A random variable s.t. $\{N = n\} \in \mathcal{F}_n \ \forall n < \infty$
(idea: decision to stop computable using information at the time of stopping)

Examples:

- $N = \inf\{n : \text{some condition on } X_n\}$ ← stopping time because $\{N = n\} = \{X_k \text{ fails for } k < n \text{ but } X_n \text{ holds}\}$ and so lies in $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$.

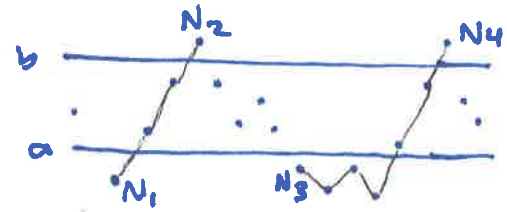
- $N = \inf\{n : X_n = 0\}$
- $N = \inf\{m : X_m \geq \lambda \text{ or } m = n\}$ ← stops early if some X_m exceeds λ.
  ↳ $P(\max_{0 \leq m \leq n} X_m \geq \lambda) = P(X_N \geq \lambda)$

Facts:

- N stopping time, then $\{N > n\} = \left(\bigcup_{k \leq n} \{N = k\}\right)^c \in \mathcal{F}_n$ too.
- X_n (sub/super) martingale $\Longrightarrow X_{N \wedge n}$ is (sub/super) martingale
  ↳ $\min(N, n)$

  Pf:
  $H_n = 1_{N \geq n} \in \mathcal{F}_{n-1}$ so H_n predictable $\rightarrow (H \cdot X)_n$ is (sub/super) martingale and $(H \cdot X)_n = X_{N \wedge n} - X_0$ and adding back X_0 preserves martingale properties.
- $\mathcal{F}_N = \{A : A \cap \{N = n\} \in \mathcal{F}_n \ \forall n\}$

Upcrossing Inequality



Set Up:

X_n submartingale $\quad a < b$

$\{N_{2k-1}\}$ indicates next time $X_n \leq a$ after last time $X_n \geq b$ $= \{\inf\{m > N_{2k-2} : X_m \leq a\}$

$\{N_{2k}\}$ indicates next time $X_n \geq b$ after last time $X_n \leq a$ $= \inf\{m > N_{2k-1} : X_m \geq b\}$

$$H_m = \begin{cases} 1 & N_{2k-1} < m \leq N_{2k} \text{ for some } k \\ 0 & \text{else} \end{cases}$$

is predictable and 1 between "upcrossings".

$U_n = \sup\{k : N_{2k} \leq n\}$ counts (completed) upcrossings up to time n.

Thm X_m submartingale

$$(b-a) E U_n \leq E(X_n - a)^+ - E(X_0 - a)^+$$

Pf: Shift to $Y_m = a + (X_m - a)^+$ (same # of upcrossings, no losses)

$(b-a) E U_n \leq (H \cdot Y_m)$ by picture.

$(1 - H \cdot Y)_n$ also submartingale $\Rightarrow E(1 - H \cdot Y)_n \geq E(1 - H \cdot Y)_0 = 0$

so $E(H \cdot Y)_n \leq E(H \cdot Y)_n + E(1 - H \cdot Y)_n = E(Y_n - Y_0)$

$$\underset{(b-a)EU_n}{\overset{\vee|}{E(H\cdot Y)_n}} \quad \cdots \leq \cdots \quad \underset{E(X_n-a)^+ - E(X_0-a)^+}{\overset{\|}{E(Y_n - Y_0)}}$$

A.S. Martingale Convergence

Result of upcrossing inequality...

Thm

$$\left.\begin{array}{l} X_n \text{ submartingale} \\ \sup E X_n^+ < \infty \end{array}\right\} \Rightarrow \begin{array}{l} X_n \xrightarrow{a.s.} X \\ \text{with } E|X| < \infty. \end{array}$$

Pf:

upcrossing inequality $\rightsquigarrow$ $E\mathcal{U}_n \le \dfrac{EX_n^+ + |a|}{b-a} < \infty$

so $E\mathcal{U}_n \uparrow E\mathcal{U} < \infty$ so for any $a<b$ only fin. many cross.

Then on $P=1$ set $\liminf X_n = \limsup X_n = \lim X_n =: X$

so converges a.s. to X.

Fatou's Lemma + $\sup EX_n^+ < \infty$ gives $EX^-, EX^+ < \infty$.

special case: $\left.\begin{array}{l} X_n \text{ supermartingale} \\ X_n \ge 0 \end{array}\right\} \Longrightarrow \begin{array}{l} X_n \to X \text{ a.s.} \\ \text{and } EX \le EX_0 \end{array}$

Key Example (not L^1-conv):

$S_n = 1 + \xi_1 + \dots + \xi_n$ $\quad P(\xi_i = \pm 1) = 1/2$ iid

$N = \inf\{n : S_n = 0\}$ $\quad X_n = S_{N \wedge n}$ is martingale and $X_n \ge 0$

so $X_n \to X_\infty$ a.s. must be 0 by $E|X_n| = EX_n = EX_0 = 1$

so $X_n \not\to X_\infty = 0$ in L^1.

Bounded Increments

Thm: $X_1, X_2, \ldots$ martingale

$|X_{n+1} - X_n| \le M < \infty$
(bounded increments)

$P(C \sqcup D) = 1.$

$C = \{\lim X_n$ exists & finite$\}$

$D = \{\liminf X_n = -\infty$ and $\limsup X_n = +\infty\}$

Pf: $N = \inf\{n : X_n \le -K\}$ $X_{n\wedge N}$ martingale

$X_{n\wedge N} + K + M \ge 0$ so converges a.s.,.

so $X_{n\wedge N} \xrightarrow{a.s.} X$ and on $\{N = \infty\}$ $X_{N\wedge n} = X_n$

so $X_n \xrightarrow{a.s.} X$ too.

On $\{\liminf X_n > -\infty\}$ then letting $K \to \infty$ eventually we have $\{N = \infty\}$ also.

same holds for $\{\limsup X_n < \infty\}$.

Doob's Decomposition

Thm: X_n submartingale

unique decomposition $X_n = M_n + A_n$

M_n martingale

A_n predictable increasing sequence $(A_0 = 0)$

Pf:

$E(X_n | \mathcal{F}_{n-1}) \geq X_{n-1}$ so define

$A_n - A_{n-1} = E(X_n | \mathcal{F}_{n-1}) - X_{n-1} \geq 0$ (increasing A_n) $(A_0 = 0)$.

$\in \mathcal{F}_{n-1}$ (predictable)

$M_n = X_n - A_n$ check martingale

$E(M_n | \mathcal{F}_{n-1}) = E(X_n - A_n | \mathcal{F}_{n-1}) = E(X_n | \mathcal{F}_{n-1}) - A_n$

$= A_n - A_{n-1} + X_{n-1} - A_n = X_{n-1} - A_{n-1} = M_{n-1}$ ✓.

Application: 2nd Borel-Cantelli II

Thm: $\mathcal{F}_n$ filtration $(\mathcal{F}_0 = \{\emptyset, \Omega\})$ $\quad \{B_n \text{ i.o.}\} = \left\{ \sum_{n=1}^{\infty} P(B_n | \mathcal{F}_{n-1}) = \infty \right\}$

$B_n \in \mathcal{F}_n$ for $n \geq 1$

Pf: $X_n = \sum_{m=1}^{n} 1_{B_m}$ submartingale

By decomposition $M_n = \sum_{m=1}^{n} 1_{B_m} - P(B_m | \mathcal{F}_{m-1})$

and $|M_{n+1} - M_n| \leq 1$ is bounded.

Evaluate both sums on C and D

C: $\sum 1_{B_m} = \infty \iff \sum P(B_m | \mathcal{F}_{m-1}) = \infty$

D: $\sum 1_{B_m} = \infty$ and $\sum P(B_m | \mathcal{F}_{m-1}) = \infty$

(only positive) (only negative)

Doob's Inequality

14.10

Thm:

X_n submartingale
N stopping time
$P(N \le k) = 1$ (for some k)

$$\left.\begin{array}{l} X_n \text{ submartingale} \\ N \text{ stopping time} \\ P(N\le k)=1 \\ \text{(for some } k\text{)}\end{array}\right\} \boxed{EX_0 \le EX_N \le EX_k}$$

Pf:

$X_{N\wedge n}$ submartingale $\longrightarrow EX_0 = EX_{N\wedge 0} \le EX_{N\wedge k} = EX_N$

$K_n = 1_{\{N<n\}}$ predictable

$(K\cdot X)_n = X_n - X_{N\wedge n}$ submartingale ✓

$EX_k - EX_N = E(K\cdot X)_k \ge E(K\cdot X)_0 = 0$

Doob's Inequality:

$$\left.\begin{array}{l} X_n \text{ submartingale} \\ \lambda > 0 \end{array}\right\} \to \lambda P(\underbrace{\max_{0\le m\le n} X_m^+ \ge \lambda}_{=A}) \le EX_n^+$$

Pf:

$N = \inf\{m : X_m \ge \lambda \text{ or } m = n\}$ on $\{\max_{0\le m\le n} X_m^+ \ge \lambda\}$, $X_N \ge \lambda$

$\lambda P(\max X_m^+ \ge \lambda) \le EX_N 1_A \le EX_n 1_A \le EX_n \le EX_n^+$.

Application (Random Walks + Kolmogorov Max Ineq.)

$S_n = \xi_1 + \dots + \xi_n \quad E\xi_m = 0$ (independent)

$\sigma_m^2 = E\xi_m^2 < \infty$

$\to X_n = S_n^2$ is martingale

$\lambda = x^2 > 0$

Doob's Inequality $\to x^2 P(\max_{1\le m\le n} |S_m| \ge x) \le ES_n^2 = \mathrm{var}(S_n)$

which is Kolmogorov's inequality.

L^p convergence (Martingales)

Using integration properties/tricks applied to Doob's Inequality $\lambda P(\max X_m^+ \geq \lambda) \leq EX_n^+$, we get

Thm (L^p Maximal Inequality)

X_n submartingale
$1 < p < \infty$

$$\boxed{E\left(\max_{0 \leq m \leq n} (X_m^+)^p\right) \leq \left(\frac{p}{p-1}\right)^p E(X_n^+)^p}$$

$\hookrightarrow$ depends only on p!

This with a.s. martingale convergence gives

Thm (L^p convergence)

$\left.\begin{array}{l} X_n \text{ martingale} \\ \sup E|X_n|^p < \infty \\ p > 1 \end{array}\right\}$ then $X_n \to X$ a.s. and in L^p.

Pf: $\sup E|X_n|^p < \infty \Rightarrow \sup EX_n^+ < \infty$ so $X_n \to X$ a.s.

L^p max ineq gives $E(\max |X_m|^p) \leq \left(\frac{p}{p-1}\right)^p E|X_n|^p$ ($|\cdot|$ instead of $+$ by double app to X_n and $-X_n$)

$\leq \left(\frac{p}{p-1}\right)^p \sup E|X_n|^p$

$< \infty$

Taking $n \to \infty$ and monotone convergence $\sup |X_n| \in L^p$.

Since $X_n \to X$ a.s. $|X_n - X|^p \leq (2 \sup |X_n|)^p$ and dominated convergence then shows $E|X_n - X|^p \to 0$.

Note: There is no L^1 maximal inequality so L^1 convergence comes about in a different way.

Uniform Integrability

Defn

collection X_n is uniformly integrable if

$$\lim_{M\to\infty}\left(\sup_{n\in\mathbb{N}} E(|X_n|; |X_n|>M)\right) = 0$$

If $M \gg 0$ s.t. $\sup < 1$ then $\sup_{n\in\mathbb{N}} E|X_n| \le \underbrace{M}_{E(|X_n|;|X_n|\le M)} + \underbrace{1}_{E(|X_n|;|X_n|\ge M)} < \infty$.

Example: $X \in L^1$ then $\{E(X|F)\}$ is uniformly integrable.
This helps show $X_n = E(X_0|F_n)$ for backwards martingales.

Sufficient Condition

$\varphi \ge 0$ with $\frac{\varphi(x)}{x} \to \infty$ as $x \to \infty$ (e.g. $\varphi(x) = x^p$)

$E\varphi(|X_i|) \le C$ for all i and fixed constant C

then $\{X_i\}$ are uniformly integrable.

Pf: $E(|X_i|; |X_i|>M) \le \sup\{\frac{x}{\varphi(x)} : x>M\} E(\varphi(|X_i|); |X_i|>M) \le C\sup\{-\} \to 0$

Connection to L^1

$\left.\begin{array}{l} E|X_n| < \infty \ \forall n \\ X_n \to X \text{ in } P \end{array}\right\} \to$ TFAE:
(i) $\{X_n\}$ are uniformly integrable
(ii) $X_n \to X$ in L^1
(iii) $E|X_n| \to E|X| < \infty$

Pf:
(i) $\Rightarrow$ (ii)

$$\varphi_M(x) = \begin{cases} \pm M & |x| \ge M \\ x & |x| \le M \end{cases}$$

$$E|X_n - X| \le \underbrace{E|X_n - \varphi_M(X_n)|}_{\le E(|X_n|; |X_n|>M) \to 0 \text{ since U.I.}} + \underbrace{E|\varphi_M(X_n) - \varphi_M(X)|}_{\to 0 \text{ since } X_n \to X \text{ in } P} + \underbrace{E|\varphi_M(X) - X|}_{\le E(|X|; |X|>M) \to 0} \to 0.$$

since $E|X| < \infty$ from U.I. so choose large M.

L^1 convergence (Martingales)

Thm X_n submartingale TFAE

(i) X_n is uniformly integrable

(ii) X_n converges in L^1 and a.s.

(iii) X_n converges in L^1

(iv) If X_n martingale, $\exists X$ s.t. $X_n = E(X|F_n)$

Pf:

(i) $\Rightarrow$ (ii)

uniform int. gives $\sup E|X_n| < \infty$ so $\sup EX_n^+ < \infty$ gives a.s. conv.

and martingale-ness gives $E|X_n| < \infty$, a.s. $\Rightarrow$ in P so U.I $\rightarrow$ L^1 conv.

(iii) $\Rightarrow$ (i)

$L^1 \rightarrow$ conv in P and $E|X_n| < \infty$ by martingaleness so equiv to U.I.

(iii) $\Rightarrow$ (iv)

If $X_n \rightarrow X$ in L^1 then $X_n = E(X|F_n)$ b/c $E(X_n; A) \rightarrow E(X; A)$

and for $A \in F_n$ and $m > n$ the martingale property gives

$E(X_m; A) = E(X_n; A)$ so $X_n = E(X|F_n)$ $\forall n$.

$\downarrow$

$E(X; A)$

Thm

$F_n \uparrow F_\infty$ (F_n increasing, $F_\infty = \sigma(\bigcup_n F_n)$)

$$E(X|F_n) \longrightarrow E(X|F_\infty) \text{ a.s. and in } L.$$

Cor/Thm:

$X_n \rightarrow X$ a.s. $|X_n| \leq Z$ with $EZ < \infty$ and $F_n \uparrow F_\infty$

$$E(X_n|F_n) \rightarrow E(X|F_\infty) \text{ a.s.}$$

Pf Idea: Use triangle inequality & bound 3 parts.

Reverse Martingale Convergence

Defn
- X_n for $n \le 0$ adapted to F_n filtration is a backwards/reversed martingale if $E(X_{n+1}|F_n) = X_n$ $n \le -1$.

Thm:
X_n backwards martingale $\longrightarrow$ $\lim_{n \to -\infty} X_n = X_{-\infty}$ exists a.s. and L^1.

Pf:
upcrossings gives $EU_\infty < \infty$ so converges in a.s.
Martingale property gives $X_n = E(X_0|F_n)$ which is a uniformly integrable collection, so converges in L^1.
Furthermore, if $F_{-\infty} = \bigcap_n F_n$, then $X_{-\infty} = E(X_0|F_{-\infty})$ by checking conditional expectation properties.

Optional Stopping Theorems

X_n submartingale, N stopping time $\Big\}$ $E(X_0) = E(X_{N\wedge 0}) \le E(X_{N\wedge n})$

Q: when does it hold that $EX_0 \le EX_N$?

Non-Example

X_n random walk on $\mathbb{Z}$, $X_0 = 1$

$N = \inf\{n : X_n = 0\}$

$X_{n\wedge N}$ martingale $\to E(X_{N\wedge n}) = EX_0 = 1$

But $EX_N = 0 \ne EX_0$.

Recall $P(N \le k) = 1$ implies $EX_0 \le EX_N$.

Lemma: X_n submart UI $\Rightarrow X_{N\wedge n}$ U.I.

Pf: U.I. $\to \sup EX^+_{N\wedge n} < \infty$ so conv a.s., $E|X_N| < \infty$. Then split U.I. term by N and show each goes to 0.

Thm: X_n U.I. submartingale $\Rightarrow EX_0 \le EX_N \le EX_\infty$.

Pf: X_n U.I means ~~[illegible]~~ $X_{N\wedge n}$ is U.I. by lemma and so $X_{N\wedge n} \to X_N$ a.s. and in L^1.

$$E(X_0 - X_N) = E(X_0 - X_{N\wedge n} + X_{N\wedge n} - X_N) \quad \forall n$$

$$\le \underbrace{E(X_0 - X_{N\wedge n})}_{\le 0 \text{, because } N\wedge n \text{ is bounded stopping time}} + \underbrace{E|X_{N\wedge n} - X_N|}_{\to 0} \le 0.$$

Thm: X_n submartingale, $E(|X_{n+1} - X_n| \,|\, F_n) \le B$ a.s., $EN < \infty$ $\Big\}$ $X_{N\wedge n}$ U.I. $\Rightarrow EX_0 \le EX_N$.

Pf:

$$|X_{N\wedge n}| \le |X_0| + \sum_{m=0}^{\infty} |X_{m+1} - X_m| 1_{N \ge m} \qquad E(\Sigma) \le BE(N) < \infty$$

so $X_{N\wedge n}$ dominated by integrable rand var, so is U.I.

and so $EX_0 \le EX_N$.

Wald's Identity (Random Walks)

4.16

Thm: (Wald's Equation)

$\xi_1, \xi_2, \ldots$ iid
$E\xi_i = \mu$
$S_n = \xi_1 + \cdots + \xi_n$
N stopping time
$EN < \infty$

$\Bigg\}\; ES_N = \mu EN$

Pf: $X_n = S_n - \mu n$

satisfies optional stopping? $E(|X_{n+1} - X_n| \mid F_n) = \overbrace{E|\xi_i - \mu|}^{=B} < \infty$ ✓

$EN < \infty$ by assumption ✓

so $0 = EX_0 = ES_N = ES_N - \mu EN$.

Application (simple symmetric random walk)

$S_0 = 0 \quad S_n = \xi_1 + \cdots + \xi_n \quad P(\xi_i = \pm 1) = 1/2$

Probability $S_n = -a$ before b?

$N = \inf\{n : S_n = -a \text{ or } b\}$

claim: $EN < \infty$

$EN = \infty \cdot P(N = \infty) + \sum_k P(N > k)$

$P(N > m(a+b)) \le \left(1 - 2^{-(\text{total})}\right)^m$

so $P(N = \infty) = \lim P(N > k) = 0$

and bound sum by geometric series.

claim: optional stopping holds

$E(|S_{n+1} - S_n| \mid F_n) = E(|\xi_{n+1}| \mid F_n)$
$= E|\xi_{n+1}| \overset{=1}{<} \infty$

so $0 = ES_0 = ES_N$.
$= aP(S_N = a) + b\underbrace{P(S_N = b)}_{=1 - P(S_N = a)}$

so solve for $P(S_N = -a) = \frac{b}{a+b}$

Claim: $EN = ab$

$X_n = S_n^2 - \sigma^2 n = S_n^2 - n$ mart.

If opt. stop. $0 = EX_0^2 = ES_N^2 - EN \Rightarrow a^2 P(S_N = a) + b^2 P(S_N = b) = EN = ab$.

Opt. stop. Holds for $N \wedge n$ and $S^2_{N \wedge n}$ bounded and $N \wedge n$ is bounded by geometric random variable so is integrable.

Countable State Space

Defn Markov Property

- $P(X_{n+1}=j \mid \underbrace{X_n=i_n, \dots, X_0=i_0}_{\text{has no effect "memoryless"}}) = P(X_{n+1}=j \mid X_n=i_n) = p(i_n,j)$

- Absorbing States have $p(x,x)=1$ (can never leave)

Examples

- Random Walk $X_n = \xi_1 + \dots + \xi_n$, $\xi_i \in \mathbb{Z}$ with dist μ

  $p(i,j) = P(\xi_n = j-i) = \mu(\{j-i\})$

- Ehrenfest Chain $S=\{0,1,\dots,r\}$ r balls split between two chamber.

  $X_n \to X_{n+1}$: pick ball and move it over

  $X_n = \#$ of balls in a particular side.

  $p(k,k+1) = \frac{r-k}{r}$ $p(k,k-1) = \frac{k}{r}$ $p(i,j)=0$ else.

Markov Properties

Defn

- A <u>transition probability</u> $p: S \times \mathcal{S} \to \mathbb{R}$ satisfies:
  (i) $A \mapsto p(x, A)$ probability measure
  (ii) $x \mapsto p(x, A)$ measurable function
- A <u>Markov chain</u> X_n (w.r.t. $\mathcal{F}_n$) and trans. prob p satisfies:
  $$P(X_{n+1} \in B \mid \mathcal{F}_n) = p(X_n, B)$$
- The <u>Markov Property</u> states that if $m < n$
  $$P(X_n \in B \mid \mathcal{F}_m) = P(X_n \in B \mid X_m).$$
- The <u>Strong Markov Property</u> extends this to stopping times. Let T be a stopping time,
  $$P(X_{T+n} \in B \mid \mathcal{F}_T) = P(X_{T+n} \in B \mid X_T) \text{ on } \{T < \infty\}$$
  $$\mathcal{F}_T = \{A \in \mathcal{F} : \forall n, \{n \geq T\} \cap A \in \mathcal{F}_n\}.$$

Recurrence and Transience

Defns

- $T_y^0 = 0$, $T_y^k = \inf\{n > T_y^{k-1} : X_n = y\}$ time of k^{th} visit to y (excluding X_0)
- $\rho_{xy} = P_x(T_y < \infty)$ probability x goes to y at some point (any visit, i.e. T_y^1)
- x is <u>recurrent</u> if $\rho_{xx} = 1$
  x is <u>transient</u> if $\rho_{xx} < 1$
- C is <u>closed</u> if $x \in C$, $\rho_{xy} > 0 \Longrightarrow y \in C$ $(P_x(X_n \in C) = 1)$
- C is <u>irreducible</u> if $x, y \in D \Longrightarrow \rho_{xy} > 0$

Facts

- $N(y) = \sum_{n=1}^{\infty} 1_{X_n = y}$ = # of (positive) visits to y. y recurrent $\Longleftrightarrow E_y N(y) = \infty$.

<u>Pf:</u>

$$E_x N(y) = \sum_{k=1}^{\infty} P_x(N(y) \geq k) = \sum_{k=1}^{\infty} P_x(T_y^k < \infty) = \sum_{k=1}^{\infty} \rho_{xy}\rho_{yy}^{k-1} = \frac{\rho_{xy}}{1-\rho_{yy}}$$

using strong mark & induction.

so $E_y N(y) = \dfrac{\rho_{yy}}{1-\rho_{yy}}$

recurrent $\Longrightarrow E_y N(y) = \dfrac{1}{1-1} = \infty$ ✓

trans $\Longrightarrow E_y N(y) = \dfrac{0}{1-0} = 0$

- "recurrence is contagious" x recurrent, $\rho_{xy} > 0 \Rightarrow$ y recurrent, $\rho_{yx} = 1$.

<u>Pf:</u>

- Take minimal chain $x \to y$, if $\rho_{yx} < 1$ then

$$\underset{\substack{\uparrow \\ x \text{ recurrent.}}}{0 =} P_x(T_x = \infty) \geq \underbrace{\rho(x, y_1) \cdots \rho(y_{k-1}, y)}_{\text{way to get to } y \text{ from } x} \underbrace{(1 - \rho_{yx})}_{\text{prob never get back}}$$ so $1 - \rho_{yx} = 0 \Longrightarrow \rho_{yx} = 1$

- choose L s.t. $p^L(y,x) > 0$. $\quad p^{L+k+n}(y,y) \geq p^L(y,x)\, p^n(x,x)\, p^k(x,y)$.

$$E_y N(y) = \sum_{n=1}^{\infty} \underbrace{p^n(y,y)}_{E_y(1_{X_n = y})} \geq \sum_{n=1}^{\infty} p^{n+L+k}(y,y) = p^L(y,x)\, p^k(x,y) \underbrace{\sum p^n(x,x)}_{= E_x N(x) = \infty}$$

so $E_y N(y) = \infty \Longrightarrow y$ is recurrent.

Irreducibility

Defns

- C is closed if $x \in C$, $\rho_{xy} > 0 \implies y \in C$
- C is irreducible if $x, y \in D \implies \rho_{xy} > 0$.

Thm:

closed + finite $\implies \exists$ recurrent state

(+ irreducible) $\implies$ (all states recurrent)

Pf: If not, $\rho_{yy} < 1 \; \forall y \in C$. And $E_x N(y) = \dfrac{\rho_{xy}}{1 - \rho_{yy}} < \infty$

since C finite

$$\infty > \sum_{y \in C} E_x N(y) = \sum_{y \in C} \sum_{n=1}^{\infty} p^n(x,y) \overset{\text{Fubini}}{=} \sum_{n=1}^{\infty} \sum_{y \in C} p^n(x,y) = \sum_{n=1}^{\infty} 1 = \infty \;\; ↯$$

And since recurrence is class property, irreducible $\implies \forall$ recurrence.

Decomposition Theorem:

$R = \{\text{all recurrent states}\}$ then $R = \bigcup_i R_i$ each R_i closed & irreducible.

Pf: Partition R into "equivalence classes" $C_x = \{y : \rho_{xy} > 0\}$.

Reflexive by recurrence ✓

Symmetric by "contagion" proof ✓

Transitive? $y \in C_x$ ($\rho_{xy} > 0$) and $z \in C_y$ ($\rho_{yz} > 0$)

then $\rho_{xz} \geq \rho_{xy} \rho_{yz} > 0$ so $z \in C_x$ ✓

Each C_x is ~~irreducible~~ closed by construction, and irreducible by trans.

Stationary Measures

Defns

- a stationary measure satisfies $\mu(y) = \sum_x \mu(x) p(x,y)$
  (this implies by expansion $\mu(y) = \sum_x \mu(x) p^n(x,y)$ also)
- a stationary distribution is also prob. meas. $\left(\sum_x \mu(x) = 1\right)$.
- a reversible measure satisfies $\underbrace{\mu(x)p(x,y) = \mu(y)p(y,x)}_{\text{Detailed Balance Condition}}$

Examples

- stat. meas. (not dist) on simple sym random walk $\mu(x)=1$
  so that $\mu(y) = \sum \mu(x)p(x,y) = 2\mu(y\pm 1)\,p(y\pm 1, y) = 2\frac{1}{2} = 1$
  but $\sum_n \mu(n) = \infty$ so not stat. dist.
- Ehrenfest Chain has stat dist $\mu(x) = 2^{-r}\binom{r}{x}$
  $\mu(x) = 2^{-r}\binom{r}{x} = 2^{-r}\binom{r}{x+1}\underbrace{p(x+1,x)}_{=\frac{x+1}{r}} + 2^{-r}\binom{r}{x-1}\underbrace{p(x-1,x)}_{=\frac{r-(x-1)}{r}}$

Results

- reversible $\Longrightarrow$ stationary (sum condition over all x)
- Existence of Stationary Meas.
  $\exists$ recurrent state $x \Longrightarrow \exists$ stat meas

  "cycle trick"
  $$\boxed{\begin{array}{l} \mu_x = E_x\left(\sum_{n=0}^{T_x-1} 1_{X_n=y}\right) \\ T_x = \inf\{n \geq 1 : X_n = x\} \end{array}}$$
  # of visits to y in $\{0,\dots,T-1\}$

  Pf Idea: shift by a step
  $\sum_z \mu(z)p(z,y) = \#$ visits to y in $\{1,\dots,T\}$
  $= \#$ visits to y in $\{0,\dots,T-1\} = \mu_x(y)$
- Uniqueness of Stationary Meas.
  irreducible & $\exists$ recurrent state $\Rightarrow$ stat. meas unique up to scaling.

  Pf Idea: expand at a
  1) $\nu(z) = \sum_y \nu(y)p(y,z) = \nu(a)p(a,z) + \sum_{y\neq a} \nu(y)p(y,z) \longrightarrow \nu(a)\mu_a(z) + \dots \geq \nu(a)\mu_a(z)$
  2) $\nu(a) = \sum_x \nu(x)\hat{p}(x,a) \geq \sum_x \nu(a)\mu_a(x)\hat{p}^n(x,a) = \nu(a)\mu_a(a) = \nu(a)$ gives term by term $=$
  irreducible $\Rightarrow \exists n$ st $p^n(x,a) > 0$ ... where $p^n(x,a) > 0$.

Stationary Distributions

Defns

- $T_x = \inf\{n \geq 1 : X_n = x\}$ $\quad \rho_{xx} = P_x(T_x < \infty)$
- x is recurrent if $\rho_{xx} = P_x(T_x < \infty) = 1.$
- x is null recurrent if $E_x T_x = \infty$ and
  positive recurrent if $E_x T_x < \infty$

Results

- irred + $\exists$ stat dist $\pi \Longrightarrow$ $\pi(x) = 1/E_x T_x.$

Pf:
1] $\exists$ some y, $\pi(y) > 0$ (since dist). First show y is recurrent by expressing $\sum_x \sum_{n=1}^{\infty} \pi(x) p^n(x,y)$ two ways:
① $\sum_n \pi(y) = \infty$ $\quad$ ② $\sum_x \pi(x) \frac{\rho_{xy}}{1-\rho_{yy}} \leq \frac{1}{1-\rho_{yy}}$ $\Big] \Rightarrow \rho_{yy} = 1$ recurrent.

2] μ_y is a stat. meas and these are unique up to scaling,

$$\sum_x \mu_y(x) = \sum_x E_y\left(\sum_{n=0}^{T-1} 1_{X_n = x}\right) = \sum_x \sum_{n=0}^{\infty} P_y(X_n = x, T_y > n) =$$
$$= \sum_{n=0}^{\infty} P_y(T_y > n) = E_y T_y$$

so $\pi(y) = \frac{\mu_y(y)}{E_y T_y} = \frac{1}{E_y T_y}$ $\quad$ irreducible $\rightarrow$ all states are recurrent so true for all x.

- If irred, TFAE
  (i) $\exists$ positive recurrent state x
  (ii) $\exists$ stationary distribution π
  (iii) all states are positive recurrent

Pf: irred makes pos a class property so (i) $\Leftrightarrow$ (iii)

(i) $\Rightarrow$ (ii) irred + (pos) rec gives μ_x and $\sum_y \mu_x(y) = E_x T_x < \infty$
so can normalize to get a stat. dist.

(ii) $\Rightarrow$ (iii) π stat dist & irred $\rightarrow$ $\pi(y) = 1/E_y T_y$ where $\pi(y) > 0$

so this implies $E_y T_y < \infty$.
and irred implies $\pi(y) > 0$ for all y.

Asymptotic Behavior

Defns

- the <u>total variation distance</u> of measures is
  $$\|\mu - \nu\| = \frac{1}{2}\sum_x |\mu(x) - \nu(x)|$$
  which defines a metric and $\mu_n \to \nu \iff \|\mu_n - \nu\| \to 0$.
  (Markov chain convergence is convergence $p^n(x,y) \to \pi(y)$)
- If x is recurrent, its <u>period, d_x,</u> is $\gcd\underbrace{\{n \geq 1 : p^n(x,x) > 0\}}_{I_x}$
  and it is a class property, i.e. $\rho_{xy} > 0 \Rightarrow d_y = d_x$.
- a chain is <u>aperiodic</u> if it is irreducible w/ $d_x = 1$.

Facts

- periodicity can prevent convergence of $p^n(x,y)$.
  <u>Pf by Example:</u>
  Ehrenfest chain – X_n even/odd $\Rightarrow X_{n+1}$ odd/even so
  $p^n(x,x) = 0$ if n odd, so $\forall$ odd n:
  $$\sum_y |p^n(x,y) - \pi(y)| = |\pi(x)| + \sum_{y \neq x} |p^n(x,y) - \pi(y)| \geq |\pi(x)| > 0$$
  so cannot $\to 0$ over n.
- class property of period ($\rho_{xy} > 0 \Rightarrow d_x = d_y$)
  <u>pf:</u> $p^K(x,y) > 0$ $\quad p^L(y,x) > 0 \longrightarrow p^{L+K}(y,y) \geq p^L(y,x)p^K(x,y) > 0$.
  so $d_y \mid L+K$.
  For any $n \in I_x$, $p^n(x,x) > 0 \to p^{L+K+n}(y,y) \geq p^L(y,x)p^n(x,x)p^K(x,y) > 0$
  so $d_y \mid L+K+n$ so $d_y \mid n \rightsquigarrow d_y \mid d_x$ (by sym $d_x = d_y$).
- If $d_x = 1$ then $\exists m_0$ s.t. $\forall m \geq m_0$ $m \in I_x$ i.e. $p^m(x,x) > 0$.
  <u>pf:</u> $K, K+1 \in I_x \leftarrow$ closed under $+$ $\rightsquigarrow 2K, 2K+1, 2K+2 \rightsquigarrow (K-1)K, (K-1)K+1, \ldots, (K-1)K+K-1$
  gcd $= 1$ $\quad 1 = a_1 i_1 + \ldots + a_n i_n = \underbrace{\sum b_k i_k}_{\in I_x} - \underbrace{\sum c_j i_j}_{\in I_x} \rightsquigarrow \underbrace{\sum b_k i_k}_{=K} = \underbrace{\sum c_j i_j + 1}_{=K+1}$.
  $b, c \geq 0$

Markov Convergence

Theorem

If p is irreducible, aperiodic ($d_x=1$) w/ stat dist π

then $p^n(x,y) \longrightarrow \pi(y)$ [i.e. the chain converges to its stationary dist.]

Proof:

X_n copy of chain w/ $p^n(x,y)$ distribution (starting at X_0).

Y_n copy of chain w/ π distribution

T = stopping time of $X=Y$ = $\inf\{m\geq 1 : X_m = Y_m\}$.

Define chain on $S\times S$ w/ $\bar{p}((x,y),(a,b)) = p(x,a)p(y,b)$.

<u>claim:</u> $\bar{p}$ irred & recurrent

• p irred & aperiodic $\Rightarrow$ $\bar{p}$ irred:

Take any (x_1,y_1) (x_2,y_2). Since p irred $\exists K,L$ s.t. $p^K(x_1,x_2)>0$, $p^L(y_1,y_2)>0$

And aperiodic says $\exists M_1,M_2$ s.t. $\forall m_i > M_i$ $p^m(\cdot_1,\cdot_2)>0$.

Then $M=\max(M_1,M_2)$

$\bar{p}^{L+K+m}((x_1,y_1),(x_2,y_2)) = p^{\overbrace{L+K+m}^{\geq M}}(x_1,x_2)p^{\overbrace{L+K+m}^{\geq M}}(y_1,y_2) > 0$. as desired.

• π stationary $\Rightarrow$ $\bar{p}$ recurrent

$\bar{\pi}(a,b) = \pi(a)\pi(b)$ is stationary $\Rightarrow$ all $\bar{\pi}(y)>0$ recurrent & irred $\Rightarrow$ all y $\bar{\pi}(y)>0$.

<u>Claim.</u> $T<\infty$ a.s.

Well $T < T_{(x,x)}$ and (x,x) recurrent so $P_{(x,x)}(T_{(x,x)}<\infty) \overset{\text{irred}}{=} 1 \rightsquigarrow T_{(x,x)}<\infty$ a.s. $\Rightarrow T<\infty$ a.s.

<u>claim:</u> $p^n \to \pi$

$P(X_n=y, T\leq n) = P(Y_n=y, T\leq n)$ by Markov (strong) Property

$$P(X_n=y) = P(X_n=y, T\leq n) + P(X_n=y, T>n)$$
$$= P(Y_n=y, T\leq n) + P(X_n=y, T>n) \leq P(Y_n=y) + P(X_n=y, T>n)$$

$\rightsquigarrow |P(X_n=y) - P(Y_n=y)| \leq P(X_n=y, T>n) + P(Y_n=y, T>n)$

$\rightsquigarrow \sum_y |p^n(x_0,y) - \pi(y)| = \sum_y |P(X_n=y) - P(Y_n=y)| \leq 2P(T>n) \longrightarrow 0$ so $p^n \to \pi$.