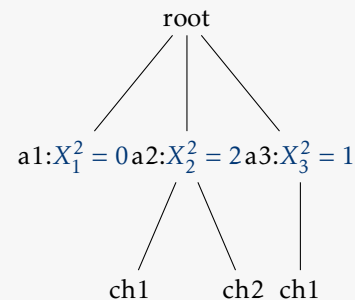


1 Chapter 0 - branching example

$$\begin{aligned} Z_+ &= [0, 1..] \quad N = [1..] \\ f(\theta) &= E(\theta^X) = \sum \theta^k P(X = k) = P(X = 0) + \sum_{k=1} \theta^k P(X = k) \\ f'(\theta) &= E(X\theta^{X-1}) = \sum k\theta^{k-1} P(X = k) \leftarrow \text{differentiate wrt } \theta \\ \text{mean} &= \mu = f'(1) = \sum k P(X = k) \quad f(1) = \sum P(X = k) = 1 \end{aligned}$$

$\{X_r^m\}$ = double series of random variables IID
 X_m^r = the children of animal m contributed to generation r
 $Z_{r+1} = X_1^{r+1} + \dots + X_{Z_r}^{r+1}$ = num children in r+1 generation = sum of the children contributed to generation r+1 by the animals

1.1 my example



so here in the picture for example :
 $Z_0 = 1 \leftarrow Z_0$ = generation 0 has 1 element
and $Z_1 = 3 = X_{Z_0=1}^1 = 3 \leftarrow Z_1$ = generation 1 has 3 children
and $Z_2 = 3 = X_1^2 + X_2^2 + X_{Z_1=3}^2 = 0 + 2 + 1 \leftarrow Z_2$ = num children in generation 2 = 3 children

1.2 calculation

- Distribution of Z_n obtained from generating $\Rightarrow f_n(\theta) = E(\theta^{Z_n}) = \sum \theta^k P(Z_n = k)$
- $f_{n+1}(\theta) = E\theta^{Z_{n+1}} = E\left(E\theta^{Z_{n+1}} | Z_n\right) = \sum E\left(\theta^{Z_{n+1}} | Z_n\right) P(Z_n = k) \leftarrow E\theta^{Z_{n+1}} | Z_n$ is the random variable here
- so now standard calc : $E\left(\theta^{Z_{n+1}} | Z_n\right) = E\left(\theta^{Z_{n+1}} | Z_n = k\right) = E\left(\theta^{X_1^{n+1} + \dots + X_{Z_n=k}^{n+1}} | Z_n = k\right)$
- but (a) num children contributed by animal m to generation n+1 = x_m^{n+1} = not dependent on other siblings x_{\dots}^{n+1} in its own generation (n), or indeed num of such siblings = Z_n , so complicated expression $E\left(\theta^{X_1^{n+1} + \dots + X_{Z_n=k}^{n+1}} | Z_n = k\right)$ = same without the conditioning = absolute expectation = $E\left(\theta^{X_1^{n+1} + \dots + X_{Z_n=k}^{n+1}}\right)$

- and moreover (b) the X_i^{n+1} are independent and independent = multiply
- so $E\left(\theta^{X_1^{n+1} + \dots + X_{Z_n=k}^{n+1}}\right) = E\theta^{X_1^{n+1}} \dots E\theta^{X_{Z_n=k}^{n+1}}$ and all the X_i^j have same distribution and $E\theta^X = f(\theta)$ and since there are $Z_n = k$ of them = $f(\theta)^k$
- so finally $E\left(\theta^{Z_{n+1}} | Z_n\right) = f(\theta)^{Z_n}$
- so tower property $E(\dots) = E\theta^{Z_{n+1}} = E f(\theta)^{Z_n} = f_n(\theta)$ by definition see first line
- so $E\theta^{Z_{n+1}} = f_{n+1}(\theta) = E f(\theta)^{Z_n} = f_n(f(\theta)) = f_n \circ f(\theta)$
- so $f_{n+1}(\theta) = f_n(\theta)$

1.3 extinction

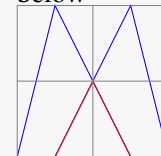
- extinction = zero children in some generation = $\pi_n = P(Z_n = 0) = f_n(0)$
- $\pi_{n+1} = P(Z_{n+1} = 0) = f_{n+1}(0) \leftarrow$ recall $f(\theta) = E(\theta^X) = \sum \theta^k P(X = k) = P(X = 0) + \sum_{k=1} \theta^k P(X = k)$
- so $\pi_{n+1} = f(\pi_n)$ and extinction = 0 eventually = $\pi = \uparrow \lim \pi_n$ and f continuous so $\pi = f(\pi)$ with $f(1) = 1$ $f(0) = P(X = 0)$ slope at 1 = $f'(1) = E(X) = \mu \leftarrow f'(\theta) = E(X\theta^{X-1})$ so $f'(1) = E(X)$
- if $\mu > 1$ then extinction with prob $\pi = f(\pi)$ before hits 1 but if $\mu < 1$ then extinction only when hits 1

1.4 martingale

- Z_{n+1} = num children in (n+1) generation only depends on $Z_n \leftarrow$ Markov
- $E(\theta^{Z_{n+1}} | Z_n) = f(\theta)^{Z_n}$ so differentiate wrt θ
- $E(Z_{n+1} \theta^{Z_{n+1}-1} | Z_n) = Z_n f(\theta)^{Z_n-1} f'(\theta)$
- and with $\theta = 1 \mapsto E(Z_{n+1} | Z_n) = Z_n \overset{\text{whatever}}{f'(1)} = \mu Z_n$
- so $E(Z_{n+1} | Z_n) = \mu Z_n \leftarrow$ not a martingale
- set $M_n = \frac{Z_n}{\mu^n} \leftarrow$ is a martingale - deflated the Z_n
- $E(M_{n+1} | Z_n) = E\left(\frac{Z_{n+1}}{\mu^{n+1}} | Z_n\right) = \frac{1}{\mu^{n+1}} E(Z_{n+1} | Z_n) = \frac{1}{\mu^{n+1}} \mu Z_n = \frac{Z_n}{\mu^n} = M_n$
- $E(M_{n+1} | Z_n) = M_n \leftarrow E(M_{n+1} | Z_n)$ Mn martingale relative to Z_n
- $E(M_{n+1}) = E(M_n) = \dots = E(M_0) = Z_0 = 1$
- \boxed{MCT} because $M_n \geq 0 \mapsto$ EXISTS ALMOST SURELY $M_\infty = \lim M_n = \lim \frac{Z_n}{\mu^n}$
- CAUTION $E(M_n) = 1$ $\lim M_n = M_\infty$ exists but if $\mu < 1 \Rightarrow M_\infty = 0$ so sometimes $E(M_\infty = \lim M_n) = 0 \neq E(M_n) = 1$ ie $E(\lim) < \lim E(\cdot)$
- $\boxed{\text{HOWEVER FATOU IS TRUE}}$ $E(\liminf Y_n) \leq \liminf E(Y_n)$ ie $E(\liminf M_n) = E(M_\infty) = E(0) = 0 \leq \liminf E(M_n) =$

$\liminf 1 = 1$

- see reminder picture about FATOU $E(\liminf) < \liminf E(\cdot)$ below



1.5 distribution of M_∞

- $M_n \xrightarrow{\text{bct}} M_\infty$ so $\exp(-\lambda M_n) \rightarrow \exp(-\lambda M_\infty)$
- \boxed{BCT} $E\left(\exp(-\lambda \frac{Z_n}{\mu^n})\right) < 1 \mapsto E(\exp(-\lambda M_\infty)) = E(\exp(-\lambda \lim M_n)) \xrightarrow{\text{bct}} \lim E(\exp(-\lambda M_n))$ and $E(\exp(-\lambda M_n)) = E\left(\exp(-\lambda \frac{Z_n}{\mu^n})\right) = f_n\left(e^{-\frac{\lambda}{\mu^n}}\right)$
- CONCLUSION $E(\exp(-\lambda M_\infty)) \xrightarrow{\text{bct}} \lim E(\exp(-\lambda M_n)) = \lim f_n\left(e^{-\frac{\lambda}{\mu^n}}\right)$

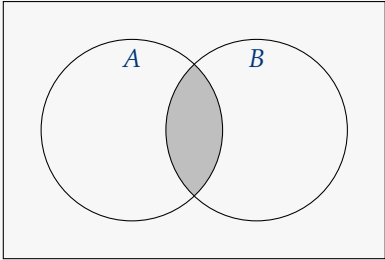
1.6 Now try a particular $P(X=k)$: $P(X = k) = pq^k$

- $P(X = k) = pq^k \Rightarrow f(\theta) = E(\theta^X) = \sum_{k=0} \theta^k P(X = k) = \sum \theta^k pq^k = \frac{p}{1-q\theta}$; and also $\mu = \frac{p}{q}$
- $\pi = f(\pi) \Rightarrow \pi = \frac{p}{1-q\pi} \Rightarrow (\pi = 1, \pi = \frac{p}{q} \iff \frac{p}{q} < 1)$
- Need fn(theta) $f(\theta) = \frac{p}{1-q\theta} \Rightarrow G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & p \\ 1 & -q \end{pmatrix} \Rightarrow G^n = (SAS^{-1})^n \Rightarrow f_n(\theta) = \frac{p\mu^n(1-\theta)+q\theta-p}{q\mu^n(1-\theta)+q\theta-p}$
- $\mu \leq 1 \Rightarrow \pi = 1$ (see earlier) and $\Rightarrow \lim f_n = 1 \Rightarrow$ all in $P(x=0) \Rightarrow$ process dies out
- $\mu > 1 \Rightarrow L(\lambda) = E(\exp(-\lambda M_\infty)) \xrightarrow{\text{bct}} \lim E(\exp(-\lambda M_n)) = \lim f_n\left(e^{-\frac{\lambda}{\mu^n}}\right)$ with fn as above it is found that $\Rightarrow L(\lambda) = \frac{p\lambda+q-p}{q\lambda+q-p} \xrightarrow{\text{Laplace}} L(\lambda) = \pi e^{-\lambda \cdot 0} + \int_0^\infty (1-\pi)^2 e^{-\lambda x} e^{-(1-\pi)x} dx \Rightarrow \begin{cases} P(M_\infty = 0) = \pi \\ P(x < M_\infty < x+dx) = (1-\pi)^2 e^{-(1-\pi)x} \Rightarrow P(M_\infty > x) = (1-\pi) \end{cases}$
- $\mu < 1 \Rightarrow$ process dies out (see above) but what is the distribution when it does not ie what is $E(\theta^{Z_n} | Z_n \neq 0)$? It is $E(\theta^{Z_n} | Z_n \neq 0) = \frac{f_n(\theta) - f_n(0)}{1 - f_n(0)} \leftarrow$ because recall $f_n(\theta) = E(\theta^{Z_n}) = \sum \theta^k P(Z_n = k)$ and we condition by dividing the intersection with the prob of the event $P(Z_n \neq 0)$
- from above can show $\mu < 1 \Rightarrow \lim P(Z_n = k | Z_n \neq 0) = (1-\mu)\mu^{k-1}$
- So MYSTERY explanation as to why for $\mu < 1 \Rightarrow E(M_n) = 1$ but $E(M_\infty) = 0$: for large n it turns out $E(Z_n | Z_n \neq 0) = \sum k P(Z_n = k | Z_n \neq 0) = \sum k(1-\mu)\mu^{k-1} = 1/(1-\mu)$ and (see above the intersection/division) $P(Z_n \neq 0) = 1 - f_n(0) = (1-\mu)\mu^n$ so

$$E(M_n) = E(M_n | Z_n \neq 0) P(Z_n \neq 0) = E\left(\frac{Z_n}{\mu^n} | Z_n \neq 0\right) P(Z_n \neq 0) = \frac{1}{(1-\mu)\mu^n} (1-\mu)\mu^n = 1$$

2 Chapter 1 - Measure spaces - sigma-algebra and pi-systems

- algebra stable under finitely many intersections / unions
- $A \cap B = (A^c \cup B^c)^c$



- σ -algebra stable under finitely many intersections / unions
- $\cap A_n = (\cup A_n^c)^c$
- $Borel(S) = B(S) = \sigma(\text{open sets in } S)$
- $Borel(\mathbb{R}) = B(\mathbb{R}) = \sigma(\pi(R))$ where $\pi(R) := \{(-\infty, x] : x \in \mathbb{R}\}$
- **Proof** $\Leftarrow (-\infty, x] = \cap_n (-\infty, x + 1/n) =$ countable intersections of open sets in B
- **Proof** \Rightarrow every borel is countable union of open intervals so sufficient to show for (a, b) and $(a, b) = \cup_n (a, b - \epsilon/n]$ and $(a, b] = (-\infty, b] \cap (\infty, a)^c$
- Additive measure $A \cap B = \emptyset \Rightarrow m(A \cup B) = m(A) + m(B)$
- Countably Additive measure disjoint sets $F_n \Rightarrow m(\cup F_n) = \sum m(F_n)$
- measure = countably additive $\Sigma \rightarrow [0, \infty]$
- finite measure = $m(A) < \infty$
- σ -finite measure = there is a sequence $\cup F_n = S$ with $m(F_n) < \infty$
- **Measure Extension theorem (Easy)** :
• π -system = family sets stable under finite intersection: $A, B \in I \Rightarrow A \cap B \in I$
• $m_1(S) = m_2(S) < \infty$ and $m_1 = m_2$ on $I \Rightarrow m_1 = m_2$ on Σ
- **Measure Extension theorem (Corollary)** : 2 measures agree on π -system \rightarrow agree on sigma algebra - and recall the case $B = B(R) = \sigma(\pi(R))$
- **Measure Extension theorem (Caratheodory)** : countably additive function m_0 on an algebra on set S can be extended to measure m on sigma (algebra) such that $m_0 = m$ on the algebra, and uniquely so if the measure is finite ie. $m_0(S) < \infty$
- **Application: Lebesgue measure** collection $F = \cup [a_k, b_k]$ of sets on $[0, 1]$ with $m_0(F) = \sum (b_k - a_k)$ and $0 \leq a_1 \leq b_1 \leq \dots \leq b_k \leq 1$ then m_0 is additive, countably additive (not trivial - see proof), and sigma (F) and $\sigma(F) = B(0, 1]$ (see proof above) so now use extension theorem to extend and create Lebesgue measure.
- **Proof sketch** : take collection F_n of disjoint elements with $F = \cup F_n$ and take $G_n = \cup_{k=1}^n F_k$ then obviously $G_n \uparrow F$ and

- $m_0(G_n) = \sum_{k=1}^n F_k$
- then just show that $m_0(G_n) \uparrow m_0(F)$ because then $m_0(F) := \lim m_0(G_n) = \lim m_0(\cup_{k=1}^n F_k) := m_0(\cup_{k=1}^\infty F_k)$
- take $H_n = F \setminus G_n$ then $H_n \downarrow \emptyset$ so just show that $m_0(H_n) \downarrow 0$
- that is same as showing if $H_n \downarrow$ and $m_0(H_n) > \epsilon \rightarrow \cap H_n \neq \emptyset$
- to do that, take $\bar{J}_k \subseteq H_k$ and $m_0(H_k \setminus J_k) \leq \epsilon 2^{-k}$ then $m_0(H_n \setminus \cap_{k \leq n} J_k) \leq m_0(\cup_{k \leq n} H_n \setminus J_k) \leq \epsilon \sum 2^{-k} \leq \epsilon$ so $m_0(\cap_{k \leq n} J_k) \geq \epsilon$
- so $\cap_{k \leq n} J_k \neq \emptyset$ so $K_n := \cap_{k \leq n} \bar{J}_k \neq \emptyset$ so $\cap_{k \leq n} H_k \neq \emptyset$
- finally it follows $\cap \bar{J}_k \neq \emptyset$ because \bar{J}_k is compact so choose a point inside K_n and find a subsequence that \rightarrow inside the compact
- end of proof
- inequalities

$$\begin{aligned} \mu(A \cup B) &\leq \mu(A) + \mu(B) \\ \mu(A \cup B) &= \mu(A) + \mu(B) - \mu(A \cap B) \Leftarrow A \cup B = A \cup (B \setminus (A \cap B)) \\ \mu(\cup_i U_i) &\leq \sum_n \mu(U_i) \\ \mu(\cup_i U_i) &= \sum_n \mu(U_i) - \sum \sum_{i < j} \mu(U_i \cap U_j) + \sum \sum \sum_{i < j < k} \mu(U_i \cap U_j \cap U_k) - \dots \end{aligned}$$

• monotone convergence of measures - UP

$$F_n \uparrow F \Rightarrow \mu(F_n) \uparrow \mu(F) \Leftarrow F_n \uparrow F \text{ means } F_n \subseteq F_{n+1} \text{ and } \cup F_n = F$$

- **Proof** $= \mu(F_1 \cup (F_2 \setminus F_1) \cup (F_3 \setminus F_2) \cup \dots \cup F_n \setminus F_{n-1}) = \sum \mu(F_j \setminus F_{j-1}) = \sum_k \mu(G_k) \uparrow \sum_\infty \mu(G_k) = \mu(F)$
- **monotone convergence of measures - DOWN - NEED FINITENESS**

$$F_n \downarrow F \text{ AND } \mu(F_n) \text{ FINITE} \Rightarrow \mu(F_n) \downarrow \mu(F)$$

- countable sum of null sets = 0
- WARNING $H_n = [n, \infty] \rightarrow H_n \downarrow \emptyset$ $\mu(H_n) = \infty$
- WARNING : EXAMPLE of $\cap \cup I_{n,k} \neq \cup \cap I_{n,k}$
- $V = Q \cap [0, 1] = [v_n]$ all rationals in $[0, 1] \subseteq G_k = U_n[v_n - \epsilon_k/2^n, v_n + \epsilon_k/2^n] = \cup I_{n,k}$ with $\epsilon_k \downarrow 0$ ie doubly infinite sum
- $V =$ countable union = measure 0
- $H = \cap G_k = \emptyset =$ measure 0 and obviously $V \subseteq H$
- however (not proved - Baire category theorem) H is **uncountable**
- cannot have countable = uncountable
- so $H = \cap G_k = \cap \cup I_{n,k} \neq \cup \cap I_{n,k} = V$
- so careful interchanging things:
- END of WARNING

3 Chapter 2 - Events, limsup, liminf

- Ω = sample space = set of outcomes, ω sample point = single outcome, $\mathcal{F} := \sigma$ -field on Ω is family of events
- ω chosen 'at random' according to law \mathbb{P} , for $F \subseteq \mathcal{F}$ then $\mathbb{P}(F)$ = probability of F

$$\text{Example coin toss } F := \left[\omega : \frac{\#(k \leq n : w_k = H)}{n} \rightarrow \frac{1}{2} \right]$$

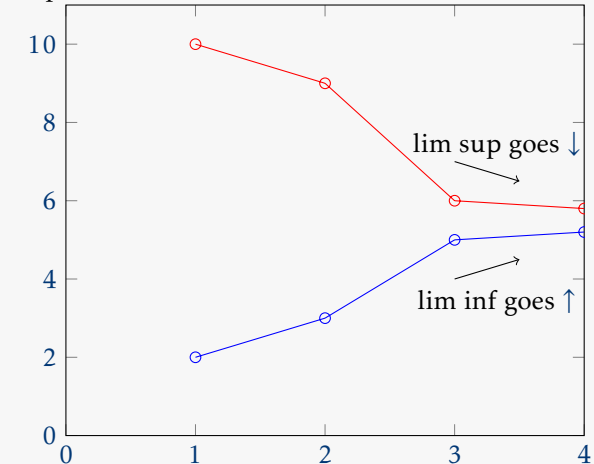
- almost surely : $F = \left[\omega : S(\omega) = \text{true} \right]$ in \mathcal{F} and has $\mathbb{P}(F) = 1$
- $\mathbb{P}(F_n) = 1 \rightarrow \mathbb{P}(\cap F_n) = 1 \Leftarrow \mathbb{P}(F_n^c) = 0 \Rightarrow \mathbb{P}(\cup F_n^c) = 0$ with $\left[\cup F_n^c \right]^c = \cap F_n$

- **EXAMPLE / WARNING:** $F_\alpha = \left[\omega : \frac{\#(k \leq n : w_{\alpha(k)} = H)}{n} \rightarrow \frac{1}{2} \right] \Rightarrow P(F_\alpha) = 1$ for all $\alpha(k)$ sequences that go from 0...1 however $\cap_{\alpha \text{ sequences}} F_\alpha = \emptyset \Leftarrow \forall \alpha \quad \exists \omega \dots$

• LIM SUP, LIM INF

$$\limsup x_n := \inf_m \left(\sup_{n \geq m} x_n \right) = \downarrow \left(\sup_{n \geq m} x_n \right) \leftarrow \text{eventual inf of all the sups}$$

$$\liminf x_n := \sup_m \left(\inf_{n \geq m} x_n \right) = \uparrow \left(\inf_{n \geq m} x_n \right) \leftarrow \text{eventual sup of all the infs}$$



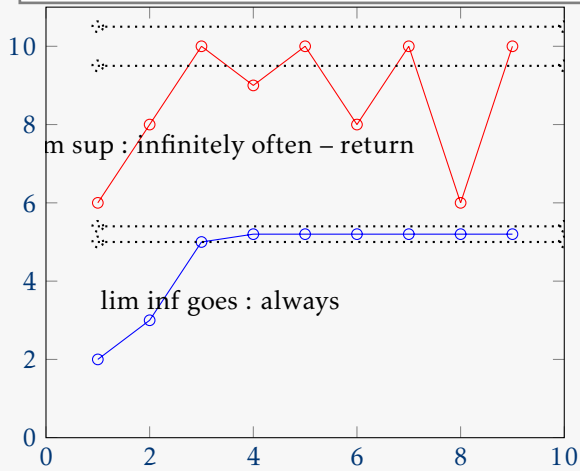
- **LIMIT** $\limsup = \liminf \Leftarrow$ limits exists
- **LIM SUP, LIM INF - for SETS**

• LIM SUP

1. E_n infinitely often
2. $:= \limsup E_n = \inf_m \left(\sup_{n \geq m} E_n \right)$
3. $= \downarrow \left(\sup_{n \geq m} E_n \right)$
4. $= \cap_m \cup_{n \geq m} E_n$
5. $= \forall m, \exists n(\omega) \geq m$ such that $\omega \in E_n(\omega)$

- for each ω , there exists $n \rightarrow$ there is always one $\omega \rightarrow$ infinitely often
- for each ω , there exists $n = \cup$
- **LIM INF**

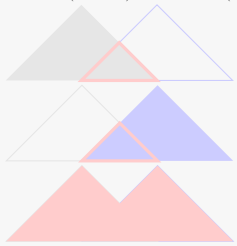
$$\begin{aligned} E_n \text{ eventually } &:= \\ \liminf E_n &= \sup_m \left(\inf_{n \geq m} E_n \right) \\ &= \uparrow \left(\inf_{n \geq m} E_n \right) \\ &= \cup_m \cap_{n \geq m} E_n \\ &= \text{all the } m(\omega) \text{ that exist, such that every } n(\omega) \geq m(\omega) : \omega \in E_n(\omega) \end{aligned}$$



- FATOU 1** $P(\limsup E_n) > \limsup P(E_n)$
- proof: $G_m = \cup_{n \geq m} E_n \rightarrow G_m \downarrow \limsup E_n \rightarrow \text{meas}(G_m) \downarrow \text{meas}(\limsup E_n) \rightarrow P(G_m) \downarrow P(\limsup E_n)$
- and also $P(G_m) = P(\cup_{n \geq m} E_n) \rightarrow P(G_m) \geq \sup_{n \geq m} P(E_n) \rightarrow P(G_m) \geq \liminf_{m \rightarrow \infty} \sup_{n \geq m} P(E_n) := \limsup P(E_n)$
- so $\rightarrow P(\limsup E_n) \geq \limsup P(E_n)$
- example: (x_1, x_2) sequence = $(0,1), (1,0), (0,1), (1,0), \dots$ then $\sum x_1 + x_2 = 1 \rightarrow \limsup \sum = 1$, but $(\limsup x_1, \limsup x_2) = (1,1) \rightarrow \sum(\limsup) = 1+1 = 2 \rightarrow \limsup(\sum) = 1 < \sum(\limsup) = 2 \rightarrow \limsup \int < \int \limsup$

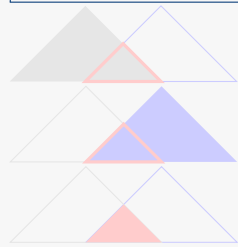
- example:
- $\text{meas}(\limsup E_n) > \limsup \text{meas}(E_n)$

$\text{meas}(\max) > \max(\text{meas})$

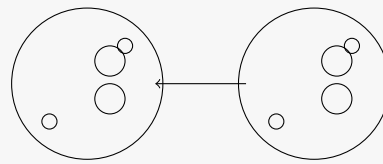


- BOREL-CANTELLI 1** $\sum P(E_n) < \infty \rightarrow P(\limsup E_n) = 0$ \leftarrow $\limsup E_n$ is a set, and has a measure/prob and here prob/measure = 0. basically if the prob ifs finite \rightarrow things that happen i.o have prob 0 (logical!)
- proof: $P(G) \leq P(G_m = \cup_{n \geq m} E_n) \leq \sum_{n \geq m} P(E_n)$
- FATOU 2** $P(\liminf E_n) < \liminf P(E_n)$
- example: (x_1, x_2) sequence = $(0,1), (1,0), (0,1), (1,0), \dots$ then $\sum x_1 + x_2 = 1 \rightarrow \liminf \sum = 1$, but $(\liminf x_1, \liminf x_2) = (0,0) \rightarrow \sum(\liminf) = 0+0 = 0 \rightarrow \liminf(\sum) = 1 > \sum(\liminf) = 0 \rightarrow \liminf \int > \int \liminf$

- example: $f_n = \chi_{n,\infty} \rightarrow \int \liminf f_n = 0 \leq \liminf \int f_n = \infty$
- hint: $\liminf x_n + \liminf y_n \leq \liminf(x_n + y_n) \leq \liminf x_n + \limsup y_n \leq \limsup(x_n + y_n) \leq \limsup x_n + \limsup y_n$
- $\text{meas}(\liminf E_n) < \liminf \text{meas}(E_n)$



4 Chapter 3 - Random variables



- note: sigma needed on target only!! f generates sigma on $\Omega \Leftrightarrow$ sets generated by f^{-1} on Ω from the target sigma. If there is a sigma on Ω as well, then measurability of f wrt that sigma-algebra \Leftrightarrow sigma(f) \in that sigma-algebra
- $h : S \rightarrow \mathbb{R}$ measurable $\Leftrightarrow h^{-1}(A) \in \Sigma \quad \forall A \in \text{Borel}$
- $h^{-1}(\cup A_\alpha) = \cup h^{-1}(A_\alpha) \quad h^{-1}(A^c) = [h^{-1}(A)]^c$
- $C \in B$ and $\sigma(C) = B \rightarrow h^{-1} : C \rightarrow \Sigma \Rightarrow h$ measurable Proof: take all element B such that $h^{-1}(B) \in \Sigma$, then these elements = σ algebra in Borel by above, and $C \in B$
- $\{h \leq s\} \in \Sigma \Rightarrow h$ measurable Proof: $C = \pi[-\infty, c]$ and $\sigma(C) = \mathbb{R}$ and use above
- sum and product and scaling of measurable = measurable Proof: eg for sum $\{h_1 + h_2 > c\} = \cup_{q \in \mathbb{Q}} \{h_1 > q\} \cap \{h_2 > c - q\}$ countable collection of Σ
- h_n measurable $\Rightarrow \inf h_n, \liminf h_n, \limsup h_n$ measurable Proof: $\{\inf h_n > c\} = \cap_n \{h_n > c\}$, define $L = \liminf h_n \quad L_n = \inf h_n$ then $\{L < c\} = \cap_n \{L_n < c\}$
- $\{s : \lim h_n(s) \in \mathbb{R}\} \in \Sigma$ Proof: the set = $\limsup < +\infty \cap \liminf < -\infty \cap g^{-1}(0)$ where $g := \limsup - \liminf$
- Strong law : $\Lambda = \{\omega : \frac{\#H}{\#tosses} \rightarrow p\} = \{\omega : L^+(\omega) = p\} \cap \{\omega : L^-(\omega) = p\}$ with $L^+ = \limsup S_n/n$, $L^- = \liminf S_n/n$ where $S_n = \text{num Hs}$, $n = \text{num tosses}$
- sigma-algebra on Ω generated by a sequence of functions** $f_n : \Omega \rightarrow \mathbb{R}$ = smallest sigma (E) on Ω such that every function in the sequence = measurable wrt that (E) ie $\sigma(f_n) = \sigma(\omega \subset \Omega : f_n(\omega) \in B)$
- Law of a random variable L_X :

$$\Omega \xrightarrow{X} \mathcal{R}$$

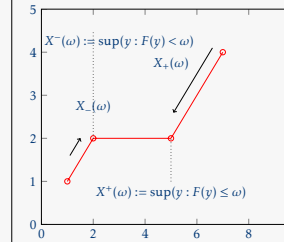
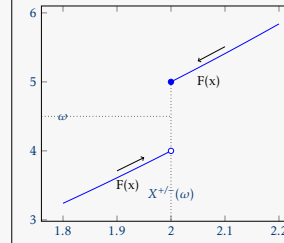
$$[0,1] \xleftarrow{\mathbb{P}} \sigma(X) \text{ or } \mathcal{F} \xleftarrow{X^{-1}} \mathcal{B}$$

$$L_X = X^{-1} \circ \mathbb{P} : \mathcal{B} \rightarrow [0,1]$$

$$F_X(c) := L_X[-\infty, c] = \mathbb{P}(X \leq c) = P(\omega : X(\omega) \leq c)$$

- distribution function F_X is right continuous** : Proof: $P(X \leq c + \frac{1}{n}) \downarrow P(X \leq c)$ because $F_n = G_n \setminus G_{n+1}$ with $G_n \downarrow G \rightarrow \mu(G_n) \downarrow \mu(G)$ because disjoint sets $\mu(G_1 \cup G_n) \uparrow \mu(G)$
- variables with prescribed law (Skorokhod)

$$\begin{aligned} X^+(\omega) &:= \inf\{z : F(z) > \omega\} = \sup\{y : F(y) \leq \omega\} \\ X^-(\omega) &:= \inf\{z : F(z) \geq \omega\} = \sup\{y : F(y) < \omega\} \end{aligned} \quad (1)$$



- $X^-(\omega) = \sup\{y : F(y) \leq \omega\}$
- so $\omega \leq F(c) \Rightarrow X^-(\omega) \leq c$ (1)
- $X^-(\omega) < c \rightarrow \omega \leq F(c)$
- so $\omega \leq F(X^-(\omega))$ because F right continuous
- so $X^-(\omega) \leq c \Rightarrow \omega \leq F(X^-(\omega)) \leq F(c)$ (2)
- so $X^-(\omega) \leq c \Leftrightarrow \omega \leq F(c)$ (1)(2)
- so $X^-(\omega) \leq c \Leftrightarrow \omega \leq F(c)$ (1)(2)
- so $P(X^-(\omega) \leq c) = F(c)$
- actually X^+ and X^- have same distribution and $P(X^+ = X^-) = 1$ Proof: definition of $X^+ \quad \omega < F(c) \Rightarrow X^+(\omega) \leq c$ so $F(c) \leq P(X^+ \leq c)$ and $X^- \leq X^+$ means $(X^+ \neq X^-) = \cup_{q \in \mathbb{Q}} X^- \leq q \leq X^+$ with $P(X^- \leq q \leq X^+) = P(X^+ \leq q) \setminus P(X^- \leq q) = F(q) - F(q) = 0$
- Sigma-algebras generated by collection of random variables** : observe the values $Y_\gamma(\omega) \Leftrightarrow$ observe $I_F(\omega)$ for $F \in \gamma = \sigma(Y_\gamma) \Leftrightarrow$ decide if F has occurred \Leftrightarrow decide if $\omega \in F$
- $\sigma(Y)$ for Y a r.v. $Y^{-1}(\text{Borel}) = \{\omega : Y(\omega) \in B\} \quad B \in \text{Borel}$
- $Z : \Omega \rightarrow \mathbb{R}$ is sigma(Y)-measurable for some $Y : \Omega \rightarrow \mathbb{R} \Leftrightarrow$ there is a Borel function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $Z = f(Y)$
- $Z : \Omega \rightarrow \mathbb{R}$ is sigma(Y_1, Y_2, \dots, Y_n)-measurable for some $Y_i : \Omega \rightarrow \mathbb{R} \Leftrightarrow$ there is a Borel function on $\mathbb{R}^n \quad f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $Z = f(Y_1, Y_2, \dots, Y_n)$
- $Z : \Omega \rightarrow \mathbb{R}$ is sigma($Y_{\text{uncountable}}$)-measurable for some $Y_{\text{uncountable}} : \Omega \rightarrow \mathbb{R} \Leftrightarrow$ there is a Borel function on $\mathbb{R}^{\mathbb{N}}$

$f: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$ such that $Z=f(Y_1, Y_2, Y_n)$ for some countable sequence inside the uncountable set. Note f of the form $\Pi_{\text{uncountable}} \text{Borel}(\mathbb{R})$ not $\text{Borel}(\mathbb{R}^{\text{uncountable}})$, the latter much bigger!

- **MONOTONE CLASS THEOREM** π systems $\rightarrow \sigma$ algebras \Leftrightarrow indicator of elements of π systems \rightarrow general measurable functions
- **MONOTONE CLASS THEOREM**: \mathcal{H} class of bounded functions $S \rightarrow \mathbb{R}$ such that

1. \mathcal{H} vector space
2. constant function $1 \in \mathcal{H}$
3. if $f_n^+ \uparrow f$ and f bounded on S then $f \in \mathcal{H}$ then
4. indicator functions of π -system $\mathcal{I} \in \mathcal{H} \rightarrow$ every bounded $\sigma(\mathcal{I})$ -measurable function $\in \mathcal{H}$

5 Chapter 4 - Independence

- independent sigma algebras $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_n \Leftrightarrow P(G_1, G_2, G_n) = P(G_1)P(G_2)P(G_n)$ for any $G_i \in \mathcal{G}_i$
- independent variables = independent sigmas generated by the variables
- independent events $E_i \Leftrightarrow$ the sigmas $\mathcal{C}_i = \{\emptyset, E_i, \Omega, E_i^c = \Omega \setminus E_i\}$ independent = the rvs I_{E_i} are independent and of course $\mathcal{C}_i = \sigma(I_{E_i})$
- independent on Pi-system \Leftrightarrow independent on sigma(pi system)
- $P(X < x, Y < y) = P(X < x)P(Y < y) \Leftrightarrow$ pi-system(x), pi-system(y) independent \Leftrightarrow the sigma(X), sigma(Y) independents $\Leftrightarrow X, Y$ independents

• **BOREL-CANTELLI BC2IndepInf** : E_n independent , $\sum P(E_n) = \infty \Rightarrow P(\limsup E_n) = 1$

• **BOREL-CANTELLI BC1NoNeedLTI** : $\sum P(E_n) < \infty \Rightarrow P(\limsup E_n) = 0$ (recall - no need for independents)

• **BOREL-CANTELLI BC2IndepInf** : E_n independent , $\sum P(E_n) = \infty \Rightarrow P(\limsup E_n) = 1$ Proof: Compute and show $P\{(\limsup E_n)^c\} = 1$. With $(\limsup E_n)^c = \liminf E_n^c = \bigcup_m \bigcap_{n \geq m} E_n^c$ and $P(\bigcap_{n \geq m} E_n^c) = \prod_{n \geq m} (1 - p_n)$ because independent and limit $r \geq n \geq m$ $r \uparrow \infty$ ok because both sides monotone and $\prod_{n \geq m} (1 - p_n) \leq \exp(-\sum p_n) = 0$ since $1 - x < \exp(-x)$

• **Example** given X_n so that $P(X_n > x) = e^{-x} \rightarrow P(X > \alpha \ln n) = e^{-\alpha \ln n} = n^{-\alpha} = \frac{1}{n^\alpha}$ so $\rightarrow P(X > \alpha \ln n)$ for infinitely many n = 0 or 1 depending if $\alpha > 1, \alpha \leq 1$. So with $\alpha = 1, L := \limsup \frac{X_n}{n} \rightarrow P(L \geq 1) \geq P(\frac{X_n}{\ln n} \geq 1, \text{infinitely often}) = 1$, but $P(L > 1 + \frac{2}{k}) = 0$ because with $\alpha = 1 + \frac{1}{k}, P(L > 1 + \frac{2}{k}) \leq P\{X_n > (1 + \frac{1}{k}) \ln n, \text{infinitely often}\} = 0$ so $P(L=1) = 1$. Also, you can repeat this : $P(X_n > \ln n + \alpha \ln \ln n) = \exp(-(\ln n + \alpha \ln \ln n)) = \frac{1}{n(\ln n)^\alpha}$ = same thing with $\alpha < 1, > 1$

- (cumulative) distribution function $P(X_n \leq x) = F(x)$
- sample path: $\omega \rightarrow X_n(\omega)$
- **Monkey typing Shakespeare** . H =type infinitely many copies, H_k =type $\geq k$ copies, $H_{k,m}$ =type $\geq k$ copies in $[1, m]$, H^m : infinite copies in $[m, \infty[$. independence between $[1, m]$ and $[m, \infty[$ so $P(H_{m,k} \cap H^m) = P(H_{m,k})P(H^m)$, but (1) $H^m = H \quad \forall m$ (2) $H_{m,k} \uparrow H_m$ (3) $H_k \downarrow H$ so $P(H_{m,k} \cap H^m) = P(H_{m,k})P(H^m) \Rightarrow P(H \cap H) = P(H) = P(H)P(H)$ so $P(H)=0, 1$. (Actually with BC2: $P(H)=1$ because if $P(X = x) = \epsilon$, so $E_n = P(\text{type Shakespeare immediately ie in } [n..N+n]) \geq \epsilon^N > 0$ so $\sum P(E_n) = \infty \Rightarrow P(\limsup E_n) = 1$
- Tail algebra of $\{X_n\} := \tau := \bigcap \tau_n$ with $\tau_n = \sigma(X_n, X_{n+1}, \dots)$
- Tail algebra \mathcal{T} has events like $F_1 := \{\lim X_k \text{ exists}\} = \{\omega : \lim X_k(\omega) \text{ exists}\}$, $F_2 := \{\sum X_k \text{ converges}\}$, $F_3 := \{\lim \frac{\sum X_k}{n} \text{ exists}\}$
- rv measurable wrt Tail algebra \mathcal{T} like $\zeta_1 = \limsup \frac{\sum X_k}{n}$
- **Example** : proof that $F_3 \in \mathcal{T}$

$$F_3^n := \lim_k \frac{X_{n+1}(\omega) + X_{n+2}(\omega) + \dots + X_{n+k}(\omega)}{k} : \omega \text{ exists}$$

then for every $n, F_3^n = F_3$. But the X_n, X_{n+1}, \dots are in the \mathcal{T}_n , so $F_3^n \in \mathcal{T}_n$

- **Kolmogorov 0-1 law** : (A) $F \in \mathcal{T} \Rightarrow P(F) = 0, 1$ (B) for any rv η on \mathcal{T} , $P(\eta = c) = 1$ for some $c \in [-\infty, +\infty]$ **Proof (A)** let $\chi_n = \sigma(X_1, X_2, \dots, X_n)$ and $\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \dots)$. Then (1) χ_n, \mathcal{T}_n are independent because pi-system $X_i < x_i$ generates χ_i and pi-system $X_j < x_j, n+1 \leq j \leq n+r, r \in \mathbb{N}$ generates \mathcal{T}_n and these pi-systems are independents. (2) χ_n, \mathcal{T} are independent because $\mathcal{T} \in \mathcal{T}_n$ (3) $\chi_\infty := \sigma(\chi_n), \mathcal{T}$ are independent because $\bigcup_n \chi_n$ is a pi-system since $\chi_n \in \chi_{n+1}$ and it generates χ_∞ (4) $\mathcal{T} \in \chi_\infty$ (5) so \mathcal{T}, \mathcal{T} independent ! (6) so $P(F) = P(F \cap F) = P(F)P(F) \quad \forall F \in \mathcal{T}$ **Proof (B)** just showed $P(\eta \leq c) = 0, 1 \quad \forall c \in \mathbb{R}$ for any $\eta \in \mathcal{T}$ so at some point it goes $0 \rightarrow 1$ so let $c = \sup\{x : P(\eta \leq x) = 0\}$ and then usual $P(\eta \leq c - 1/n) = 0 \quad P(\eta \leq c + 1/n) = 1$ and $P(\cup) = 0$ and $P(\cap) = 1$ so $P(=c) = 1$

- **Note branching example** : M_∞ is in tail algebra and not deterministic but then again the Z_n are not independent
- **Example** let $P(Y_i = -1) = P(Y_i = 1) = \frac{1}{2}$ and let $X_n = Y_0 Y_1 \dots Y_n$. the X_n are independent. Define $\Gamma = \sigma(X_1, X_2, \dots)$ and $\mathcal{T}_n = \sigma(X_n, X_{n+1}, \dots)$. Then we can prove $\Lambda := \bigcap_n \sigma(\mathcal{T}_n, \Gamma) \neq \zeta := \sigma(\bigcap_n \mathcal{T}_n, \Gamma)$ **Proof**: $\mathcal{T}_0 = \sigma(X_1, X_2, \dots), \mathcal{T}_1 = \sigma(X_2, X_3, \dots), \mathcal{T}_{n-1} = \sigma(X_n, X_{n+1}, \dots), Y_0 = \frac{X_n}{Y_1 Y_2 \dots Y_n}$ so (a) $Y_0 \in \Gamma$ and also $\in \mathcal{T}_{n-1} \rightarrow Y_0 \in \Lambda := \bigcap_n \sigma(\Gamma, \mathcal{T}_{n-1})$ (b) if $\Pi = \bigcap$ -stable then independence wrt $\Pi \Leftrightarrow$ sigma(Π), and $G := \bigcap_n \mathcal{T}_n \cap \Gamma$ is \cap -stable, and Y_0 independent of G because $P\{Y_0 \in A \cap g\} = P(Y_0 \in A)P(g), \forall A \in \text{Borel}, \forall g \in G$ because any g

is $g = g_1 \cup g_2, g_1 \in \Gamma, g_2 \in \bigcap_n \mathcal{T}_{n-1}$ and Y_0 indep of Γ and $P(g_2) = 0, 1$ since tail-algebra (Kolmogorov 0-1)

6 Chapter 5 - Integration

- notation : $\int_\Omega f d\mu = \mu(f, \Omega) = \mu(f) = \int_\Omega f(s) \mu(ds)$ and $\int_A f d\mu = \mu(f, A) = \int_A f(s) \mu(ds) = \mu(f I_A)$
- +ve simple = +ve finite sum of indicator functions = $f = \sum_1^m a_k I_{A_k}, A_k \in \Sigma, a_k \in [0, +\infty]$ and $\mu(f) = \sum a_k \mu(A_k)$, general $f \in m\Sigma \rightarrow \mu(f) = \sup \mu(\text{simple}), \text{simple} \leq f$

- **MON** $f_n \in m\Sigma, f_n \uparrow f \Rightarrow \mu(f_n) \uparrow \mu(f) := \int_S f_n(s) \mu(ds) \uparrow \int_S f(s) \mu(ds)$

(Recall **MON** for measure 1.10.a: $F_n \in \Sigma, F_n \uparrow F \Rightarrow \mu(F_n) \uparrow \mu(F)$ \leftarrow recall $(F_n \uparrow F)$ means $F_n \in F_{n+1}, \cup F_n = F$)

- for a given f . can build f_n using staircase function $a^{(r)} = 0, \frac{i-1}{2^r}, r$ for $x = 0, \frac{i-1}{2^r} < x \leq \frac{i}{2^r} \leq r (i \in \mathbb{N}), x > r$ and set $f_n = a^{(n)} \circ f$ Note $a^{(r)}$ is left-continuous so (reverse) $f_n \uparrow f \Rightarrow a^{(r)}(f_n) \uparrow a^{(r)}(f)$

- **MON A.E.** $f_n \in m\Sigma, f_n \uparrow f \text{ A.E.} \Rightarrow \mu(f_n) \uparrow \mu(f)$ **Proof**:

- **FATOU** - illi $\mu(\liminf f_n) \leq \liminf \mu(f_n) \quad \forall f_n \in m\Sigma \leftarrow$ Recall: $P(\liminf E_n) < \liminf P(E_n)$ **Proof**: define $g_k :=$

$$\inf_{n \geq k} f_n \rightarrow \liminf_n f_n \stackrel{(*)}{=} \uparrow \lim g_k, \forall n \geq k : f_n \geq g_k \rightarrow \mu(f_n) \geq \mu(g_k), \Rightarrow \mu(g_k) \leq \inf_{n \geq k} \mu(f_n) \text{ so } \mu(\liminf_n f_n) \stackrel{(*)}{=} \uparrow \lim_k \mu(g_k) \leq \uparrow \lim_k \inf_{n \geq k} \mu(f_n) := \liminf_n \mu(f_n)$$

- **FATOU2** - needs a dominating FINITENESS $\mu(g)$

$$f_n, g \in m\Sigma, f_n \leq g \quad \forall n, \mu(g) < \infty \Rightarrow \mu(\limsup f_n) \leq \mu(g)$$

Proof: $h_n := g - f_n$ then **FATOU1** $\mu(\liminf h_n) \leq \liminf \mu(h_n) \rightarrow \mu\{\liminf(g - f_n)\} \leq \liminf \mu\{g - f_n\}$, now $\limsup f_n = -\liminf(-f_n)$ so removing g (ok since finite): $\mu\{\liminf(-f_n)\} \leq \liminf \mu(-f_n) \rightarrow -\liminf \mu(-f_n) \leq \mu\{-\liminf(-f_n)\} \rightarrow \limsup \mu(f_n) \leq \mu\{\limsup f_n\}$

- $f \in m\Sigma, f = f^+ - f^- := \max(f, 0) - \max(-f, 0), |f| = f^+ + f^-$ with $f^+, f^- \in m\Sigma^+$
- **INTEGRABLE** for $f \in m\Sigma$, f is μ -integrable $:= f \in \mathbb{L}^1(S, \Sigma, \mu)$ if $\mu(|f|) = \mu(f^+) + \mu(f^-) < \infty$, and then we set $\int f d\mu = \mu(f^+) - \mu(f^-)$
- $|\mu(f)| \leq \mu(|f|)$ like **FATOU2**
- **DOM** - L^1 convergence

1. $f_n, f \in m\Sigma$
2. $f_n(s) \rightarrow f(s)$
3. $|f_n(s)| \leq g(s) \quad \forall s \in S, n \in \mathbb{N}$
4. $\mu(g) < \infty$

5. then $\Rightarrow f_n \rightarrow f$ in $L^1(S, \Sigma, \mu)$
6. ie $\mu\{|f_n - f|\} \rightarrow 0$
7. note therefore $\mu(f_n) \rightarrow \mu(f)$

Proof: $|f_n - f| \leq 2g, \mu(g) < \infty \xRightarrow{\text{FATOU2}} \limsup \mu|f_n - f| \leq \mu\{\limsup |f_n - f|\} = \mu(0) = 0 \leftarrow |\mu(f_n) - \mu(f)| = |\mu(f_n - f)| \leq \mu|f_n - f|$

• **SCHEFFE** $f_n, f \in L^{1,+}, f_n \rightarrow f$ a.e then $\mu|f_n - f| \rightarrow 0 \Leftrightarrow \mu(f_n) \rightarrow \mu(f)$ **Proof** $\Rightarrow \mu|f_n - f| \leq \mu|f_n - f| \rightarrow 0$ so done. **Proof** \Leftarrow so suppose $\mu(f_n) \rightarrow \mu(f)$ and by hyp $f_n \rightarrow f$

then (a) $\text{because}(f_n - f)^- \leq f \xRightarrow{\text{DOM}} \mu\{(f_n - f)^-\} \rightarrow 0$ (b) $\mu\{(f_n - f)^+\} = \mu(f_n - f; f_n \geq f) = \mu(f_n - f) - \mu(f_n - f; f_n < f)$ and $|\mu(f_n - f; f_n < f)| \leq \mu(f_n - f)^- \rightarrow 0$

• standard machine: prove for indicator, prove by linearity for SF+ = finsite sum, use (MON) for $m\Sigma+$, set $h = h^+ - h^-$
• $f \in m\Sigma^+, (f\mu)(A) := \mu(f; A) := \mu(fI_A)$ and $f\mu$ is measure.
• $\{h(f\mu)\}(A) := (f\mu)(hI_A) = \mu\{fhI_A\} = \{hf(\mu)\}(A)$

• **RADON-NIKODYM** μ, λ σ -finite on (S, Σ) then if $\mu(f) = 0 \Rightarrow \lambda(F) = 0 \forall F \Rightarrow \lambda = f\mu$ for some $f \in m\Sigma^+$

7 Chapter 6 - Expectation

• $E(X) := \int_{\Omega} X dP = \int_{\Omega} X(\omega)P(d\omega)$ for $X \in L^1 = L^1(\Omega, F, P)$

1. if X_n, X are such that $X_n \xrightarrow{\text{a.s.}} X$ ie $P(X_n \rightarrow X) = 1$
2. **MON-Prob** $0 \leq X_n \uparrow X \Rightarrow E(X_n) \uparrow E(X) \leq \infty \Leftrightarrow f_n \in m\Sigma^+, f_n \uparrow f \Rightarrow \mu(f_n) \uparrow \mu(f) := \int_S f_n(s)\mu(ds) \uparrow \int_S f(s)\mu(ds)$
3. **FATOU-Prob** $X_n \geq 0 \Rightarrow E(X) \leq \liminf E(X_n) \Leftrightarrow \mu(\liminf f_n) \leq \liminf \mu(f_n) \forall f_n \in m\Sigma^+$
4. **DOM-Prob** $|X_n(\omega)| \leq Y(\omega), E(Y) < \infty \Rightarrow E|x_n - X| \rightarrow 0 \Leftrightarrow$ see long one above
5. **SCHEFFE - Prob** $E|X_n| \rightarrow E|X| \Rightarrow E|X_n - X| = 0 \Leftrightarrow f_n, f \in L^{1,+}, f_n \rightarrow f$ a.e then $\mu|f_n - f| \rightarrow 0 \Leftrightarrow \mu(f_n) \rightarrow \mu(f)$
6. **BDD-DOMK** $|X_n(\omega)| \leq K \Rightarrow E|X_n - X| \rightarrow 0 \Leftrightarrow$ Note $E(K) < \infty$ because $P(\Omega) = 1$

• notation $E(X, F) := \int_F X(\omega)P(d\omega) := E(XI_F)$

• **MARKOV**

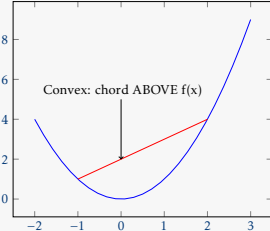
1. $Z \in mF, g(\text{nondec}) : \mathbb{R} \rightarrow [0, \infty]$ Borel
2. then $g \circ Z \in mF+$ and so

3. $\Rightarrow E(gZ) \geq E(gZ; Z \geq c) \geq g(c)P(Z \geq c) \Leftrightarrow$ examples $Z \in mF+, g(x) = x \Rightarrow E(gZ) = E(Z) \geq cP(Z \geq c)$, another example $X \in L1$ so $|X|$ as in first example, so $E(|X|) \geq cP(|X| \geq c)$, final example $g(x) = e^{\theta x}, \theta \geq 0$ applied to $E(g(Z)) \geq g(c)P(Z \geq c) \Leftrightarrow \frac{E(g(Z))}{g(c)} \geq P(Y \geq c)$ is $e^{-\theta c} E(e^{\theta Y}) \geq P(Y \geq c) \Leftarrow$ choose a good c !

• useful results to remember

1. $X \in mF+, E(X) < \infty$ ie integrable, then obviously $P(X < \infty) = 1 \Leftarrow$ remember Markov above $X \in mF+, P(Z \geq c) \leq \frac{E(Z)}{c}$
2. $Z_k \in mF+$ then $E(\sum Z_k) \stackrel{\text{linear}}{=} \sum E(Z_k) \stackrel{\text{MON}}{\leq} \infty$
3. $Z_k \in mF+$ with $E(\sum Z_k) < \infty \Rightarrow \sum Z_k < \infty$ (a.s) and therefore $\Rightarrow Z_k \rightarrow 0$ (a.s) \Leftarrow use the above 2 items
4. note : **BOREL-CANTELLI BC1NoNeedLTI** : $\sum P(E_n) < \infty \Rightarrow P(\limsup E_n) = 0$ is in fact due to the above: let F_k such that $\sum P(F_k) < \infty$ and let $Z_k = I_{F_k}$, then $E(Z_k) = P(F_k)$ ($\Leftarrow P(1_A) = P(A)$ by definition!) and $\sum I_{F_k}$ = number of F_k which occur

• **CONVEX** c convex: $c(px + qy) \leq pc(x) + qc(y)$ with $0 \leq p = 1 - q \leq 1$ example $x^2, |x|, e^{\theta x} \theta \in \mathbb{R}$ and any function $c'' \geq 0$



• **JENSEN**

1. $c : G \rightarrow \mathbb{R}$
2. $E|X| < \infty$
3. $P(X \in G) = 1$
4. $E|c(X)| < \infty$
5. then $\Rightarrow c(E(X)) \leq E(c(X))$ intuition $\Leftarrow c(px + qy) \leq pc(x) + qc(y)$ so $\Leftarrow c(p_1x_1 + p_2x_2 + \dots + p_nx_n) \leq p_1c(x_1) + p_2c(x_2) + \dots + p_nc(x_n)$
6. Proof: c convex $\Delta_{u,v} = \frac{c(v) - c(u)}{v - u}$ so $\Delta_{u,v} \leq \Delta_{v,w}$ so $\uparrow \lim_{u \uparrow v} \Delta_{u,v} = D^-(v) \leq D^+(v) = \downarrow \lim_{w \downarrow v} \Delta_{v,w}$ so for $m \in [D^-(v), D^+(v)]$ then $c(x) \geq m(x - v) + c(v), \forall x \in G$, in particular $v = \mu = E(x) \in G$ and substitute $x \rightarrow X$ (a.s) $c(X) \geq m(X - \mu) + c(v)$ and take $E(\cdot)$ then $E(c(X)) \geq m(E(X) - \mu) + c(\mu)$

• **Norm** : $\|Y\|_p = E(|Y|^p)^{\frac{1}{p}}$ for $E(|X|^p) < \infty$

• monotone $L^p \|Y_p\| \leq \|Y_r\|$ with $1 \leq p \leq r < \infty$ **Proof:** use truncation to make bounded $X_n(\omega) = \{|Y(\omega)| \wedge n\}^p$, bounded so $X_n, X_n^{\frac{r}{p}} \in L^1$, use convex $c(x) = x^{\frac{r}{p}}$ and JENSEN $c(E(X)) \leq E(c(X))$ so $\{E(X_n)\}^{\frac{r}{p}} \leq E(X_n^{\frac{r}{p}}) = E(\{|Y(\omega)| \wedge n\}^r)$ which is $\leq \{E|Y|\}^r$, so $E(X_n)^{\frac{1}{p}} \leq E|Y|^{\frac{1}{r}} = \|Y_r\|$ but (a) (MON) $X_n \uparrow Y^p$ so $E(X_n) \uparrow E(|Y|^p)$ so $E(X_n)^{\frac{1}{p}} \uparrow E(Y^p)^{\frac{1}{p}} = \|Y\|_p$ so $\|Y\|_p \leq \|Y\|_r$

• **SCHWARTZ** $|E(XY)| \leq E(|XY|) \leq \|X\|_2 \|Y\|_2$ **Proof** by truncation : let $X_n = X \wedge n$ and $Y_n = Y \wedge n$ and $0 \leq E\{(aX_n + bY_n)^2\} \geq 0$ ie $E(a^2X_n^2 + 2abX_nY_n + b^2Y_n^2) \geq 0$ with the $B^2 - 4AC < 0$ (since no solution) so $\{2E(X_nY_n)\}^2 < 4E(X_n)^2E(Y_n)^2 < 4E(X)^2E(Y)^2$ and then let $n \uparrow \infty$ with MON like before

• **Variance/Covariance**

1. $VAR = E\{(X - \mu_X)(X - \mu_X)\} = COVAR(X, X)$
2. $COVAR(X, Y) = E\{(X - \mu_X)(Y - \mu_Y)\}$
3. $\langle X, Y \rangle = E(XY)$
4. $\text{correl} = \cos \theta = \frac{\langle U, V \rangle}{\|U\|_2 \|V\|_2}$

• **Completeness of L^p :**

1. L^p is complete (Cauchy series converge and limit inside)
2. X_n Cauchy sequence in $L^p \Leftrightarrow \sup_{r,s \geq k} \|X_r - X_s\|_p \rightarrow 0$
3. X_n Cauchy sequence $\Leftrightarrow \exists X \in L^p \therefore X_n \rightarrow X \in L^p$ ie $\|X_r - X\|_p \rightarrow 0$
4. Proof: choose sequence $\|X_{\alpha_{n+1}} - X_{\alpha_n}\|_p < 2^{-n}$ then with $p=1$ and monotonicity $E|X_{\alpha_{n+1}} - X_{\alpha_n}| = \|X_{\alpha_{n+1}} - X_{\alpha_n}\|_1 \leq \|X_{\alpha_{n+1}} - X_{\alpha_n}\|_p$ so $E \sum |X_{\alpha_{n+1}} - X_{\alpha_n}| < \infty$, so $\sum \{X_{\alpha_{n+1}} - X_{\alpha_n}\}$ converge almost surely (absolutely in fact), so $X_{\alpha_n}(\omega)$ converges, then set $X = \limsup X_n$, so X is measurable and $X_{\alpha_n} \rightarrow X$. Then this X is the one to use for L^p completeness : for $r > \alpha_n$ we have $E(|X_r - X_{\alpha_t}|^p) = \|X_r - X_{\alpha_t}\|_p^p \leq 2^{-np}$ for some $t \geq n$ and $r \geq \alpha_t$. Then with Fatout ('illi') : $E(|X_r - X|^p) \leq \liminf_{t \rightarrow \infty} E(|X_r - X_{k_t}|^p) \leq 2^{-np}$, so $X_r - X$ in L^p , so X in L^p since L^p is a vector space and $X_r \rightarrow X$

• **Law of a rv** $\Lambda_X(B) = P(X \in B)$

• $Eh(X) = \Lambda_X(h) = \int_{\mathbb{R}} h(x)\Lambda_X(dx)$, with $h = 1_B$, this is a definition ie $Eh(X) = E1_B = P(X \in B) = \Lambda_X(B) = \int_{\mathbb{R}} 1_B(x)\Lambda_X(dx) = \int_B \Lambda_X(dx)$

• **PDF of a rv** there is an f such that $P(X \in B) = \int_B f_X(x)dx$ with dx meaning $Leb(dx)$ and $\frac{d\Lambda_X}{dLeb} = f_X$

- HÖLDER (p,q,mult)** let $p, q > 1, \frac{1}{p} + \frac{1}{q} = 1, f \in m\Sigma, f \in L^p, h \in L^q, \mu(|f|^p) = \|f\|_p^p < \infty$ then $|\mu(fh)| \leq \mu(|fh|) \leq \|f\|_p \|h\|_q$
Proof say $f, h \geq 0, \mu(f^p) > 0$ then define a prob measure $P = \frac{f^p \mu}{\mu(f^p)}$ [\Leftarrow recall $\mu(f^p) := \mu(f^p, \Omega) = \int_{\Omega} f^p d\mu$ and also recall $f^p \mu$ is a measure ie $(f^p \mu)(A) := \mu(f^p; A) = \int_A f^p d\mu$ so that the division normalises to =1], now use $u(s) := \frac{h(s)}{f(s)^{p-1}}, f(s) > 0$ otherwise, then because $P(u^q) \leq P(u^q)$ [\Leftarrow recall JENSEN $c(E(X)) \leq E(c(x))$ with $c(x) = x^q$ convex since $q > 1$], then $\mu(|fh|) \leq \|f\|_p \|h\|_q$

- Intuition** Holder (p=q=2) = Cauchy schwartz . Holder = $\|fg\|_1 \leq \|f\|_p \|g\|_p$ Cauchy = $\|fg\|_1 \leq \|f\|_2 \|g\|_2$ ie $(\int fg)^2 \leq \int f^2 \cdot \int g^2$ and norm $\int g^q = \|g\|_q^q, \int f^p = \|f\|_p^p, \int fg = \|fg\|_1$

- HÖLDER** $|E(XY)| \leq E(|XY|) \leq E(|X|^p)^{\frac{1}{p}} E(|Y|^q)^{\frac{1}{q}}$ Proof : $Q(\Lambda) = \frac{E(1_{\Lambda} X^p)}{C=E(X^p)}, Z = \frac{Y}{X^{p-1}} 1_{X>0}, (E_Q(Z))^q \leq E(Z^q)$ then $\frac{1}{C^q} E(XY)^q = \frac{1}{C^q} E(YX^{1-p} X^p)^q = (E_Q(YX^{1-p}))^q \leq E_Q((YX^{1-p})^q) = \frac{1}{C} E((YX^{1-p})^q X^p) = \frac{1}{C} E(Y^q X^{q(1-p)} X^p) = \frac{1}{C} E(Y^q X^{p-q}) = \frac{1}{C} E(Y^q)$ so $E(XY) \leq C^{\frac{q-1}{q}} E(Y^q)^{\frac{1}{q}} = C^{\frac{1}{p}} E(Y^q)^{\frac{1}{q}} = E(X^p)^{\frac{1}{p}} E(Y^q)^{\frac{1}{q}}$

- MINKOWSKI (p,sum,triangle)** $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ with $f, g \in L^p$ **Proof** use Holder $\mu(|f+g|^p) = \mu(|f+g|^{p-1} |f+g|) \leq \mu(|f+g|^{p-1} |f|) + \mu(|f+g|^{p-1} |g|) \leq \|f+g\|_p^{p-1} \|f\|_p + \|f+g\|_p^{p-1} \|g\|_p$ JENSEN x2 $\|f+g\|_p \leq \|f+g\|_p^{p-1} (\|f\|_p + \|g\|_p)$ $\|f+g\|_p \leq \|f\|_p + \|g\|_p$

8 Chapter 7 - Easy strong law

- Independence=Multiply** $E(XY) = E(X)E(Y)$ **Proof:** use staircase function [\Leftarrow Recall $a^{(r)} = 0, \frac{i-1}{2^r}, r$ for $x = 0, \frac{i-1}{2^r} < x \leq \frac{i}{2^r} \leq r(i \in \mathbb{N}), x > r$] and set $f^{(r)} = a^{(r)} \circ f \uparrow f$ (MON) on disjoint partitions and $a^{(r)}(X) = \sum a_i I_{A_i}, a^{(r)}(Y) = \sum b_i I_{B_i}$ on $A_i, B_i \in \sigma(X), \sigma(Y)$ and $P(A_i \cap B_i) = P(A_i)P(B_i)$
- Independence** $\text{Cov}(XY) = 0$ and $\text{Var}(X+Y) = 0$
- Strong Law - simple form - 4th moments -avg=0** X_1, X_2, \dots indeps, $E(X_i) = 0, E(X_i^4) \leq K \Rightarrow$ let $S_n = X_1 + \dots + X_n$ then $P(\frac{S_n}{n} \rightarrow 0) = 1$ ie $\frac{S_n}{n} \rightarrow 0$ a.s \Leftarrow note the average is 0 in this form here. **Proof** terms with single X_i ie $E(X_i X_j^2 X_k) = E(X_i X_j^3) = E(X_i X_j X_k X_l) = 0$ with indep and $E(X_i) = 0$, terms like $E(X_i^2 X_j^2) = E(X_i^2)E(X_j^2) \leq K$ because $E(X_i^2)^2 \leq E(X_i^4) \leq K$ so $E(S_n^4) = E((X_1 + \dots + X_n)^4) = E(\sum_k X_k^4 + 6 \sum \sum_{i < j} X_i^2 X_j^2) \leq nK + 6 \frac{n(n-1)}{2} K \leq 3Kn^2$ and these

are sums of +ve rvs so $E(\sum (\frac{S_n}{n})^4) \leq 3K \sum \frac{1}{n^2} < \infty$ so (again) see sum of +ve rvs) $\sum (\frac{S_n}{n})^4 < \infty$ a.s $\Rightarrow \frac{S_n}{n} \rightarrow 0$ a.s

- Chebyshev** $P(|X - \mu| > c) \leq \frac{\text{Var}(X)}{c^2}$ **Proof** $P(|X - \mu| > c) = P((X - \mu)^2 > c^2) \stackrel{\text{markov}}{\leq} \frac{E(X - \mu)^2}{c^2}$ **Example** Bernoulli $P(X_i = 1) = p, P(X_i = 0) = 1 - p$ then $E(X_i) = 1 \cdot p + 0 \cdot (1 - p) = p$ and $\text{Var}(X_i) = E(X_i^2) - E(X_i)^2 = p - p^2 = p(1 - p)$ so $E(S_n) = np, \text{Var}(S_n) \stackrel{\text{indep.}}{=} \sum \text{Var}(X_i) = np(1 - p) \leq n/4$ so $E(S_n/n) = p, \text{Var}(\frac{S_n}{n}) = \frac{\text{Var} S_n}{n^2} \leq \frac{1}{4n}$ so **Chebyshev** $P(|\frac{S_n}{n} - p| > c) \leq \frac{1}{4nc^2}$

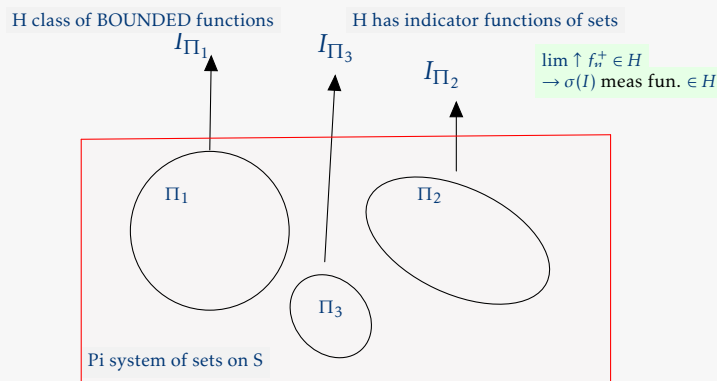
- Example Weierstrass polynomial** $\sup_{s \in [0,1]} |B(x) - f(x)| \leq \epsilon$ for some B polynomial, given f **uniformly** continuous [0,1] **Proof** $P(S_n = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ so define $B_n(x), x \in [0,1]$ by $B_n(p) := E f(\frac{S_n}{n}) = \sum_0^n f(\frac{k}{n}) \binom{n}{k} p^k (1 - p)^{n-k}$ then

$|B_n(p) - f(p)| = |E f(\frac{S_n}{n}) - f(p)| \stackrel{\text{jensen}}{\leq} E(Y_n := |f(\frac{S_n}{n}) - f(p)|)$ with $Z_n := \frac{S_n}{n} - p$ then Z_n small $\rightarrow Y_n$ small and $E(Y_n) = E(Y_n; Z_n \leq \delta) + E(Y_n; Z_n > \delta) \leq \text{small} \times P(Z_n \leq \delta) + 2KP(Z_n > \delta)$ see above $\leq \text{small} \times P(Z_n \leq \delta) + 2K \frac{1}{4n\delta^2}$ ie small for n big enough.

9 Chapter 8 - Product measures

- MONOTONE CLASS THEOREM** : H is a CLASS OF BOUNDED FUNCTIONS such that

- H is a vector space
 - 1 is in H
 - f_n^+ is such that $f_n^+ \uparrow f$ then $f \in H \Leftarrow$ H closed under increasing limits of +ve functions
- THEN
- H contains the I_x of all sets a pi-system on (S) \Rightarrow H contains all bounded σ -(pi-system) measurable functions on S



- statement** : can interchange $\int \int f$ if

- $f = f^+$ ie f positive, in which case $\int \int f^+$ can be ∞

OR IF

- $\int \int |f|$ finite

- where interchange means :

$$\int_{S_1} \left\{ \int_{S_2} f(s_1, s_2) \mu(ds_2) \right\} \mu_1(ds_1) = \int_{S_2} \left\{ \int_{S_1} f(s_1, s_2) \mu(ds_1) \right\} \mu_1(ds_2)$$

- product measure** : $(S_1, \Sigma_1), (S_2, \Sigma_2), S := S_1 \times S_2, \rho_1(s_1, s_2) := s_1, \rho_2(s_1, s_2) := s_2$ then $\Sigma = \Sigma_1 \times \Sigma_2 = \sigma(\rho_1, \rho_2)$ ie sets like $\rho_1^{-1}(B_1) = B_1 \times S_2$ and sets like $\rho_2^{-1}(B_2) = S_1 \times B_2$ with $B_1 \in \Sigma_1, B_2 \in \Sigma_2$
- product measure** : generated by cartesian products (sigma algebra) \times (whole spaces)
- outline** :

- $B_1 \times S_2 \cap B_2 \cap S_1 = B_1 \times B_2$
- $I = \{B_1 \times B_2\}$ is a pi-system which generates $\Sigma = \Sigma_1 \times \Sigma_2$
- then define classes H of functions with useful properties eg

- H class of functions :

- $s_1 \mapsto f(s_1, s_2)$ is Σ_1 measurable
- $s_2 \mapsto f(s_1, s_2)$ is Σ_2 measurable
- $\Rightarrow H = b\Sigma$ (by monotone class theorem)

- define H class of functions :

- $I_1^f(s) = \int_{S_2} f(s_1, s_2) \mu_2(ds_2)$, and $I_1^f()$ is $b\Sigma_1$ measurable
- $I_2^f(s_2) = \int_{S_1} f(s_1, s_2) \mu_1(ds_1)$, and $I_2^f()$ is $b\Sigma_2$ measurable
- $\int_{S_1} I_1^f(s_1) \mu_1(ds_1) = \int_{S_2} I_2^f(s_2) \mu_1(ds_2)$
- $\Rightarrow H = b\Sigma$ (by monotone class theorem)

- Fubini** :

- for $F \in \Sigma$ and indicator $f := I_F$, define $\mu(F) := \int_{S_1} I_1^f(s_1) \mu_1(ds_1) = \int_{S_2} I_2^f(s_2) \mu_2(ds_2)$
- μ = product measure = a set function, and a measure on (S, Σ)
- $\mu = \mu_1 \times \mu_2, (S, \Sigma, \mu) = (S_1, \Sigma_1, \mu_1) \times (S_2, \Sigma_2, \mu_2)$ (\Leftarrow confusing notation since $\Sigma = \Sigma_1 \times \Sigma_2$ is not actually a cartesian product of sigma algebras)
- μ is unique measure $\therefore \mu(A_1 \times A_2) = \mu(A_1) \times \mu(A_2)$ for $A_1 \in \Sigma_1, A_2 \in \Sigma_2$
- if $f \in m\Sigma^+, \mu(f) = \int_{S_1} I_1^f(s_1) \mu_1(ds_1) = \int_{S_2} I_2^f(s_2) \mu_2(ds_2)$
- if $f \in m\Sigma, \mu(|f|) < \infty \stackrel{\text{same}}{\Rightarrow} \mu(f) = \int_{S_1} I_1^f(s_1) \mu_1(ds_1) = \int_{S_2} I_2^f(s_2) \mu_2(ds_2) \Leftarrow f = f^+ - f^-$ etc

- Warning** : must be σ -finite measure :

- $S_1 = S_2 = [0, 1], \Sigma_1 = \Sigma_2 = \mathcal{B}([0, 1]), \mu_1 = \Lambda = \text{Leb}, \mu_2 = m = \text{counting measure}$
- let $\Delta = (x, y) \in S_1 \times S_2 : x = y$ ie $\Delta = \{(x, x) : x \in [0, 1]\}$
- then $I = [0, 1]$

$$\begin{aligned} 4. & \int_I \left(\int_I \mathbf{1}_\Delta(x, y) dm(y) \right) d\lambda(x) \\ &= \int_I \left(\int_I \mathbf{1}_{\{x\}}(y) dm(y) \right) d\lambda(x) \\ &= \int_I m(\{x\}) d\lambda(x) \\ &= \int_I 1 \times d\lambda(x) \\ &= \lambda(I) = 1 \end{aligned}$$

$$\begin{aligned} 5. & \int_I \left(\int_I \mathbf{1}_\Delta(x, y) d\lambda(x) \right) dm(y) \\ &= \int_I \left(\int_I \mathbf{1}_{\{y\}}(x) d\lambda(x) \right) dm(y) \\ &= \int_I \lambda(\{y\}) dm(y) \\ &= \int_I 0 dm(y) = 0 \end{aligned}$$

- indicators on Σ -algebra \Rightarrow can get away with just standard machine, indicators on Π -system \Rightarrow MUST USE Monotone class. Above cannot build F_n of sets $F_n \uparrow \Delta$

• Example: $E(X) = \int_{\mathbb{R}^+} P(X \geq x) dx$

- $\Delta := (\omega, x) : 0 \leq x \leq X(\omega), h := 1_A$
- $$\begin{aligned} & \int_{\Omega} \left(\int_{\mathbb{R}^+} \mathbf{1}_\Delta(\omega, x) d\Lambda(x) \right) dP(\omega) \\ &= \int_{\Omega} \left(\int_0^{X(\omega)} 1 \times d\Lambda(x) \right) dP(\omega) \\ &= \int_{\Omega} X(\omega) dP(\omega) \\ &= E(X) \end{aligned}$$
- $$\begin{aligned} & \int_{\mathbb{R}^+} \left(\int_{\Omega} \mathbf{1}_\Delta(\omega, x) dP(\omega) \right) d\Lambda(x) \\ &= \int_{\mathbb{R}^+} \left(\int_{\Omega} \mathbf{1}_{X(\omega) \geq x} \times dP(\omega) \right) d\Lambda(x) \\ &= \int_{\mathbb{R}^+} P(X \geq x) dx \end{aligned}$$
- so $E(X) = \int_{\mathbb{R}^+} P(X \geq x) dx$

• joint law, joint distribution, joint pdf

- $\text{Law}_{X,Y} : P[(X, Y) \in \Gamma]$
- joint distribution : $F_{X,Y} : P(X \leq x, Y \leq y)$
- joint probability density : $f_{X,Y} : P[(X, Y) \in \Gamma] = \int_{\Gamma} f_{X,Y}(z) \mu(dz) = \int_{\mathbb{R}} \int_{\mathbb{R}} I_{\Gamma}(x, y) f_{X,Y}(x, y) dx dy$
- $f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y) dy = \text{pdf for } X$

• independence : joint law, joint distribution, joint pdf

- X, Y independent
- $\text{Law}_{X,Y} = \text{Law}_X \times \text{Law}_Y$
- $F_{X,Y}(x, y) = F_X(x) \times F_Y(y)$
- $f_{X,Y}(x, y) = f_X(x) \times f_Y(y)$ (when that f exists !)

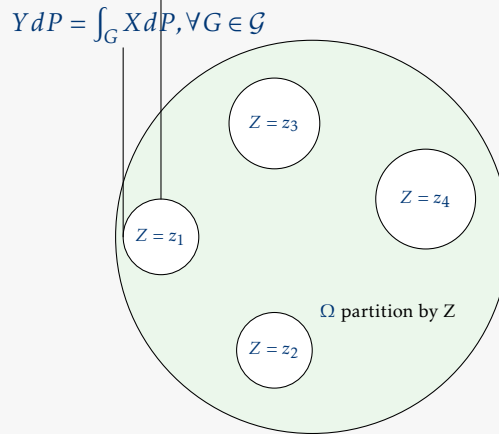
• $B(R)^n = B(R^n)$ because **Proof:**

- $B(R)^n = \sigma(\rho_i : 1 \leq i \leq n) \subseteq B(R^n)$ where ρ_i are the coordinate maps $\rho_i(x_1, \dots, x_n) = x_i$
- $B(R^n)$ is generated by countable union of open hypercubes ie $B(R^n) = \sigma(\Pi_1^n(a_i, b_i))$ and these $\Pi \subset B(R)^n$

10 Chapter 9 - Martingales I : Conditional expectation

- standard stuff $P(X = x_i | Z = z_j) := \frac{P(X=x_i, Z=z_j)}{P(Z=z_j)}$
- standard stuff $E(X|Z = z_j) := \sum_i x_i P(X = x_i | Z = z_j)$

$Y(\omega) := E(X|Z = z_j) := y_j$ constant on this



- $\mathcal{G} = \sigma(Z)$ generated by $Z =$ sets on $\Omega : \{Z \in B\}, B \in \text{Borel} =$ power set of the n $Z = z_i$ atoms $= 2^n$ sets
- $Y(\omega) := E(X|Z = z_j) := y_j$ constant on each Z atom ie measurable
- $$\begin{aligned} \int_{Z=z_j} Y dP &= y_j P(Z = z_j) = E(X|Z = z_j) P(Z = z_j) \\ &= \sum_i x_i P(X = x_i | Z = z_j) P(Z = z_j) = \sum_i x_i P(X = x_i, Z = z_j) = \int_{Z=z_j} X dP \end{aligned}$$
- with $G_j = Z = z_j$ we have $E(YI_{G_j}) = E(XI_{G_j})$ and then adding up all the G_j we have $E(YI_G) = E(XI_G)$ ie $\int_G Y dP = \int_G X dP, \forall G \in \mathcal{G}$
- KOLMOGOROV: Conditional Expectation existence theorem** given X such that:
 - $E|X| < \infty$ ie $|X|$ integrable
 - $\mathcal{G} \subset \mathcal{F}$
 - $\Rightarrow \exists Y$.:

- Y is \mathcal{G} -measurable
- $E|Y| < \infty$ ie $|Y|$ is integrable
- $E(YI_G) = E(XI_G) \forall G \in \mathcal{G}$
- any another such Y' will obey $P(Y = Y') = 1$

• **Proof outline** : 2 methods

- proof method 1: Radon-Nikodym
- proof method 2: Least-squares

(a) prove a.s uniqueness of $E(X|\mathcal{G})$ (assuming it exists) : $E(Y; G) = E(Y'; G) = E(X; G) \Rightarrow E(Y - Y'; G) = 0$. if $Y \neq Y'$ a.s then label Y, Y' such that $P(Y > Y') > 0$. Now consider that $\{Y > Y' + \frac{1}{n}\} \uparrow \{Y > Y'\}$, so for some n , $P(Y - Y' > \frac{1}{n}) > 0$, but that set $\{(Y - Y') > \frac{1}{n}\} \in \mathcal{G}$ since both G -meas. and $E(Y - Y'; Y - Y' > \frac{1}{n}) \geq \frac{1}{n} P(Y - Y' > \frac{1}{n}) > 0 \Rightarrow$ contradiction

(b) prove existence of $E(X|\mathcal{G})$ when $X \in L^2 := L^2(\Omega, \mathcal{F}, \mathbb{P})$: let $\mathcal{K} := L^2(\Omega, \mathcal{G}, \mathbb{P})$ with $\mathcal{G} \subset \mathcal{F}$. \mathcal{K} complete in L^2 -norm, take orthog. proj ie set $Y = \inf_{W \in L^2(\mathcal{K})} E[(X - W)^2]$ which we know exists and is in L^2 , and also it is orthog. proj ie $\langle X - Y, Z \rangle = 0 \forall Z \in L^2(\mathcal{G})$, take $Z = 1_G$ so that $E(X; G) = E(Y; G) \forall G$ ie the l.s inf is as desired.

(c) prove existence of $E(X|\mathcal{G})$ in general ie for $X \in L^1$ (since X integrable) : write $X = X^+ - X^-$, prove if for X^+ , choose bounded bX_n (eg by truncation) such that $bX_n \uparrow X$, bounded so $bX_n \in L^2$, set $Y_n = E(bX_n|\mathcal{G})$, set $Y(\omega) := \limsup Y_n(\omega)$, then $Y \in m\mathcal{G}$ and $Y_n \uparrow Y \xrightarrow{\text{MON-Prob}} E(Y; \mathcal{G}) = E(X; \mathcal{G}) \Leftarrow$ Recall : $\text{MON-Prob } 0 \leq X_n \uparrow X \Rightarrow E(X_n) \uparrow E(X) \leq \infty$. So to finish using Mon-Prob we need to show $0 \leq Y_n \uparrow$.

(d) so to finish we prove : $U^+ \Rightarrow E(U^+|\mathcal{G}) \geq 0$ Proof: let $W = E(U^+|\mathcal{G})$, assume $P(W < 0) > 0$ then for some n $P(G := W < -\frac{1}{n}) > 0$ and $0 \leq E(U^+; \mathcal{G}) = E(W; \mathcal{G}) < -\frac{1}{n}$ contradiction;

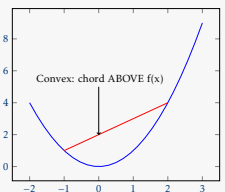
- $E(YI_G) = E(XI_G) \forall G \in \mathcal{G}$
- Notation : such a Y is written $Y = E(X|\mathcal{G})$
- Notation : $E(X|Z) \equiv E[X|\sigma(Z)] \equiv E(X|Z_1, Z_2, \dots)$
- Notation : $E[X|\text{trivial sigma}] = E[X|\sigma(\emptyset, \Omega)] = E(X|\text{no information}) = E(X)$ ie $Y(\omega) = E(X|\mathcal{G})(\omega) = \text{constant } \forall \omega = E(X)$

• conditional prob density: $f_{X|Z}(x|z) := \begin{cases} \frac{f_{X,Z}(x,z)}{f_Z(z)} & \text{if } f_Z(z) \neq 0 \\ 0 & \text{otherwise} \end{cases}$

• Propertie of Conditional Expectations:

- $E[E(X|\mathcal{G})] = E(X)$ **Proof:** $\Leftarrow E(Y; \Omega) \stackrel{\Omega \in \mathcal{G}}{=} E(X; \Omega) = E(X)$

2. X is \mathcal{G} -measurable $\Rightarrow E(X|\mathcal{G}) = XE(1|\mathcal{G}) = X$ **Proof:**
 \Leftarrow use definition $Y := E(X|\mathcal{G})$ is such that $\int_{G \in \mathcal{G}} Y dP = \int_{G \in \mathcal{G}} X dP$ and X will serve as this
3. cLIN $E[aX + bY|\mathcal{G}] = aE[X|\mathcal{G}] + bE[Y|\mathcal{G}]$ **Proof** also use definition
4. c+VE $X \geq 0 \Rightarrow E(X|\mathcal{G}) \geq 0$ **Proof** : see proof for $U^+ \Rightarrow E(U^+|\mathcal{G}) \geq 0$ above
5. cMON+ $0 \leq X_n \uparrow X \Rightarrow E(X_n|\mathcal{G}) \uparrow E(X|\mathcal{G})$ **Proof:** same as above for Least-Squares
6. cFATOU+ $0 \leq X_n \Rightarrow E(\liminf X_n|\mathcal{G}) \leq \liminf E(X_n|\mathcal{G})$ a.s. **Proof:** same as regular MON \rightarrow FATOU proof
7. cDOM $|X_n(\omega)| \leq V(\omega)$ with $E(V) < \infty$ and $X_n \rightarrow X \Rightarrow E(X_n|\mathcal{G}) \rightarrow E(X|\mathcal{G})$ a.s. **Proof:** same as regular FATOU \rightarrow DOM proof
8. cJENSEN $\text{convex} : \mathbb{R} \mapsto \mathbb{R}$ with $E|\text{convex}(X)| < \infty$

$$\Rightarrow \text{convex}(E(X|\mathcal{G})) \leq E(\text{convex}(X)|\mathcal{G})$$
 Proof:
 $\text{convex} \rightarrow \text{continuous} \rightarrow \text{convex}(x) = \sup_n (a_n x + b_n)$ for some sequence $(a_n, b_n) \rightarrow \forall n c(x) \geq a_n x + b_n \xRightarrow{\text{c+VE, cLIN}} \forall n E[c(X)|\mathcal{G}] \geq \sup_n a_n E(X|\mathcal{G}) + b_n = c[E(X|\mathcal{G})]$

9. cJENSEN Corollary $\|E(X|\mathcal{G})\|_p \leq \|X\|_p$ **Proof:** use $c(x) = |x|^p, \forall p \geq 1$ in above and then take $\sqrt[p]{x} = x^{1/p}$ of both sides to get the $\|\cdot\|_p$
10. TOWER $E[E(X|\mathcal{G})|\mathcal{H}] = E(X|\mathcal{H})$ Notation $E[X|\mathcal{G}|\mathcal{H}] := E[E(X|\mathcal{G})|\mathcal{H}]$ **Proof** $\mathcal{H} \subset \mathcal{G}$ so (definition) $E(X|\mathcal{G})$ and X are good candidates for $E(\cdot|\mathcal{H})$
11. TAKE OUT $G \mathcal{G}$ -measurable $\Rightarrow E(GX|\mathcal{G}) = GE(X|\mathcal{G})$ **Proof:** for $Y = E(X|\mathcal{G})$, prove that $E(ZY; G) = E(ZX; G)$ for any $G \in \mathcal{G}$ and any Z which is \mathcal{G} -measurable using 'standard machine' (ie starting with $Z = 1_G$, then linearity for $Z \in \text{SimpleFun}^+(\Omega, \mathcal{G}, \mathbb{P})$, then MON for $Z \in (\mathcal{M}\mathcal{G})^+$

12. \mathcal{H} indep of $\sigma(\sigma(X), \mathcal{G})$
 $\Rightarrow E[X|\sigma(\mathcal{G}, \mathcal{H})] = E(X|\mathcal{G})$ **Proof:** \mathcal{H} indep of $\sigma(\sigma(X), \mathcal{G})$ then $X1_G, H$ independent for $G \in \mathcal{G}$ so

$$E(X; G \cap H) = E[(X1_G)1_H] \stackrel{\text{indep}}{=} E(X1_G)E(1_H) = E(X1_G)P(H).$$
Similarly for $Y = E(X|\mathcal{G}) : E[(Y1_G)(1_H)] = E(Y1_G)E(1_H)$ so $E(X; G \cap H) = E(Y; G \cap H)$ so the measures $\mu_1 : F \mapsto E(X; F), \mu_2 : F \mapsto E(Y; F)$ on $\sigma(\mathcal{G}, \mathcal{H})$ agree on pi-system $G \cap H, G \in \mathcal{G}, H \in \mathcal{H} \Rightarrow$ agree on $\sigma(\mathcal{G}, \mathcal{H})$
 - Example Compute $E(X_1 | \frac{X_1 + X_2 + \dots + X_n}{n})$ with X_1, \dots, X_n iid same as $X : E(X_1|S_n) = E(X_2|S_n) \dots = E(X_n|S_n)$ since $E(X_1; S_n \in B) = \int \dots \int_{s_n \in B} x_1 \mu(dx_1) \mu(dx_2) \dots \mu(dx_n)$; so $E(X_1|S_n) = E(X_2|S_n) \dots = E(X_n|S_n) = \frac{1}{n} E(X_1 + \dots + X_n|S_n) = \frac{1}{n} E(S_n|S_n) = \frac{1}{n} S_n E(1|S_n) = \frac{S_n}{n}$
- 11 Chapter 10 - Martingales II : Martingales
 - filtration : $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n \subseteq \mathcal{F}$
 - $\mathcal{F}_\infty := \sigma(\cup_n \mathcal{F}_n) \subseteq \mathcal{F}$
 - normally $\mathcal{F}_n = \sigma(W_0, W_1, \dots, W_n)$ for some stoch. process W
 - information about ω at time $n : W_0(\omega), W_1(\omega), \dots, W_n(\omega)$
 - $X : \{X_i\}$ adapted to $\{\mathcal{F}_i\}$ if for each n, X_n is \mathcal{F}_n -measurable
 - Martingale definition : X Martingale relative to \mathcal{F}
 1. X adapted to \mathcal{F}
 2. $E|X_n| < \infty$ for all n
 3. $E(X_n|\mathcal{F}_{n-1}) = X_{n-1}$
 - supermartingale decreases : $E(X_n|\mathcal{F}_{n-1}) \leq X_{n-1}$
 - submartingale increases : $E(X_n|\mathcal{F}_{n-1}) \geq X_{n-1}$
 - Examples of martingales :
 1. Sum of iid with $E(X_k) = 0$ ie $S_0 := 0, S_n := X_1 + \dots + X_n, \mathcal{F}_n := \sigma(X_1, X_2, \dots, X_n)$ then martingale **Proof:** $E(S_n|\mathcal{F}_{n-1}) = E(S_{n-1}|\mathcal{F}_{n-1}) + E(X_n|\mathcal{F}_{n-1}) = S_{n-1} + E(X_n) = S_{n-1}$
question : $\lim S_n$ exists when ? [answer: Kolmogorov 3-series theorem]
 2. Product of iid with $E(X_k) = 1$ ie $M_n := X_1 \dots X_n, E(X_i) = 1, \mathcal{F}_n := \sigma(X_1, \dots, X_n)$ is martingale. **Proof:** $E(M_n|\mathcal{F}_{n-1}) = E(M_{n-1}X_n|\mathcal{F}_{n-1}) = M_{n-1}E(X_n|\mathcal{F}_{n-1}) \stackrel{\text{indep.}}{=} M_{n-1}E(X_n) = M_{n-1}$
question : $\lim M_n$ exists when ? [answer: MCT : +ve rvs so $M_\infty := \lim M_n$ exists as]
question : $\lim E(M_\infty) = 1$ when ? [answer: Kakutani theorem on product martingale]
 3. Accumulating data : $\xi \in L^1 \rightarrow M_n := E(\xi|\mathcal{F}_n)$ is martingale **Proof:** $E(M_n|\mathcal{F}_{n-1}) = E(\xi|\mathcal{F}_n|\mathcal{F}_{n-1}) = E(\xi|\mathcal{F}_{n-1}) = M_{n-1}$

- question : $\lim M_n$? [answer: Levy upward theorem]
 $M_n \rightarrow M_\infty := E(\xi|\mathcal{F}_\infty)$
question : $\xi = E(\xi|\mathcal{F}_\infty)$ when ? [answer : $\sum \frac{1}{c_k^2}$]
- game: $X_n - X_{n-1}, n \geq 1$ = net win per unit stake i game n.
 - fair game: $E(X_n - X_{n-1}|\mathcal{F}_{n-1}) = 0$
 - previsible process: $C_n, n \geq 1$ is \mathcal{F}_{n-1} -measurable. C_0 does not exist.
 - C_n stake on game n.
 - Winnings in game n = $C_n(X_n - X_{n-1})$
 - total Winnings up to game n := $Y_n := \sum_{1 \leq k \leq n} C_k(X_k - X_{k-1}) := (C \bullet X)_n$ [note: similar to $\int C dX$]
 - cannot beat the system
 1. if X martingale, C bounded non-negative previsible ie $C_n(\omega) \geq 0, |C_n(\omega)| \leq K$ then $(C \bullet X)$ martingale null at 0
 2. if X martingale, C bounded previsible then $(C \bullet X)$ martingale null at 0
 3. same if bounded $\Leftrightarrow L^2$ and X in L^2
 4. **Proof:** $E(Y_n - Y_{n-1}|\mathcal{F}_{n-1}) = C_n E(X_n - X_{n-1}|\mathcal{F}_{n-1})$
 - stopping time (A): T_n stopping time if $\{T \leq n\} = \{\omega : T(\omega) \leq n\} \in \mathcal{F}_n, \forall n \in [0, \infty]$ note ∞ is included
 - stopping time (B): same equivalent condition as $\{T = n\} = \{\omega : T(\omega) = n\} \in \mathcal{F}_n, \forall n \in [0, \infty]$ **Proof:** if (A) then $\{T = n\} = \{T \leq n\} \setminus \{T \leq n-1\} \in \mathcal{F}_n$ if (B) $\{T \leq n\} \cup \bigcup_{0 \leq k \leq n-1} \{T = k\} \in \mathcal{F}_n$ since $\mathcal{F}_k \subseteq \mathcal{F}_n, k \leq n$
 - example of stopping time: time of first entry of A := adapted process $\{A_n\}$ into B : $T = \inf\{n \geq 0 : A_n \in B\}$
 - example of NOT a stopping time: $T = \sup\{n \leq 10 : A_n \in B\}$
 - stopped process :
 1. let X martingale
 2. let T stopping time,
 3. let stake process C^T be 0 after T ie $C_n^T = I_{n \leq T} = \begin{cases} 1 & \text{if } n \leq T(\omega) \\ 0 & \text{after} \end{cases}$
 4. wins process at time $n : (C^T \bullet X)_n = X_{T \wedge n} - X_0$ ie win is frozen after a while and dstays like that forever
 5. let X^T = process stopped (frozen) at $T : X_n^T(\omega) := X_{T(\omega) \wedge n}(\omega) \Leftarrow \omega$ determines when to stop ad also the value of X at that stopping point
 6. then obviously $C^T \bullet X = X^T - X_0$ (wins are frozen after a while)
 7. $C^T = 0, 1 \leq n \leq T, \geq 0$ so bounded and non-negative/previsible because set to 0 once T is reached ie $\{C_n(T) = 0\} = \{T \leq n-1\} \in \mathcal{F}_{n-1}$
 - X martingale, T stopping time $\Rightarrow X^T$ martingale
 - in particular $E(X_{T \wedge n}) = E(X_0)$
 - VERY IMPORTANT : Sometimes $E(X_{T \wedge n}) \neq E(X_T)$

- example : random walk on \mathbb{Z}^+ starting at 0, first time hitting 1. Can prove $P(T < \infty) = 1$. Then (as per above) $E(X_{T \wedge n}) = E(X_0) = 0$ but $E(X_T) = 1$ so $E(X_{T \wedge n}) \neq E(X_T)$
- Doobs optional stopping theorem : gives conditions for $E(X_T) = E(X_0)$
- Doobs optional stopping theorem

- let T stopping time, X martingale
- then IF one of the following holds, $E(X_T) = E(X_0)$ [not just $E(X_{T \wedge n})$] :
 - T is certainly bounded ie $\exists N : T(\omega) \leq N$ **Proof:** set $n=N$ in above
 - X is certainly bounded and T a.s finite ie $\exists K : |X_n(\omega)| \leq K, \forall n, \forall \omega$ **Proof:** BDD so can use $n \rightarrow \infty$
 - $E(T)$ bounded aka integrable and X increment bounded ie $E(T) < \infty$ and $\exists K : |X_n(\omega) - X_{n-1}(\omega)| \leq K, \forall n, \forall \omega$ **Proof:** $|X_{T \wedge n} - X_0| = |\sum_{k=1}^{T \wedge n} (X_k - X_{k-1})| \leq KT$ and so can use DOM

- COROLLARY : Doobs optional stopping theorem

- let M martingale with bounded increment K_1
- let C previsible with bound K_2
- let T stopping time with $E(T) < \infty$

then

- $E(C \bullet M)_T = 0$ **Proof** $E(C \bullet M)_T = E(C \bullet M)_0 := 0$

- can use this for ABRACADABRA problem and show $E(T) = 26^{11} + 26^4 + 26$
- When is $E(T) < \infty$? : let stopping time T such that for some N and some $\epsilon : P(T \leq n + N | \mathcal{F}_n) > \epsilon \Rightarrow E(T) < \infty$
- hitting time for random walk

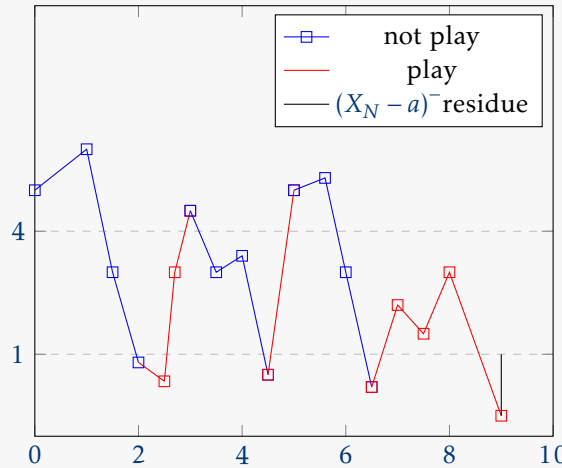
- let $P(X = +1) = P(X = -1) = \frac{1}{2}$
- let $S_0 := 0, S_n := X_1 + \dots + X_n$
- let $T = \inf\{n : S_n = 1\}$
- let $\mathcal{F}_n = \sigma(X_1, \dots, X_n) = \sigma(S_0, S_1, \dots, S_n)$
- compute $E(e^{\theta X}) = \frac{1}{2}(e^{\theta} + e^{-\theta}) = \cosh \theta$
- so $E[(\text{sech } \theta)e^{\theta X_n}] = 1, \forall n$
- so (see prod of +ve martingales) $M_n^{\theta} = (\text{sech } \theta)^n e^{\theta S_n}$ is martingale
- T stopping time so $EM_{T \wedge n}^{\theta} = 1 = E[(\text{sech } \theta)^{T \wedge n} e^{\theta S_{T \wedge n}}]$
- now let $\theta > 0$
- $S_{T \wedge n} \leq 1$ so $e^{\theta S_{T \wedge n}} \leq e^{\theta}$, so $M_{T \wedge n}^{\theta} \leq e^{\theta}$ ie bounded
- note $M_{T \wedge n}^{\theta} \rightarrow M_T^{\theta}$ as $n \rightarrow \infty$

- so can use BDD to let $n \rightarrow \infty$ in $EM_{T \wedge n}^{\theta} = 1$ to get $EM_T^{\theta} = 1 = E[(\text{sech } \theta)^T e^{\theta}]$
- so $E[(\text{sech } \theta)^T] = e^{-\theta}$
- so let $\theta \downarrow 0$ then because $\text{sech } \theta < 1$ we have

$$\begin{cases} (\text{sech } \theta)^T \uparrow 1, \text{ when } T < \infty \\ (\text{sech } \theta)^T \uparrow 0, \text{ when } T = \infty \end{cases}$$
- both bounded so either MON or BDD gives $EI_{T < \infty} = 1 = P(T < \infty)$
- let $\alpha = \text{sech } \theta$ then from above $E[(\text{sech } \theta)^T] = E(\alpha^T) = e^{-\theta}$
- then $E(\alpha^T) = \sum \alpha^n P(T = n) = e^{-\theta} = \frac{1 - \sqrt{1 - \alpha^2}}{\alpha}$
- so find $P(T = k)$ by expansion

12 Chapter 11 - Martingales III : Martingale Convergence theorem

play when below a=1, stop when above b=4



- C previsible strategy: play when below a=1, stop when above b=4
- $C_1 := I_{X_0 < a}, C_n := I_{C_{n-1}=1} I_{X_{n-1} \leq b} + I_{C_{n-1}=0} I_{X_{n-1} < a} \nleftrightarrow$ ie if previous = 1, then stay 1 while $\leq b$, if previous = 0 then set to 1 when $< a$
- let $Y = C \bullet X$ = gains process with that game. Y follows X when played, constant otherwise. $Y_0(\omega) = 0$
- number of upcrossings $U_N[a, b]$ = largest k in partition that catches upcrossings = k in finest partition with $0 \leq s_1 < t_1 \dots s_{N-1} < t_{N-1} \leq N$ that catches all the upcrossings ie $X(s_i)(\omega) < a, X(t_i)(\omega) > b$
- $Y_N(\omega) \geq (b - a)U_N[a, b](\omega) - |X_N(\omega) - a| \nleftrightarrow$ upcrossings - last bit. Note -|first bit| = ignored
- Doob's upcrossing lemma : $(b - a)EU_N[a, b] \leq E[(X_N - a)^-]$ **Proof** C previsible (see formula above), bounded (it is 0,1), +ve (same), and $Y = C \bullet X$, so Y supermartingale with $Y_0 = 0$ so $E(Y_N) \leq 0$ then use above
- corollary : let X supermartingale bounded in L^1 meaning $\sup_n E(|X_n|) < \infty$, and let $U_{\infty} \uparrow U_N[a, b]$, then

- $\Rightarrow (b - a)EU_{\infty} \leq |a| + \sup_n E(|X_n|) < \infty$ **Proof:** see above $(b - a)EU_N[a, b] \leq |a| + E(|X_N|) \leq |a| + \sup_n E(|X_n|)$ then use MON since all the functions positive.

- so $P(U_{\infty}[a, b] = \infty) = 0$ **Proof** $EU_{\infty}[a, b]$ finite and U +ve and see sums of +ve rvs

- Doob Martingale Convergence theorem : X supermartingale bounded in L^1 ie $\sup_n E(|X_n|) < \infty \Rightarrow X_{\infty} := \lim X_n := \limsup_n X_n$ exists, and X_{∞} is \mathcal{F}_{∞} -measurable **Proof** $A := \{\omega : X_n(\omega) \text{ does not converge to a limit in } [-\infty, +\infty]\} = \{\omega : \liminf X_n(\omega) < \limsup X_n(\omega)\} = \bigcup_{a, b \in \mathbb{Q}, a < b} \{\omega : \liminf X_n(\omega) < a < b < \limsup X_n(\omega)\} = \bigcup A_{a, b}$ but $A_{a, b} \subseteq \{\omega : U_{\infty}[a, b](\omega) = \infty\}$ so (above) $P(A_{a, b}) = 0$ so (countable \bigcup) $P(A) = 0$ so $X_{\infty} := \lim X_n$ exists in $[-\infty, +\infty]$ a.s. But actually it is finite a.s because (FATOU) $E|X_{\infty}| = E(\liminf |X_n|) \leq \liminf E|X_n| \leq \sup_n E|X_n| < \infty$ so $P(X_{\infty}) = 1$

- WARNING** : yes $X_{\infty} := \lim X_n$ exists but sometimes $E(X_{\infty}) \neq E(X_n)$ - see branching process example

- Corollary if X non-negative supermartingale then $X_{\infty} := \lim X_n$ exists a.s **Proof** bounded in L^1 since $E|X_n| = E(X_n) \leq E(X_0)$

13 Chapter 12 - Martingales III : Martingale Bounded in L2

- Three series theorem, strong law of large numbers, Levy's extension of BC
- bounded in L2 (ie $\sup_n \|M_n\|_2 < \infty$) \Rightarrow bounded in L1 (monotonicity of norms) and L2 easier because of formula $E(M_n^2) = E(M_0^2) + \sum_{k=1}^n E[(M_k - M_{k-1})^2]$ (proof below)
- $M = (M_n : n \geq 0)$ martingale in L^2 ie $\forall n, E(M_n^2) < \infty$ so since martingale $E(M_v | \mathcal{F}_u) = M_u$ then $M_v - M_u \perp L^2(\mathcal{F}_u)$ so $M_n = M_0 + \sum_{k=1}^n (M_k - M_{k-1})$ sum of orthogonal terms in L2 norm ie $E(M_n^2) = E(M_0^2) + \sum_{k=1}^n E[(M_k - M_{k-1})^2]$
- M martingale in L^2 ie $M_n \in L^2$ then (a) M bounded $\Leftrightarrow \sum E[(M_k - M_{k-1})^2] < \infty$ and then (b) $M_n \xrightarrow{L^2} M_{\infty}$ **Proof** (a) $E(M_n^2) = E(M_0^2) + \sum_{k=1}^n E[(M_k - M_{k-1})^2]$ so bounded one side \Leftrightarrow bounded the other . (b) M bounded in $L^2 \xRightarrow{\text{monotonicity}} \text{Doob convergence}$

bounded in L^1 so $\Rightarrow \exists M_{\infty} := \lim M_n$. Now pythagoras $E(M_{n+r} - M_n)^2 = \sum_{k=n+1}^{n+r} E[(M_k - M_{k-1})^2]$, let $r \uparrow \infty$ and $E(M_{\infty} - M_n)^2 \stackrel{\text{FATOU}}{\leq} \sum_{k \geq n+1} E[(M_k - M_{k-1})^2]$ so $\lim_n E[(M_{\infty} - M_n)^2] = 0$ so $M_n \xrightarrow{L^2} M_{\infty}$

- Sum of zero-mean RV with converging sum of Vars converges
- if X_k such that $E(X_k) = 0, \sigma_k^2 := \text{Var}(x_k) < \infty$ then

- $\sum \sigma_k^2 < \infty \Rightarrow \sum X_k$ converges a.s

Partial converse : if $\exists K \therefore |X_k(\omega)| \leq K$ then $\sum X_k$ converges a.s $\Rightarrow \sum \sigma_k^2 < \infty$

2. Proof (1) let $M_n = \sum_{k=1}^n X_k$. Now sum of IID RV \Rightarrow martingale as we know already and $E(M_k - M_{k-1})^2 = E(X_k^2) = \sigma_k^2$. Now since $E(M_n^2) = E(M_0^2) + \sum_{k=1}^n E(M_k - M_{k-1})^2$ then $E(M_n^2) = \sum \sigma_k^2 := A_n$, and since $\sum \sigma_k^2 < \infty$ then M bounded in L^2 , so $\lim M_n$ exists as,

• **Proof (2)** $E[(M_k - M_{k-1})^2 | \mathcal{F}_{k-1}] = E(X_k^2 | \mathcal{F}_{k-1})$ X_k indep of \mathcal{F}_{k-1}
 $E(X_k^2) = \sigma_k^2$. Separately M_{k-1} is \mathcal{F}_{k-1} -measurable so $\sigma_k^2 = E(M_k^2 | \mathcal{F}_{k-1}) - 2M_{k-1}E(M_k | \mathcal{F}_{k-1}) + M_{k-1}^2 = E(M_k^2 | \mathcal{F}_{k-1}) - M_{k-1}^2$ which shows that $N_n := M_n^2 - \sum_{k=1}^n \sigma_k^2$ is martingale. let $T := \inf(r : |M_r| > c)$ then N^T martingale so $EN_n^T = E(M_n^T)^2 - EA_{T \wedge n} = EN_0 = 0$ with $A_n = \sum_{k=1}^n \sigma_k^2$, but $|M_T - M_{T-1}| = |X_T| \leq K$ if T finite, so $|M_n^T| \leq K + c$ so $(\star)EA_{T \wedge n} \leq (K + c)^2$. Now $\sum X_k$ converges \Rightarrow partial sums bounded so for c big enough so $P(T = \infty) > 0$ so from $(\star) \Rightarrow A_\infty := \sum \sigma_k^2 < \infty$

• random signs: (a_n) seq. of real numbers and $\epsilon_n : P(\epsilon_n = \pm 1) = \frac{1}{2}$ then from above:

1. $\sum \epsilon_n a_n$ converges $\Leftrightarrow \sum a_n^2 < \infty$
2. $\sum \epsilon_n a_n$ oscillates $\Leftrightarrow \sum a_n^2 = \infty$

• can prove above without need for zero-mean variable by symmetrising ie can prove: $X_n(\omega)$ bounded by K then $\sum X_k$ converges $\Rightarrow \sum E(X_k)$ converges and $\sum \text{Var}(X_n) < \infty \Leftrightarrow$ no requirement for $E(X_k) = 0$ **Proof sketch:** copy sample space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*) = (\Omega, \mathcal{F}, \mathbb{P}) \times (\Omega^@, \mathcal{F}^@, \mathbb{P}^@)$ and symmetrise with let $\omega^* = (\omega, \omega^@)$, $X_k^*(\omega^*) := X_n(\omega)$, $X_k^{*@}(\omega^*) := X_k^@(\omega^@)$, $Z_k^*(\omega^*) := X_k^*(\omega^*) - X_k^{*@}(\omega^*)$ which now has 0-mean and is bounded because X is.

• **KOLMOGOROV 3-SERIES** : X_n iid sequence then $\sum_k X_k$ converges

1. $\sum_k X_k$ converges

• \Leftrightarrow all 3 hold for some K :

1. $\sum_n P(|X_n| > K) < \infty$
2. $\sum_n E(X_n^K)$ converges
3. $\sum_n \text{Var}(X_n^K) < \infty$
4. where $X_n^K(\omega) := \begin{cases} X_n(\omega) & \text{if } |X_n(\omega)| \leq K \\ 0 & \text{otherwise} \end{cases}$

• **Proof \Leftarrow (if)** So for some K the 3 properties hold. $\sum P(X_n \neq X_n^K) = P(|X_n| > K) < \infty$ so (BC1 – see aside) $P(X_n = X_n^K)$ for all but finitely many n - ie eventually=1

• Recall BC1 : $\sum P(E_n) < \infty \Rightarrow P(\limsup E_n) = P(E_n, \text{i.o.}) = 0$

• Aside: $(X_n \neq X_n^K, \text{i.o.}) = \{\omega : X_n(\omega) \neq X_n^K(\omega), \text{i.o.}\}$
ie $= \{\omega : \forall n \in \mathbb{N}, \exists m > n \text{ with } X_m(\omega) \neq X_n^K(\omega)\}.$

$\{\omega : \forall n \in \mathbb{N}, \exists m > n \text{ with } X_m(\omega) \neq X_n^K(\omega)\}$. The negation of $\forall n \in \mathbb{N}, \exists m > n$ with $X_m(\omega) \neq X_n^K(\omega)$ is $\exists n \in \mathbb{N}$ such that $\forall m > n X_m(\omega) = X_n^K$. SO $\{\omega : X_n(\omega) \neq X_n^K(\omega), \text{i.o.}\}^c = \{\omega : X_n(\omega) = X_n^K(\omega), \text{eventually}\}$

• Aside alternative proof : $0 = P(\limsup\{X_n \neq X_n^K\}) = P(\cap_m \cup_{n>m} \{X_n \neq X_n^K\}) = 1 - P(\cup_m \cap_{n>m} \{X_n = X_n^K\}) \Rightarrow 1 = P(\liminf\{X_n = X_n^K\})$.

• Anyway so $P(X_n = X_n^K)$ for all but finitely many n - ie eventually=1. So we only need to show that $\sum X_n^K$ converges a.s, and actually (property 2), only need to show that $\sum Y_n^K$ converges where $Y_n^K := X_n^K - E(X_n^K)$. Y is zero mean, finite var, so use (3) and earlier above.

• **Proof \Rightarrow (only if)** So suppose $\sum X_n$ converges a.s. Then $X_n \rightarrow 0$ a.s. Let any $K > 0$ then $|X_n| > K$ for finitely many n , so (BC2) means prop (1) holds. [Recall BC2 : E_n independent, $\sum P(E_n) = \infty \Rightarrow P(\limsup E_n) = 1$]. Since $X_n = X_n^K$ for all but finitely many n , $\sum X_n^K$ converges a.s. then use 'symmetrising' above.

- Cesaro
- Kronecker
- Strong law with variance constraint
- Kolmogorov truncation lemma
- Kolmogorov SLLN

14 Chapter 13 - Martingales IV : Uniform integrability

15 Chapter 14 - Martingales V : UI Martingales

16 Chapter 15 - Martingales VI : Examples

17 Chapter 16 - Characteristic functions I : Basics

18 Chapter 17 - Characteristic functions II : Weak convergence

19 Chapter 18 - Characteristic functions III : Central Limit theorem