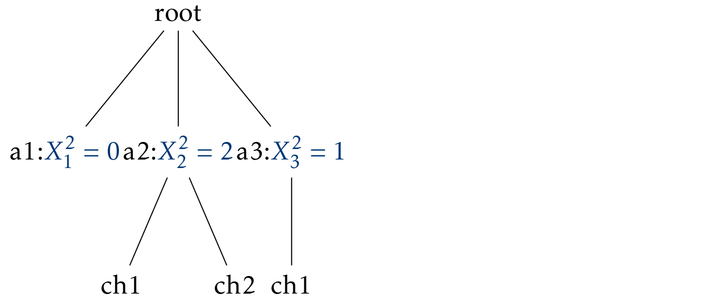
1 Chapter 0 - branching example

$\mathbf{Z}+ = [0,1..]$ $\quad \mathbf{N} = [1..]$
$f(\theta) = E(\theta^X) = \sum \theta^k P(X=k) = P(X-0) + \sum_{k=1}\theta^k P(X=k)$
$f'(\theta) = E(X\theta^{X-1}) = \sum k\theta^{k-1}P(X=k) \leftarrow$ differentiate wrt θ
mean $= \mu = f'(1) = \sum_k kP(X=k)$ $\quad f(1) = \sum P(X=k) = 1$

$\{X_r^m\}$ = double series of random variables IID
X_m^r = the children of animal m contributed to generation r
$Z_{r+1} = X_1^{r+1} + ... + X_{Z_r}^{r+1}$ = 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 :
$Z_0 = 1 \leftarrow$ Z0 = generation 0 has 1 element
and $Z_1 = 3 == X^1_{Z_0=1} = 3 \leftarrow$ Z1 = generation 1 has 3 children
and $Z_2 = 3 = X_1^2 + X_2^2 + X^2_{Z_1=3} = 0+2+1 \leftarrow Z_2$ = num children in generation 2 = 3 children

1.2 calculation

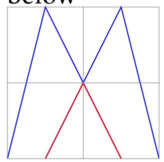
- Distribution of $\boxed{Z_n}$ obtained from generating $\Rightarrow f_n(\theta) = E(\theta^{Z_n}) = \sum \theta^k \boxed{P(Z_n = k)}$
- $f_{n+1}(\theta) = E\theta^{Z_{n+1}} = E\left(\boxed{E\theta^{Z_{n+1}}|Z_n}\right) = \sum \boxed{E\left(\theta^{Z_{n+1}}|Z_n\right)} P(Z_n = k) \leftarrow \boxed{E\theta^{Z_{n+1}}|Z_n}$ is the random variable here
- so now standard calc : $E\left(\theta^{Z_{n+1}}|Z_n\right) = E\left(\theta^{Z_{n+1}}|Z_n = k\right) = E\left(\theta^{X_1^{n+1}+..X^{n+1}_{Z_n=k}}|Z_n = k\right)$
- but (a) num children contributed by animal m to generation n+1 = x_m^{n+1} = not dependent on other siblings $x^{n+1}_{...}$ in its own generation (n), or indeed num of such siblings = Z_n , so complicated expression $E\left(\theta^{X_1^{n+1}+..X^{n+1}_{Z_n=k}}|Z_n = k\right)$ = same without the conditioning = absolute expectation = $E\left(\theta^{X_1^{n+1}+..X^{n+1}_{Z_n=k}}\right)$ and moreover (b) the $X^{n+1}_{.}$ are independent and independent = multiply
- so $E\left(\theta^{X_1^{n+1}+..X^{n+1}_{Z_n=k}}\right) = E\theta^{X_1^{n+1}}...E\theta^{X^{n+1}_{Z_n=k}}$ and all the X_i^j have same distribution and $E\theta^X = f(0)$ 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 $E(...) = E\theta^{Z_{n+1}} = Ef(\theta)^{Z_n} = f_n(\theta)$ by definition see first line
- so $E\theta^{Z_{n+1}} = f_{n+1}(\theta) = Ef(\theta)^{Z_n} = f_n(f(\theta)) = f_n \circ f(\theta)$
- so $\boxed{f_{n+1}(\theta) = f_n(\theta)}$

1.3 extinction

- extinction = zero children in some generation = $\pi_n = P(Zn = 0) = fn(0)$
- $\pi_{n+!} = P(Z_{n+1} = 0) = f_{n+1}(0) \leftarrow$ 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 $\boxed{\pi = f(\pi)}$ with $f(1) = 1$ $\quad f(0) = P(X = 0)$ slope at 1 $= f'(1) = E(X) = \mu \longleftarrow \boxed{f'(\theta) = E(X\theta^{X-1})}$ so $f'(1) = E(X)$
- if $\mu > 1$ then extinction with prob $\pi = f(\pi)$ before hits 1 but if 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 $\theta = 1 \rightarrowtail E(Z_{n+1}|Z_n) = Z_n 1^{whatever} f'(1) = \mu Z_n$
- 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 Zn
- $E(M_{n+1}|Z_n) = E(\frac{Z_{n+1}}{\mu^{n+1}}|Z_n) = \frac{1}{\mu^{n+1}}E(Z_{n+1}|Z_n) = \frac{1}{\mu^{n+1}}\mu Z_n = \frac{z_n}{\mu^n} = M_n$
- $E(M_{n+1}|Z_n) = M_n \leftarrow E(M_{n+1}|Z_n)$ **M**n martingale relative to **Z**n
- $E(M_{n+1}) = E(M_n) = ... = E(M_0) = Z_0 = 1$
- $\boxed{MCT}$ because Mn >= 0 $\rightarrowtail$ EXISTS ALMOST SURELY $M_\infty = \lim M_n = \lim \frac{Z_n}{\mu^n}$
- CAUTION $E(M_n) = 1$ $\quad \lim M_n = M_\infty$ exists but if $\mu <= 1 \rightarrow= M_\infty = 0$ so sometimes $E(M_\infty = \lim M_n) = 0 \neq E(M_n) = 1$ ie $E(lim) <> \lim E(.)$
- $\boxed{\text{HOWEVER FATOU IS TRUE}}$ $E(\liminf Y_n) <= \liminf E(Y_n)$ ie $E(\liminf M_n) = E(M_\infty) = E(0) = 0 <= \liminf E(M_n) = \liminf 1 = 1$
- see reminder picture about FATOU E (lim inf) < lim inf E (.) below



1.5 distribution of M_∞

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

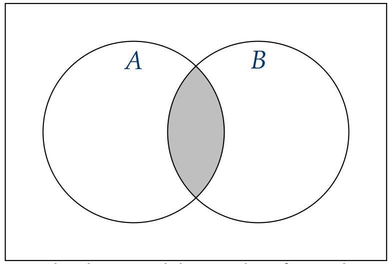
1.6 Now try a particular P(X=k) : $P(X = k) = pq^k$

- $P(X = k) = pq^k \Rightarrow f(\theta) = E\left(\theta^X\right) = \sum_{k=0}\theta^k P(X = k) = \sum_{k=0}\theta^k pq^k = \frac{p}{1-q\theta}$; and also $\mu = \frac{q}{p}$
- $\pi = f(\pi) \Rightarrow \pi = \frac{p}{1-q\pi} \Rightarrow (\pi = 1, \pi = \frac{p}{q} \iff \frac{p}{q} < 1)$
- Need fn(theta) $f(\theta) = \frac{p}{1-q\theta} \Rightarrow G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & p \\ 1 & -q \end{pmatrix} \Rightarrow G^n = (SAS^{-1})^n \Rightarrow f_n(\theta) = \frac{p\mu^n(1-\theta)+q\theta-p}{q\mu^n(1-\theta)+q\theta-p}$
- $\mu <= 1 \Rightarrow \pi = 1$ (see earlier) and $\Rightarrow \lim f_n = 1 \Rightarrow$ all in P(x=0) $\Rightarrow$ process dies out
- $\mu > 1 \Rightarrow L(\lambda) = E(exp(-\lambda M_\infty)) \overset{bct}{=} \lim E(exp(-\lambda M_n)) = \lim f_n\left(e^{\frac{-\lambda}{\mu^n}}\right)$ with fn as above it is found that $\Rightarrow L(\lambda) = \frac{p\lambda+q-p}{q\lambda+q-p} \overset{Laplace}{\Rightarrow} L(\lambda) = \pi e^{-\lambda.0} + \int_0^\infty (1-\pi)^2 e^{-\lambda x} e^{-(1-\pi)x} dx \Rightarrow \begin{cases} P(M_\infty = 0) = \pi \\ P(x < M_\infty < x+dx) = (1-\pi)^2 e^{-(1-\pi)x} \Rightarrow P(M_\infty > x) = (1-\pi) \end{cases}$ [cut off]
- $\mu < 1 \Rightarrow$ process dies out (see above) but what is the distribution when it does not ie what is $E\left(\theta^{Z_n}|Z_n \neq 0\right)$? It is $E\left(\theta^{Z_n}|Z_n \neq 0\right) = \frac{f_n(\theta)-f_n(0)}{1-f_n(0)} \Leftarrow$ 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 $E(Z_n|Z_n \neq 0) = \sum kP(Z_n = k|Z_n \neq 0) = \sum k(1-\mu)\mu^{k-1} = 1/(1-\mu)$ and (see above the intersection/division) $P(Z_n) \neq 0 = 1 - f_n(0) = (1-\mu)\mu^n$ so

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

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

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



- $\sigma-$algebra stable under finitely many interesections / unions
- $\cap A_n = (\cup A^c)^c$
- $Borel(S) = B(S) = \sigma(\text{open sets in S})$
- $Borel(\mathbb{R}) = B(\mathbb{R}) = \sigma(\pi(R))$ 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 $(a,b) = \cup_n(a, b-\epsilon.1/n]$ 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 $\Sigma -> [0, \infty]$
- 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:$A, B \in I \to A \cap B \in I$
- $m1(S) = m2(S) < \infty$ and $m1 = m2$ on I $\Rightarrow m1 = m2$ on Σ
- **Measure Extension theorem (Corollary)** : 2 measures agree on pi-system -> agree on sigma algebra - and recall the case $B = B(R) = \sigma(\pi(R))$
- **Measure Extension theorem (Caratheodory)** : countably additive function m0 on an algebra on set S can be extended to measure m on sigma (algebra) such that m0 = m on the algebra, and uniquely so if the measure is finite ie. m0 (S) < ∞
- **Application: Lebesgue measure** collection $F = \cup[a_k, b_k]$ of sets on [0,1] with $m0(F) = \sum(b_k - a_k)$ and $0 \le a_1 \le b_1 \le ...bk \le 1$ then m0 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 Fn of disjoint elements with $F = \cup F_n$ and take $G_n = \cup_{k=1}^{k=n} F_n$ then obviously $G_n \uparrow F$ and $m0(G_n) = \sum_1^n F_k$
- then just show that $m0(Gn) \uparrow m0(F)$ because then $m0(F) :=\uparrow \lim m0(Gn) =\uparrow \lim m0(\cup_{k=1}^{k=n} F_n) := m0(\cup_{k=1}^{k=\infty} F_n)$
- take $H_n = F \setminus G_n$ then $H_n \downarrow \emptyset$ so just show that $m0(H_n) \downarrow 0$
- that is same as showing if $Hn \downarrow$ and $m0(H_n) > \epsilon \to \cap H_n \neq \emptyset$
- to do that , take $\overline{J_k} \subseteq H_k$ and $m0(H_k \setminus J_k) \le \epsilon 2^{-k}$ then $m0(H_n \setminus \cap_{k\le n} J_k) \le m0(\cup_{k \le n} H_n \setminus J_k) \le \epsilon \sum 2^{-k} \le \epsilon$ so $mo(\cap_{k\le n} J_k) \ge \epsilon$
- so $\cap_{k \le n} J_k \neq \emptyset$ so $K_n := \cap_{k\le n}\overline{J_k} \neq \emptyset$ so $\cap_{k \le n} H_k \neq \emptyset$
- finally it follows $\cap \overline{J_k} \neq \emptyset$ because $\overline{J_k}$ is compact so choose a point inside K_n and find a subsequence that $\to$ inside the compact
- end of proof
- inequalities

$\mu(A \cup B) \le \mu(A) + \mu(B)$
$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) \Leftarrow A \cup B = A \cup (B \setminus (A \cap B))$
$\mu(\cup_n U_i) \le \sum_n \mu(U_i)$
$\mu(\cup_n U_i) = \sum_n \mu(U_i) - \sum\sum_{i<j} \mu(U_i \cap U_j) + \sum\sum\sum_{i<j<k} \mu(U_i \cap U_j \cap U_k) - ...$

- **monotone convergence of measures - UP**

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

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

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

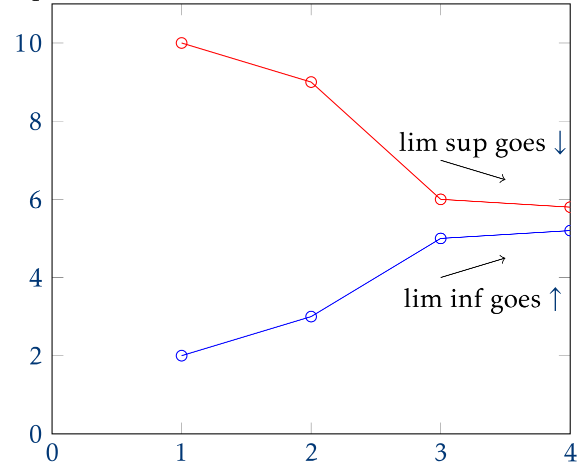
- countable sum of null sets = 0
- WARNING $H_n = [n, \infty] \to H_n \downarrow \emptyset \quad \mu(H_n) = \infty$
- WARNING : EXAMPLE of $\cap \cup I_{n,k} \neq \cup \cap I_{n,k}$
- $V = Q \cap [0,1] = [v_n]$ all rationals in [0,1] $\subseteq G_k = U_n[v_n - \epsilon_k/2^n, v_n + \epsilon_k/2^n] = \cup I_{n,k}$ 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
- Example coin toss F: = $\left[\omega : \frac{\#(k \le n \quad w_k = H)}{n} \to \frac{1}{2}\right]$
- almost surely : $F = [\omega : S(\omega) = true]$ in $\mathcal{F}$ and has P(F) =1
- $P(F_n) = 1 \to P(\cap F_n) = 1 \Leftarrow P(F_n^c) = 0 \Rightarrow P(\cup F_n^c) = 0$ with $\left[\cup F_n^c\right]^c = \cap F_n$
- EXAMPLE / WARNING: $F_\alpha = \left[\omega : \frac{\#(k \le n \quad w_{\alpha(k)} = H)}{n} \to \frac{1}{2}\right] \Rightarrow$ $P(F_\alpha) = 1$ for all $\alpha(k)$ sequences that go from 0...1 however $\bigcap_{\alpha \text{ sequences}} F_\alpha = \emptyset \Longleftarrow \forall \alpha \quad \exists \omega ...$
- LIM SUP , LIM INF

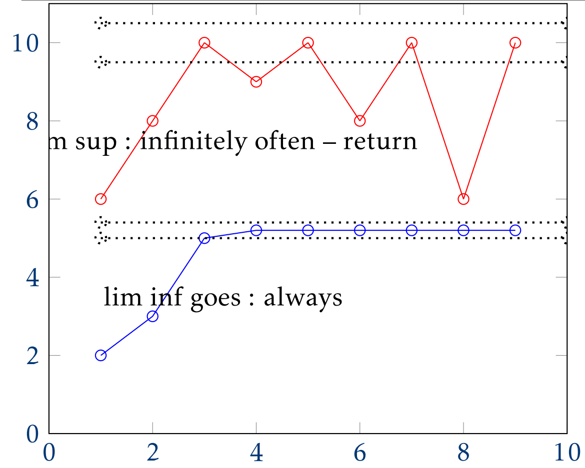
$\boxed{\limsup x_n := \inf_m \left(\sup_{n \ge m} x_n\right) =\downarrow \left(\sup_{n\ge m} x_n\right)} \leftarrow$ eventual inf of all the sups

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



- LIMIT $\limsup = \liminf \Longleftarrow$ limits exists
- LIM SUP , LIM INF - for SETS
-
- LIM SUP
  1. E_n infinitely often
  2. $:= \limsup E_n = \inf_m \left(\sup_{n\ge m} E_n\right)$
  3. $=\downarrow \left(\sup_{n \ge m} E_n\right)$
  4. $= \cap_m \cup_{n \ge m} E_n$
  5. $= \forall m, \exists n(\omega) \ge m$ such that $\omega \in E_{n(\omega)}$
- for each .., there exists ... $\to$ there is always one ... $\to$ infinitely often
- for each = $\cap$, there exists = $\cup$
- LIM INF

E_n eventually :=
$\liminf E_n = \sup_m \left(\inf_{n \ge m} E_n\right)$
$=\uparrow \left(\inf_{n\ge m} E_n\right)$
$= \cup_m \cap_{n \ge m} E_n$
= all the $m(\omega)$ that exist, such that every $n(\omega) \ge m(\omega) : \omega \in E_{n(\omega)}$

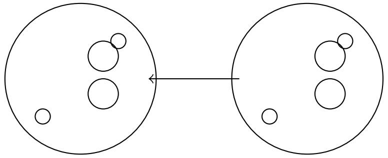


- FATOU 1 $\boxed{P(\limsup E_n) > \limsup P(E_n)}$
- proof: $G_m = \cup_{n \geq m} E_n \rightarrow G_m \downarrow \limsup E_n \rightarrow meas(G_m) \downarrow meas(\limsup E_n) \rightarrow P(G_m) \downarrow P(\limsup E_n)$
- and also $P(G_m) = P(\cup_{n\geq m} E_n) \rightarrow P(G_m) >= sup_{n>=m} P(E_n) \rightarrow P(G_m) \geq\downarrow \lim_m sup_{n>=m} P(E_n) := \limsup P(E_n)$
- so $\rightarrow P(\limsup E_n) \geq \limsup P(E_n)$
- example: (x1,x2) sequence = (0,1),(1,0),(0,1),(1,0),.... then sum x1+x2 = 1 -> limsup sum =1 , but (limsup x1,limsup x2) = (1,1) -> sum(limsup) = 1+1 = 2 $\rightarrowtail \limsup(sum) = 1 <= sum(\limsup) = 2 \rightarrowtail \limsup \int <= \int \limsup$
- example:
- $\boxed{meas(\limsup E_n) > \limsup meas(E_n)}$
- meas(max) > max (meas)

- BOREL-CANTELLI 1 $\boxed{\sum P(E_n) < \infty \rightarrow P(\limsup E_n) = 0}$ $\leftarrow$ limsup En is a set , and has a measure/prob and here prob/measure = 0. basically if the prob ifs finite -> things that happen i.o have prob 0 (logical!)
- proof: $P(G) <= P(G_m = \cup_{n\geq m} E_n) <= \sum_{n>=m} P(E_n)$
- FATOU 2 $\boxed{P(\liminf E_n) < \liminf P(E_n)}$
- example: (x1,x2) sequence = (0,1),(1,0),(0,1),(1,0),.... then sum x1+x2 = 1 -> liminf sum =1 , but (liminf x1,liminf x2) = (0,0) -> sum(liminf) = 0+0 = 0 $\rightarrowtail \liminf(sum) = 1 >= sum(\liminf) = 0 \rightarrowtail \liminf \int >= \int \liminf$

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

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 $\leftrightarrow h^{-1}(A) \in \Sigma \quad \forall A \in$ Borel
- $h^{-1}(\cup A_\alpha) = \cup_\alpha h^{-1}(A_\alpha) \quad h^{-1}(A^c) = \left[h^{-1}(A)\right]^c$
- $C \in B$ 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 $C \in B$
- $\{h <= s\} \in \Sigma \Rightarrow h$ 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 $\{h1 + h2 > c\} = \cup_{q\in Q}\{h1 > q\} \cap \{h2 > c - q\}$ countable collection of Σ
- $\boxed{h_n \text{ measurable} \Rightarrow \inf h_n, \liminf h_n, \limsup h_n \text{measurable}}$ Proof: $\{\inf h_n > c\} = \cap_n \{h_n > c\}$, define $L = \liminf h_n \quad L_n = \inf h_n$ then $\{L < c\} = \cap_n \{L_n < c\}$
- $\boxed{\{s : \lim h_n(s) \in \mathbb{R}\} \in \Sigma}$ 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 Sn/n , L- = liminf Sn/n where Sn = num Hs . n= num tosses
- **sigma-algebra on Ω generated by a sequence of functions** $f_n : \Omega -> \mathbb{R}$= smallest sigma (E) on Ω such that every function in the sequence = measurable wrt that (E) ie $\sigma(f_n) = \sigma(\omega \subset \Omega : f_n(\omega) \in \mathcal{B})$
- Law of a random variable L_X:

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

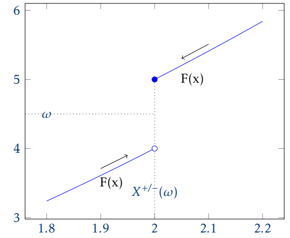
$$[0,1] \xleftarrow{\mathbb{P}} \sigma(X) \text{ or } \mathcal{F} \xleftarrow{X^{-1}} \mathcal{B}$$

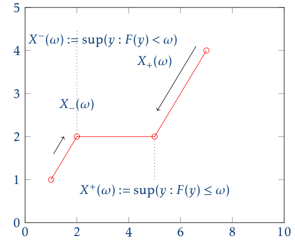
$$L_X = X^{-1} \circ \mathbb{P} : \mathcal{B} \rightarrow [0,1]$$

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

- **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)

$$X^+(\omega) := inf(z : F(z) > w) = sup(y : F(y) \leq \omega) \qquad (1)$$
$$X^-(\omega) := inf(z : F(z) \geq w) = sup(y : F(y) < \omega)$$





- $X^-(\omega) = sup(y : F(y) \leq \omega)$
- 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+ $\omega < F(c) \Rightarrow X^+(\omega) \leq c$ so $F(c) \leq P(X+ \leq c)$ and X-<= X+ means $(X+ \neq X-) = \cup_{q\in Q} X^- \leq q \leq X^+$ 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 $= Y^{-1}(\text{Borel}) = \left\{\{w : Y(w) \in B\} \quad B \in \text{Borel}\right\}$
- $Z : \Omega \rightarrow \mathbb{R}$ 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(Y1,Y2,..Yn)-measurable for some $Y_i : \Omega \rightarrow \mathbb{R} \Leftrightarrow$ there is a Borel function on Rn $f : \mathbb{R}^{\ltimes} \rightarrow \mathbb{R}$ such that Z=f(Y1,Y2,...,Yn)
- $Z : \Omega \rightarrow \mathbb{R}$ is sigma($Y_{\text{uncountable}}$)-measurable for some $Y_{\text{uncountable}} : \Omega \rightarrow \mathbb{R} \Leftrightarrow$ there is a Borel function on RN

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

- **MONOTONE CLASS THEOREM** π systems $\to \sigma$ algebras $\Leftrightarrow$ indicator of elements of π systems $\to$ general mesurable functions
- **MONOTONE CLASS THEOREM** : $\mathcal{H}$ class of bounded functions $S \to \mathbb{R}$ 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 $\mathcal{I} \in \mathcal{H} \to$ 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 Ei $\Leftrightarrow$ the sigmas $\zeta_i = \{\emptyset, E_i, \Omega, E_i^c = \Omega \backslash E_i\}$ independent = the rvs I_{E_i} are independent and of course $\zeta_i = \sigma(I_{E_i})$
- 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() 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 $P\left\{(\limsup E_n)^c\right\} = 1$. With $(\limsup E_n)^c = \liminf E_n^c = \cup_m \cap_{n \geq m} E_n^c$ and $P(\cap_{n \geq m} E_n^c) = \prod_{n \geq m}(1 - p_n)$ becasue independent and limit $r \geq n \geq m \quad r \uparrow \infty$=ok because both sides monotone and $\prod_{n \geq m}(1 - p_n) \leq \exp(-\sum p_n) = 0$ isnce $1 - x < \exp(-x)$
- Example given Xn so that $P(X_n > x) = e^{-x} \to P(X > \alpha \ln n) = e^{-\alpha \ln n} = n^{-\alpha} = \frac{1}{n^\alpha}$ so $\to P(X > \alpha \ln n$ for infinitely many n$) =$ 0 or 1 depending if $\alpha > 1, \alpha \leq 1$. So with $\alpha = 1, L := \limsup \frac{X_n}{n} \to P(L \geq 1) \geq P(\frac{X_n}{\ln n} \geq 1$,infinitely often$) = 1$, but $P(L > 1 + \frac{2}{k}) = 0$ because with $\alpha = 1 + \frac{1}{k}$, $P(L > 1 + \frac{2}{k}) \leq P\left\{X_n > (1 + \frac{1}{k})\ln n, \text{infinitely often}\right\} = 0$ 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: $\omega \to X_n(\omega)$
- 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(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 $\{X_n\} := \tau := \cap \tau_n$ with $\tau_n = \sigma(X_n, X_{n+1}, ...)$
- 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 $F_3 \in \mathcal{T}$

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

then for every n, $F_3^n = F_3$. But the $X_n, X_{n+1}, ..$ 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 $\chi_n = \sigma(X_1.X_2..., X_n)$ and $\mathcal{T}_n = \sigma(X_{n+1}.X_{n+2}...)$ .Then **(1)** $\chi_n, \mathcal{T}_n$ are independent because pi-system $X_i < x_i$ generates χ_i and pi-system $X_j < x_j, n+1 \leq j \leq n+r, r \in \mathbf{N}$ generates $\mathcal{T}_n$ and these pi-systems are independents. **(2)** $\chi_n, \mathcal{T}$ are independent because $\mathcal{T} \in \mathcal{T}_n$ **(3)** $\chi_\infty := \sigma(\chi_n), \mathcal{T}$ are independent because $\cup_n \chi_n$ 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 poitn it goes 0 -> 1 so let $c = sup\left\{x : \left\{P(\eta <= x) = 0\right\}\right\}$ and then usual $P(\eta \leq c - 1/n) = 0$ $P(\eta \leq c + 1/n) = 1$ and $P(\cup) = 0$ and $P(\cap) = 1$ so P(=c)=1
- **Note branching example** : M_∞ is in tail alebra 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 $X_n = Y_0 Y_1 ... Y_n$. the X_n are independent. Define $\Gamma = \sigma(X_1, X_2, ...)$ and $\mathcal{T}_n = \sigma(X_n, X_{n+1}, ...)$. Then we can prove $\Lambda := \cap_n \sigma(\mathcal{T}_n, \Gamma) \neq \zeta := \sigma(\cap_n \mathcal{T}_n, \Gamma)$ **Proof**: $\mathcal{T}_0 = \sigma(X_1, X_2, ..), \mathcal{T}_1 = \sigma(X_2, X_3, ..), \mathcal{T}_{n-1} = \sigma(X_n, X_{n+1}, ..), Y_0 = \frac{X_n}{Y_1 Y_2 ... Y_n}$ so (a) $Y_0 \in \Gamma$ and also $\in \mathcal{T}_{n-1} \to Y_0 \in \Lambda := \cap_n \sigma(\Gamma, \mathcal{T}_{n-1})$ (b) if $\Pi = \cap$-stable then independence wrt $\Pi \Leftrightarrow$ sigma(Π), and G:=$\cap_n \mathcal{T}_n \cap \Gamma$ is $\cap$-stable, and Y_0 independent of G because $P\left\{(Y_0 \in A) \cap g\right\} = P(Y_0 \in A)P(g), \forall A \in Borel, \forall g \in G$ because any g is $g = g_1 \cup g_2, g_1 \in \Gamma, g_2 \in \cap_n \mathcal{T}_{n-1}$ 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 $f \in m\Sigma+ \to \mu(f) = \sup \mu(simple), simple \leq f$
- MON $\boxed{f_n \in m\Sigma+, f_n \uparrow f \Rightarrow \mu(f_n) \uparrow \mu(f) := \int_S f_n(s)\mu(ds) \uparrow \int_S f(s}$

  (Recall MON for measure 1.10.a: $F_n \in \Sigma, F_n \uparrow F \Rightarrow \mu(F_n) \uparrow \mu(F) \longleftarrow$ 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 $x = 0, \frac{i-1}{2^r} < x \leq \frac{i}{2^r} \leq r(i \in \mathbf{N}), x > r$ 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. $\boxed{f_n \in m\Sigma+, f_n \uparrow f \text{ A.E. } \Rightarrow \mu(f_n) \uparrow \mu(f)}$ **Proof:**
- FATOU - illi $\boxed{\mu(\liminf f_n) \leq \liminf \mu(f_n) \ \forall f_n \in m\Sigma+}$ $\longleftarrow$ Recall: $\boxed{P(\liminf E_n) < \liminf P(E_n)}$ **Proof:** define $g_k := \inf_{n \geq k} f_n \to \liminf_n f_n \overset{(*)}{=} :\uparrow \lim g_k, \forall n \geq k : f_n \geq g_k \to \mu(f_n) \geq \mu(g_k), \Rightarrow \mu(g_k) \leq \inf_{n \geq k} f_n$ so $\mu(\liminf_n f_n) \overset{(*)}{=} \uparrow \lim_k \mu(g_k) \leq \uparrow \lim_k \inf_{n \geq k} \mu(f_n) := \liminf_n \mu(f_n)$
- FATOU2 – needs a dominating FINITENESS $\mu(g)$

$$\boxed{f_n, g \in m\Sigma+, f_n \leq g \forall n, \mu(g) < \infty \Rightarrow \mu(\limsup f_n) \geq \limsup \mu(f_n)}$$

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

- $f \in m\Sigma, f = f^+ - f^- := \max(f, 0) - \max(-f, 0), |f| = f^+ + f^-$ 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. $f_n(s) -> f(s)$
  3. $|f_n(s)| \leq g(s) \ \forall s \in S, n \in \mathbf{N}$
  4. $\mu(g) < \infty$

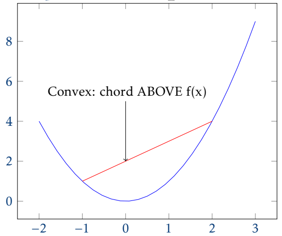
5. then $\Rightarrow f_n \to f$ in $L^1(S,\Sigma,\mu)$
6. ie $\mu\{|f_n - f|\} \to 0$
7. note therefore $\mu(f_n) \to \mu(f)$

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

- SCHEFFE $f_n, f \in L^{1,+}, f_n \to f$ a.e then $\mu|f_n - f| \to 0 \Leftrightarrow \mu(f_n) \to \mu(f)$ **Proof** $\hookrightarrow$ $|\mu(f_n - f)| \leq \mu|f_n - f| \to 0$ so done. **Proof** $\hookleftarrow$ so suppose $\mu(f_n) \to \mu(f)$ and by hyp $f_n \to f$ then (a) $because (f_n - f)^- \leq f \overset{DOM}{\Rightarrow} \mu\{(f_n - f)^-\} \to 0$ (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 $|\mu(f_n - f; f_n < f)| \leq \mu(f_n - f)^- \to 0$
- 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 $\mu(f) = 0 \to \lambda(F) = 0\ \forall F \Rightarrow \lambda = f\mu$ 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 $X_n \overset{a.s.}{\Rightarrow} X$ ie $P(X_n \to X) = 1$
  2. MON-Prob $\boxed{0 \leq X_n \uparrow X \Rightarrow E(X_n) \uparrow E(X) \leq \infty}$ $\hookleftarrow$ $f_n \in m\Sigma+, f_n \uparrow f \Rightarrow \mu(f_n) \uparrow \mu(f) := \int_S f_n(s)\mu(ds) \uparrow \int_S f(s)\mu(ds)$
  3. FATOU-Prob $\boxed{X_n \geq 0 \Rightarrow E(X) \leq \liminf E(X_n)}$ $\hookleftarrow$ $\mu(\liminf f_n) \leq \liminf \mu(f_n)\ \forall f_n \in m\Sigma+$
  4. DOM-Prob $\boxed{|X_n(\omega)| \leq Y(\omega), E(Y) < \infty \Rightarrow E|x_n - X| \to 0}$ $\hookleftarrow$ see long one above
  5. SCHEFFE - Prob $\boxed{E|X_n| \to E|X| \Rightarrow E|X_n - X| = 0}$ $\hookleftarrow$ $f_n, f \in L^{1,+}, f_n \to f$ a.e then $\mu|f_n - f| \to 0 \Leftrightarrow \mu(f_n) \to \mu(f)$
  6. BDD-DOMK $\boxed{|X_n(\omega)| \leq K \Rightarrow E|X_n - X| \to 0}$ $\hookleftarrow$ Note $E(K) < \infty$ because $P(\Omega) = 1$
- notation $E(X,F) := \int_F X(\omega)P(d\omega) := E(XI_F)$
- MARKOV
  1. $Z \in mF, g(nondec) : \mathbb{R} -> [0,\infty]\ Borel$
  2. then $g \circ Z \in mF+$ and so
  3. $\Rightarrow$ $\boxed{E(gZ) \geq E(gZ; Z \geq c) \geq g(c)P(Z \geq c)}$ $\hookleftarrow$ 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 $e^{-\theta c}E(e^{\theta Y}) \geq P(Y \geq c)$ $\hookleftarrow$ choose a good c !
- useful results to remember
  1. $X \in mF+, E(X) < \infty$ ie integrable , then obviously $P(X < \infty) = 1$ $\hookleftarrow$ remember Markov above $X \in mF+, P(Z \geq c) \leq \frac{E(Z)}{c}$
  2. $Z_k \in mF+$ then $E(\sum Z_k) \overset{linear}{=} \sum E(Z_k) \overset{MON}{\leq} \infty$
  3. $Z_k \in mF+$ with $E(\sum Z_k) < \infty \Rightarrow \sum Z_k < \infty$ (a.s) and therefore $\Rightarrow Z_k \to 0$ (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 $E(Z_k) = P(F_k)$ ($\hookleftarrow E(1_A) = P(A)$ 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. $c : G \to \mathbb{R}$
  2. $E|X| < \infty$
  3. $P(X \in G) = 1$
  4. $E|c(X)| < \infty$
  5. then $\boxed{\Rightarrow c(E(X)) \leq E(c(X))}$ intuition $\hookleftarrow$ $c(px + qy) \leq pc(x) + qc(y)$ so $\hookleftarrow$ $c(p_1x_1 + p_2x_2 + ..p_nx_n) \leq p_1c(x_1) + p_2c(x_2) + ... + p_nc(x_n)$
  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 $x \to X$ (a.s) $c(X) \geq m(X - \mu) + c(v)$ and take E(.) 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 $E(X_n)^{\frac{1}{p}} \leq E|Y|^{r\frac{1}{r}} = \|Y_r\|$ 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. $< X, Y >= E(XY)$
  4. correl = $cos\theta = \frac{<U,V>}{\|U\|_2\|V\|_2}$
- **Completeness of** L^p **:**
  1. L^p is complete (Cauchy series converge and limit inside)
  2. X_n Cauchy sequence in $L^p \Leftrightarrow \sup_{r,s\geq k} \|X_r - X_s\|_p \to 0$
  3. X_n Cauchy sequence $\Leftrightarrow \exists X \in L^p \therefore X_n \to X \in L^p$ ie $\|X_r - X\|_p \to 0$
  4. Proof: choose sequence $\|X_{\alpha_{n+1}} - X_{\alpha_n}\|_p < 2^{-n}$ then with p=1 and monotonocity $E|X_{\alpha_{n+1}} - X_{\alpha_n}| = \|X_{\alpha_{n+1}} - X_{\alpha_n}\|_1 \leq \|X_{\alpha_{n+1}} - X_{\alpha_n}\|_p$ 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 $X_{\alpha_n} \to X$. Then this X is the one to use for Lp 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') : $\mathbb{E}(|X_r - X|^p) \leq \liminf_{t\to\infty} \mathbb{E}(|X_r - X_{k_t}|^p) \leq 2^{-np}$, so $X_r - X$ in L^p , so X in L^p since L^p is a vector space and $X_r \to X$
- **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 $|\mu(fh)|\le\mu(|fh|)\le\|f\|_p\|h\|_q$ **Proof** say $f,h\ge0,\mu(f^p)>0$ then define a prob measure $P=\frac{f^p\mu}{\mu(f^p)}$ [$\hookleftarrow$ 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 $u(s):=\frac{h(s)}{f(s)^{p-1}},f(s)>0;0$ otherwise, then because $P(u)^q\le P(u^q)$ [$\hookleftarrow$ recall JENSEN $c(E(X))\le E(c(x))$ with $c(x)=x^q$ convex since $q>1$] , then $\mu(|fh|)\le\|f\|_p\|hI_{f>0}\|_q\le\|f\|_p\|h\|_q$
- **Intuition** Holder (p=q=2) = Cauchy schwartz . Holder = $\|fg\|_1\le\|f\|_p\|g\|_p$ Cauchy = $\|fg\|_1\le\|f\|_2\|g\|_2$ ie $\left(\int fg\right)^2\leqslant\int f^2\cdot\int g^2$ and norm $\int g^q=\|g\|_q^q$, $\int f^p=\|f\|_p^p$, $\int fg=\|fg\|_1$
- HÖLDER $|E(XY)|\le E(|XY|)\le E(|X|^p)^{\frac{1}{p}}E(|X|^q)^{\frac{1}{q}}$ Proof : $Q(\Lambda)=\frac{E(1_\Lambda X^p)}{C=E(X^p)}$, $Z=\frac{Y}{X^{p-1}}1_{X>0},(E_Q(Z))^q\le E(Z^q)$ then $\frac{1}{C^q}E(XY)^q=\frac{1}{C^q}E(YX^{1-p}X^p)^q=(E_Q(YX^{1-p}))^q\overset{\text{JENSEN}}{\le}E_Q((YX^{1-p})^q)=\frac{1}{C}E((YX^{1-p})^qX^p)=\frac{1}{C}E(Y^qX^{q(1-p)}X^p)=\frac{1}{C}E(Y^qX^{-p}X^p)=\frac{1}{C}E(Y^p)$ so $(E(XY))^q\le C^{q-1}E(Y^q)$ so $E(XY)\le C^{\frac{q-1}{q}}E(Y^q)^{\frac{1}{q}}=C^{\frac{1}{p}}E(Y^q)^{\frac{1}{q}}=E(X^p)^{\frac{1}{p}}E(Y^q)^{\frac{1}{q}}$
- MINKOWSKI (p,sum,triangle) $\|f+g\|_p\le\|f\|_p+\|g\|_p$ with $f,g\in L^p$ **Proof** use Holder $\mu(|f+g|^p)=\mu(|f+g|^{p-1}|f+g|)\le\mu(|f+g|^{p-1}|f|)+\mu(|f+g|^{p-1}|g|)\overset{\text{JENSENx2}}{\le}\|f\|_p\ \||f+g|^{q-1}\|_q+\|g\|_p\ \||f+g|^{q-1}\|_q$

8 Chapter 7 - Easy strong law

- **Independence=Multiply** $E(XY)=E(X)E(Y)$ **Proof:** use staircase function [$\hookleftarrow$ Recall $a^{(r)}=0,\frac{i-1}{2^r},r$ for $x=0,\frac{i-1}{2^r}<x\le\frac{i}{2^r}\le r(\ i\in\mathbf{N}),x>r$] and set $f^{(r)}=a^{(r)}\circ f\uparrow f(MON)$ on disjoint partitions and $a^{(r)}(X)=\sum a_iI_{A_i},a^{(r)}(Y)=\sum b_iI_{B_i}$ on $A_i,B_i\in\sigma(X),\sigma(Y)$ and $P(A_i\cap B_i)=P(A_i)P(B_i)$
- **Independence** $\mathrm{Cov}(XY)=0$ and $\mathrm{Var}(X+Y)=0$
- **Strong Law - simple form - 4th moments -avg=0** $X_1,X_2...$ indeps, $E(X_i)=0,E(X_i^4)\le K\Rightarrow$ let $S_n=X_1+..+X_n$ then $P(\frac{S_n}{n}\to0)=1$ ie $\frac{S_n}{n}\to0$ a.s $\hookleftarrow$ note the average is 0 in this form here. **Proof** terms with single $X_{(.)}^1$ ie $E(X_iX_j^2X_k)=E(X_iX_j^3)=E(X_iX_jX_kX_l)=0$ with indep and $E(X_i)=0$, terms like $E(X_i^2X_j^2)=E(X_i^2)E(X_j^2)\le K$ because $E(X_i^2)^2\overset{\text{Jensen}}{\le}E(X_i^4)\le K$ so $E(S_n^4)=E((X_1+..X_n)^4)=E(\sum_k X_k^4+6\sum\sum_{i<j}X_i^2X_j^2)\le nK+6\frac{n(n-1)}{2}K\le3Kn^2$ and these are sums of +ve rvs so $E(\sum(\frac{S_n}{n})^4)\le3K\sum\frac{1}{n^2}<\infty$ so (again see sum of +ve rvs) $\sum(\frac{S_n}{n})^4<\infty$ a.s $\Rightarrow\frac{S_n}{n}\to0$ a.s
- **Chebyshev** $P(|X-\mu|>c)\le\frac{\mathrm{Var}(X)}{c^2}$ **Proof** $P(|X-\mu|>c)=P((X-\mu)^2>c^2)\overset{\text{markov}}{\le}\frac{E(X-\mu)^2)}{c^2}$ **Example** Bernoulli $P(X_i=1)=p,P(X_i=0)=1-p$ then $E(X_i)=1.p+0.(1-p)=p$ and $\mathrm{Var}(X_i)=E(X_i{}^2)-E(X_i)^2=p-p^2=p(1-p)$ so $E(S_n)=np,\mathrm{Var}(S_n)\overset{\text{indep.}}{=}\sum\mathrm{Var}(X_i)=np(1-p)\le n/4$ so $E(S_n/n)=p,\mathrm{Var}(\frac{S_n}{n})=\frac{\mathrm{Var}\,S_n}{n^2}\le\frac{1}{4n}$ so **Chebyshev** $P(|\frac{S_n}{n}-p|>c)\le\frac{1}{4nc^2}$
- **Example Weierstrass polynomial** $\sup_{s\in[0,1]}|B(x)-f(x)|\le\epsilon$ 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 $B_n(p):=Ef(\frac{S_n}{n})=\sum_0^n f(\frac{k}{n})\binom{n}{k}p^k(1-p)^{n-k}$ then $|B_n(p)-f(p)|=|Ef(\frac{S_n}{n})-f(p)|\overset{\text{jensen}}{\le}E(Y_n:=|f(\frac{S_n}{n})-f(p)|)$ with $Z_n:=\frac{S_n}{n}-p$ then Z_n small $\to$ Y_n small and $E(Y_n)=E(Y_n;Z_n\le\delta)+E(Y_n;Z_n>\delta)\le$ small x $P(Z_n\le\delta)+2KP(Z_n>\delta)\overset{\text{see above}}{\le}$ small x $P(Z_n\le\delta)+2K\frac{1}{4n\delta^2}$ ie small for n big enough.

9 Chapter 8 - Product measures

- statement : can interchange $\int\int$ if $f+$ (in which case 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) x (whole spaces)

10 Chapter 9 - Martingales I : Conditional expectation

11 Chapter 10 - Martingales II : Martingales

12 Chapter 11 - Martingales III : Martingale Convergence theorem

13 Chapter 12 - Martingales III : Martingale Bounded in L2

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