Lecture 6 : Independence

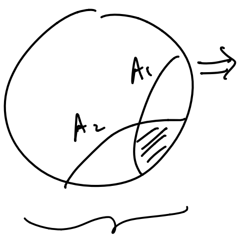
1

Let $(\Omega, \mathcal{F}, P)$ be a probability space

sample space $\uparrow \Omega$, σ-field $\uparrow \mathcal{F}$, prob measure $\uparrow P$

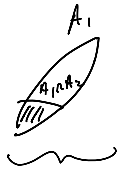
Suppose $A_1, A_2 \in \mathcal{F}$ with $P(A_1) > 0$ & $P(A_2) > 0$.
Recall from undergrad probability

$$P(A_2 | A_1) \equiv \text{prob of } A_2 \text{ given } A_1 := \frac{P(A_1 \cap A_2)}{P(A_1)}$$



Sample space Ω

$\Rightarrow$ Sample a random $\omega \in \Omega$ according to P. Your told $\omega \in A_1$ but nothing else $\Rightarrow$



New sample space.

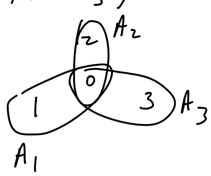
New prob $\frac{P(\cdot \cap A_1)}{P(A_1)}$

A_1 is <u>independent</u> of A_2 if $P(A_2|A_1) = P(A_2)$

i.e. if $P(A_1 \cap A_2) = P(A_1)P(A_2)$.

<u>Question</u>: How to make sense of independence among a collection of events (possibly uncountably many)? Is pairwise independence enough?

e.g. $\Omega = \{0,1,2,3\}$, $\mathcal{F} = 2^\Omega$, $P = $ unif on Ω.



$$i \neq j \Rightarrow P(\underbrace{\underbrace{A_i \cap A_j}_{=\{0\}}}_{=1/4}) = \underbrace{P(A_i)}_{1/2}\underbrace{P(A_j)}_{1/2}$$

so A_1, A_2, A_3 are pairwise independent.
But A_1, A_2, A_3 are not jointly indep:

$$P(A_1 \cap A_2 \cap A_3) = 1/4$$
$$P(A_1)P(A_2)P(A_3) = \frac{1}{8}$$

(Note: $P(A_1 | A_2 \cap A_3) = 1 \neq P(A_1)$)

2

e.g. Let $A_1, \ldots, A_n$ be events (i.e. $A_i \in \mathcal{F}$) s.t. $A_1 = \emptyset$. Then

$$P(A_1 \cap \cdots \cap A_n) = 0 = P(A_1) \cdots P(A_n)$$

So the full factorization criterian will not work as a def of independence either

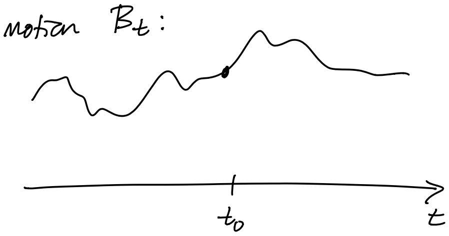
Here is the "right" def of indep for a collection of events.

Def: A collection of events $\{A_k\}_{k \in \mathcal{K}}$ are <u>independent events</u> iff $\forall$ finite $\mathcal{H} \subset \mathcal{K}$

$$P\Big(\bigcap_{h \in \mathcal{H}} A_h\Big) = \prod_{h \in \mathcal{H}} P(A_h).$$

Note: $\mathcal{K}$ is allowed to be <u>any</u> index set.

we will also need the notion of independent σ-fields to make sense of things like the strong markov property of Brownian motion B_t:



$\sigma\langle B_t : t < t_0 \rangle$ is indep of $\sigma\langle B_t : t > t_0 \rangle$ given $\sigma\langle B_{t_0} \rangle$.

Def: Let $\mathcal{K}$ be an arbitrary index set. $\forall k \in \mathcal{K}$, let $\mathcal{A}_k$ be a collection of events. The $\mathcal{A}_k$'s are <u>independent collections</u> if $\{A_k\}_{k \in \mathcal{K}}$ are independent events for each choice $A_k \in \mathcal{A}_k$.

Thm: Let $\mathcal{A}_k, \mathcal{B}_k$ be collections of events for each $k \in \mathcal{K}$ (arb index set). Then

(i) (subclasses):

If $\mathcal{A}_k \subset \mathcal{B}_k$ $\forall k \in \mathcal{K}$ & the $\mathcal{B}_k$'s are indep then the $\mathcal{A}_k$'s are indep.

(ii) (augmentation):

$\mathcal{A}_k$'s are indep. iff $\mathcal{A}_k \cup \{\Omega\}$'s are indep.

(iii) (simplified product):

If $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ all contain Ω then the $\mathcal{A}_k$'s are indep iff

$$P\left(\bigcap_{k=1}^n A_k\right) = \prod_{k=1}^n P(A_k)$$

$\forall A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n$.

Proof:

(i): trivial

(ii): $\Leftarrow$ follows by "subclasses".

For $\Rightarrow$ choose $A_k \in \mathcal{A}_k \cup \{\Omega\}$ & finite $\mathcal{H} \subset \mathcal{K}$. Let $\mathcal{K}_0 = \{k : A_k \in \mathcal{A}_k\}$.

$$\therefore P\left(\bigcap_{h \in \mathcal{H}} A_h\right)$$

$$= P\left(\bigcap_{h \in \mathcal{H} \cap \mathcal{K}_0} A_h\right), \quad A_k = \Omega \text{ when } k \in \mathcal{K} - \mathcal{K}_0$$

$$= \prod_{h \in \mathcal{H} \cap \mathcal{K}_0} P(A_h), \quad \mathcal{A}_k\text{'s indep}$$

$$= \prod_{h \in \mathcal{H}} P(A_h), \quad P(A_k) = P(\Omega) = 1 \text{ when } k \in \mathcal{K} - \mathcal{K}_0$$

$\therefore$ $\mathcal{A}_k \cup \{\Omega\}$'s are indep.

(iii) $P\left(\bigcap_{k=1}^n A_k\right) = \prod_{k=1}^n P(A_k)$

$$\Rightarrow P\left(\bigcap_{h \in \mathcal{H}} A_h\right) = \prod_{h \in \mathcal{H}} P(A_h)$$

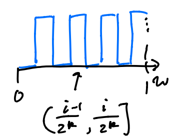
for $\mathcal{H} \subset \{1, 2, \dots, n\}$ by replacing A_k with $\Omega \in \mathcal{A}_k$, $k \notin \mathcal{H}$.

QED

e.g. Coin flip model from lecture 1:

$\Omega = (0,1]$, $\mathcal{F} = \mathcal{B}((0,1])$, P = uniform measure.

$X_k(\omega) :=$ k^{th} binary digit of ω =

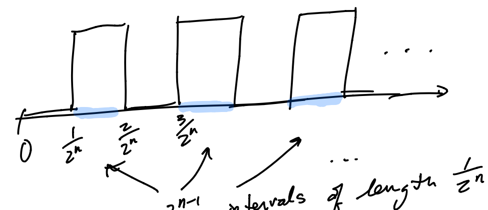


$H_k := \{\omega \in \Omega : X_k(\omega) = 1\}$

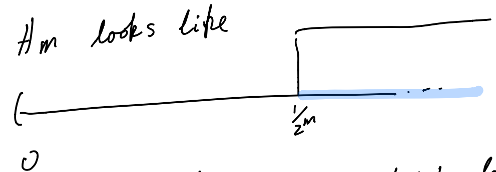
event of flipping a heads on the k^{th} toss if we want X_k to model fair coin flips.

Claim: $H_1, H_2, H_3, \dots$ are indep events.

Proof: $H_n = \left(\frac{1}{2^n}, \frac{2}{2^n}\right] \cup \left(\frac{3}{2^n}, \frac{4}{2^n}\right] \cup \left(\frac{5}{2^n}, \frac{6}{2^n}\right] \cdots$



If $m < n$ then H_m looks like



$\therefore$ $H_n \cap H_m$ = union of half of the disjoint intervals that make up H_n

Let $1 \le i_1 < i_2 < \dots < i_n$ & show

$$P(H_{i_1} \cap \dots \cap H_{i_n}) = \underbrace{P(H_{i_1}) \cdots P(H_{i_n})}_{= \frac{1}{2^n}}$$

Now $H_{i_n}\cap H_{i_{n-1}}\cap\cdots\cap H_{i_1} = \frac{2^{i_n-1}}{2^{n-1}}$ disjoint intervals of length $\frac{1}{2^{i_n}}$

2^{i_n-1} intervals of length $\frac{1}{2^{i_n}}$

reduce the # of intervals by ½ for each further intersection

$\therefore\ P(H_{i_n}\cap H_{i_{n-1}}\cap\cdots\cap H_{i_1}) = \frac{2^{i_n-1}}{2^{n-1}}\cdot\frac{1}{2^{i_n}} = \frac{1}{2^n}$

as was to be shown QED.

π-generators are enough & ANOVA

At this point checking two σ-fields are indep would be a daunting task since we have no representation for general events in a σ-field.

The following thm helps this.

Thm (π-generators are enough):

Let $\mathcal{A}_k \subset \mathcal{F}$, $\forall k\in\mathcal{K}$. Then

$\mathcal{A}_k$'s are indep π-systems

$\Longrightarrow \sigma\langle\mathcal{A}_k\rangle$'s are independent.

Proof: Let $\mathcal{B}_k := \mathcal{A}_k \cup \{\Omega\}$

Suppose the $\mathcal{A}_k$'s are indep π-sys

$\therefore$ the $\mathcal{B}_k$'s are indep π-sys, by augmentation

$\therefore\ \forall$ distinct $k_1, k_2, \dots, k_n \in \mathcal{K}$

the $\mathcal{B}_{k_1}, \mathcal{B}_{k_2}, \dots, \mathcal{B}_{k_n}$ are indep π-sys

show $\sigma\langle\mathcal{B}_{k_1}\rangle, \mathcal{B}_{k_2}, \dots, \mathcal{B}_{k_n}$ are indep π-sys

and we will be done (by induction)

By the simplified product criterion this is equivalent to showing

$$P(B_1\cap\cdots\cap B_n) = P(B_1)\cdots P(B_n) \quad (*)$$

$\forall\, B_1\in\sigma\langle\mathcal{B}_{k_1}\rangle,\ B_2\in\mathcal{B}_{k_2}, \dots, B_n\in\mathcal{B}_{k_n}$

Fixing $B_2, \dots, B_n$ let

$$\mathcal{L} := \{B_1\in\mathcal{F} : (*) \text{ holds}\}$$

& show $\sigma\langle\mathcal{B}_{k_1}\rangle \subset \mathcal{L}$.

- $\mathcal{B}_{k_1}\subset\mathcal{L}$: yes, since $\mathcal{B}_k$'s are indep.
- $\Omega\in\mathcal{L}$: yes, since $\Omega\in\mathcal{B}_k\ \forall k\in\mathcal{K}$.
- $B\in\mathcal{L}\Rightarrow$

$$P(B^c\cap \underbrace{B_2\cap\cdots}_{A}) = P(B_2\cap\cdots) - P(B\cap B_2\cap\cdots)$$

since $P(B^c\cap A) = P(A - B\cap A)$

$$= P(\Omega\cap B_2\cap\cdots) - P(B\cap B_2\cap\cdots)$$

$$= P(\Omega)\cdot P(B_2)\cdots - P(B)P(B_2)\cdots$$

since $\Omega, B\in\mathcal{L}$

$$= \underbrace{[P(\Omega) - P(B)]}_{P(B^c)} P(B_2)\cdots P(B_n)$$

$\Rightarrow B^c\in\mathcal{L}$

- $\underbrace{A_1, A_2, \dots}_{\text{disjoint}} \in\mathcal{L}$

$$\Rightarrow P\Big(\big(\bigcup_k A_k\big)\cap B_2\cap\cdots\cap B_n\Big)$$

$$= \sum_k P(A_k\cap B_2\cap\cdots\cap B_n)$$

$$= \sum_k P(A_k)P(B_2)\cdots P(B_n)$$

$$= P(B_2)\cdots P(B_n)\underbrace{\Big[\sum_k P(A_k)\Big]}_{P(\bigcup_k A_k)}$$

$\Rightarrow \bigcup_k A_k \in\mathcal{L}$

$\therefore$ $\mathcal{L}$ is a λ-sys & B_{k_1} is a π-sys.

$\therefore$ $\sigma\langle B_{k_1}\rangle \subset \mathcal{L}$.

QED.

e.g. coin flip example showed

$H_1, H_2, \dots$ are indep

since $\{H_k\}$ is a π-sys for each k,

$\sigma\langle H_1\rangle, \sigma\langle H_2\rangle, \dots$ are indep

σ-fields (where $\sigma\langle H_k\rangle = \{\emptyset, \Omega, H_k, H_k^c\}$)

$\therefore$ Any sequence $H_1, H_2^c, H_3, H_4^c, H_5^c, \dots$ are indep.

tails on the 5^{th} toss

To motivate the next thm let

A = the event $\sum_{k=1}^{n}(1-2X_{2k})=0$ for infinitely many n

B = the event $\sum_{k=1}^{n}(1-2X_{2k+1})=0$ for infinitely many n

is A indep of B?

Thm (ANOVA): Matrix of π-systems

$$\begin{matrix} \mathcal{A}_{11} & \mathcal{A}_{12} & \mathcal{A}_{13} & \cdots \\ \mathcal{A}_{21} & \mathcal{A}_{22} & \mathcal{A}_{23} & \cdots \\ \mathcal{A}_{31} & \mathcal{A}_{32} & \mathcal{A}_{33} & \cdots \\ \vdots & & & \ddots \end{matrix}$$

Let $R_i = \sigma\langle \underbrace{\mathcal{A}_{i1}, \mathcal{A}_{i2}, \dots}_{i^{th}\text{ Row}}\rangle$

Then

all the $\mathcal{A}_{ik}$'s are indep $\iff$ (i) R_k's are indep
(ii) the $\mathcal{A}_{ik}$'s within each row are independent

Proof:

($\Longrightarrow$) Suppose all the $\mathcal{A}_{ik}$'s are indep.

$\therefore$ (ii) clearly holds

To show (i) note

$R_k = \sigma\langle \underbrace{\mathcal{A}_{k1} \cup \mathcal{A}_{k2} \cup \cdots}_{\text{would like to use } \pi\text{-generators but this isn't a } \pi\text{-sys}} \rangle = \sigma\langle P_k\rangle$

where P_k = the closure of $\mathcal{A}_{k_1} \cup \mathcal{A}_{k_2} \cup \cdots$ under finite intersection

Clearly P_k's are π-systems.

Lets show the P_k's are indep.

(9)

Select one P_k from $\mathcal{P}_k$ and note:

$$\underbrace{P_{k_1} \cap \cdots \cap P_{k_n}}$$

write this as $\underbrace{(A_1 \cap \cdots)}_{\text{Row } k_1} \cap \underbrace{(B_1 \cap \cdots)}_{\text{Row } k_2} \cap \underbrace{(C_1 \cap \cdots)}_{\text{Row } k_3} \cap \cdots$

P_{k_1} — each event in here is from a unique $\mathcal{A}_{k_2, j}$

merging (via "$\cap$") multiple sets from the same $\mathcal{A}_{k_j}$ if necessary ... still a $\mathcal{A}_{ki}$ set by π-sys assumption

Now,

$$P(P_{k_1} \cap \cdots \cap P_{k_n}) = P(A_1) \cdots P(B_1) \cdots P(C_1) \cdots$$
$$= P(P_{k_1}) \cdots P(P_{k_n})$$

e.g.

$$P(\overbrace{\underbrace{A_1 \cap A_2}_{\text{both in } \mathcal{A}_{k_1,1}}}^{P_{k_1}} \cap \overbrace{\underbrace{B_1}_{\in \mathcal{A}_{k_2,1}} \cap \underbrace{B_2}_{\in \mathcal{A}_{k_2,2}}}^{P_{k_2}})$$

$$= P(\underbrace{A_1 \cap A_2}_{\in \mathcal{A}_{k_1,1} \text{ by } \pi\text{-sys}}) P(B_1) P(B_2)$$

since $\mathcal{A}_{k_1,1}$ & $\mathcal{A}_{k_2,1}$ & $\mathcal{A}_{k_2,2}$ are indep.

$$= P(A_1 \cap A_2) P(B_1 \cap B_2)$$

since $\mathcal{A}_{k_2,1}$ & $\mathcal{A}_{k_2,2}$ are indep.

$$= P(P_{k_1}) P(P_{k_2})$$

$\therefore$ $\mathcal{P}_k$'s are indep π-sys.

$\therefore$ The σ-fields $\mathcal{R}_k := \sigma\langle \mathcal{P}_k \rangle$ are independent by π-generators.

(10)

($\Leftarrow$)

Suppose the row σ-fields $\mathcal{R}_k$ are indep & the $\mathcal{A}_{k_i}$'s within each Row are indep.

Let $\mathcal{H}$ be a finite set of (Row, col) index tuples

For each $(i,k) \in \mathcal{H}$ choose one $A_{ik} \in \mathcal{A}_{ik}$.

$$\therefore P\Big(\bigcap_{(i,k)\in\mathcal{H}} A_{ik}\Big) = P\Big(\bigcap_{\substack{\text{Rows } i \\ \text{in } \mathcal{H}}} \overbrace{\bigcap_{\substack{\text{Cols } k \\ \text{s.t. } (i,k)\in\mathcal{H}}} A_{ik}}^{\in \mathcal{R}_i}\Big)$$

$$\overset{R_i\text{'s indep}}{=} \prod_{\substack{\text{Rows } i \\ \text{in } \mathcal{H}}} P\Big(\bigcap_{\substack{\text{Cols } k \\ \text{s.t. } (i,k)\in\mathcal{H}}} A_{ik}\Big)$$

$$\overset{\text{within Row indep}}{=} \prod_{\substack{\text{Rows } i \\ \text{in } \mathcal{H}}} \prod_{\substack{\text{Cols } k \\ \text{s.t. } (i,k)\in\mathcal{H}}} P(A_{ik})$$

$$= \prod_{(i,k)\in\mathcal{H}} P(A_{ik})$$

QED

(11)

Kolmogorov's 0-1 law

Let $\mathcal{A}_1, \mathcal{A}_2, \ldots$ be a sequence of collections of $\mathcal{F}$-sets (i.e. $\mathcal{A}_k \subset \mathcal{F}$)

Def: The tail σ-field of the $\mathcal{A}_k$'s is defined as

$$\mathcal{T} := \bigcap_{m=1}^{\infty} \sigma\langle \mathcal{A}_m, \mathcal{A}_{m+1}, \ldots \rangle$$
$$= \{ A \in \mathcal{F} : A \in \sigma\langle \mathcal{A}_m, \mathcal{A}_{m+1}, \ldots \rangle \ \forall m \}$$

($\mathcal{T}$ is a σ-field for the same reason $\sigma\langle \mathcal{E} \rangle$ is)

e.g. For the coin flip model from lecture 1 we have

$$\frac{S_n}{n} \to 0 \iff \frac{X_1 + \cdots + X_n}{n} \to \tfrac{1}{2}$$
$$\iff \frac{X_1 + \cdots + X_{m-1} + X_m + \cdots + X_n}{n} \to \tfrac{1}{2}$$
$$\iff \frac{X_m + \cdots + X_n}{n} \to \tfrac{1}{2} \quad \forall m$$

where $\left\{ \frac{X_m + \cdots + X_n}{n} \xrightarrow{n \to \infty} \tfrac{1}{2} \right\} \in \sigma\langle H_m, H_{m+1}, \ldots \rangle$ $\forall m$

$\therefore N = \left\{ \frac{S_n}{n} \to 0 \right\} \in$ tail σ-field generated by $H_1, H_2, \ldots$

(12)

Thm (Kolmogorov's 0-1 law)

If $\mathcal{A}_1, \mathcal{A}_2, \ldots$ are indep π-systems then $\forall A \in \mathcal{T}$, $P(A) = 0$ or $P(A) = 1$.

tail σ-field generated by the $\mathcal{A}_k$'s

Proof: show A is independent of itself.

$\mathcal{A}_1, \ldots, \mathcal{A}_{m-1}, \mathcal{A}_m, \mathcal{A}_{m+1} \ldots$ are indep π-sys.

$\therefore \sigma\langle \mathcal{A}_1 \rangle, \ldots \sigma\langle \mathcal{A}_{m-1} \rangle, \sigma\langle \mathcal{A}_m, \mathcal{A}_{m+1}, \ldots \rangle$ are indep π-sys by anova.

$\therefore \sigma\langle \mathcal{A}_1 \rangle, \ldots, \sigma\langle \mathcal{A}_{m-1} \rangle, \mathcal{T}$ are indep π-sys by subclasses.

holds for all m

$\therefore \sigma\langle \mathcal{A}_1 \rangle, \sigma\langle \mathcal{A}_2 \rangle, \ldots, \mathcal{T}$ are indep π-sys by the finite selection requirement of the def of indep.

$\therefore \sigma\langle \mathcal{A}_1, \mathcal{A}_2, \ldots \rangle, \mathcal{T}$ are indep π-sys by Anova.

$\therefore \mathcal{T}, \mathcal{T}$ are indep π-sys by subclasses

$\therefore \forall A \in \mathcal{T}$, $P(A \cap A) = P(A)P(A)$

$\therefore P(A) = 0$ or 1.

QED

Borel-Cantelli and Fatou

Let $A_1, A_2, \ldots \in \mathcal{F}$.

Def:

$$\{A_n \text{ i.o.}_n\} := \{\omega \in \Omega : \omega \in A_n \text{ infinitely often in } n\}$$

$$:= \underbrace{\bigcap_{m=1}^{\infty} \bigcup_{n \geq m} A_n}_{\forall m \, \exists n \geq m \text{ s.t. } \omega \in A_n.}$$

$$\{A_n \text{ a.a.}_n\} := \{\omega \in \Omega : \omega \in A_n \text{ for all but finitely many } n\}$$

$$:= \underbrace{\bigcup_{m=1}^{\infty} \bigcap_{n \geq m} A_n}_{\exists m \text{ s.t. } \forall n \geq m,\ \omega \in A_n}$$

Note: $\{A_n \text{ i.o.}_n\} \in \mathcal{F}$ & $\{A_n \text{ a.a.}_n\} \in \mathcal{F}$

Sometimes people write

$\limsup_{n \to \infty} A_n$ for $\{A_n \text{ i.o.}_n\}$

$\liminf_{n \to \infty} A_n$ for $\{A_n \text{ a.a.}_n\}$

since

indicator of A_n

$$\limsup_n I_{A_n}(\omega) = I_{\{A_n \text{ i.o.}_n\}}(\omega)$$

$$\liminf_n I_{A_n}(\omega) = I_{\{A_n \text{ a.a.}_n\}}(\omega)$$

Some Facts:

$$\{A_n \text{ a.a.}_n\} \subset \{A_n \text{ i.o.}_n\}$$

$$\{A_n \text{ a.a.}_n\}^c = \{A_n^c \text{ i.o.}_n\}$$ & vice-versa

$$\{A_n \text{ a.a.}_n\} = \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} A_n$$

$$= \Big(\bigcap_{n \geq 1} A_n\Big) \cup \Big(\bigcap_{n \geq 2} A_n\Big) \cup \cdots$$

these grow since your removing restrictions

$$= \bigcup_{m=k}^{\infty} \bigcap_{n \geq m} A_n, \text{ for any } k$$

since anything in the first $k-1$ terms are included in the latter.

$\in$ tail σ-field generated by $\{A_1\}, \{A_2\}, \ldots$

$$A_n \uparrow A \Rightarrow A = \bigcup_{m=1}^{\infty} A_m \ \& \ A_1 \subset A_2 \subset \cdots$$

$$\Rightarrow A = \bigcup_{m=1}^{\infty} A_m = \bigcup_{m=1}^{\infty} \underbrace{\bigcap_{n \geq m} A_n}_{= A_m} = \{A_n \text{ a.a.}_n\}$$

$$\{A_n \text{ i.o.n}\} = \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} A_n$$

$$= \Big(\bigcup_{n\geq 1} A_n\Big) \cap \Big(\bigcup_{n \geq 2} A_n\Big) \cap \cdots$$

these decrease as sets

$$= \bigcap_{m=k}^{\infty} \bigcup_{n \geq m} A_n, \quad \forall k$$

since the restrictions found in the first $k-1$ terms is already in the k^{th} term.

$\in$ tail σ-field generated by $\{A_1\}, \{A_2\}, \ldots$

$$A_n \downarrow A \Leftrightarrow A_n^c \uparrow A^c$$

$$\Rightarrow A^c = \{A_n^c \text{ a.a.n}\}$$

$$\Rightarrow A = \{A_n \text{ i.o.n}\}$$

Note: The 0-1 law already implies

$A_1, A_2 \ldots \in \mathcal{F}$ are indep

$$\Rightarrow P(A_n \text{ i.o.n}) = 0 \text{ or } 1$$

$$P(A_n \text{ a.a.n}) = 0 \text{ or } 1.$$

Thm (First Borel Cantelli lemma)

$$\underbrace{\sum_{n=1}^{\infty} P(A_n) < \infty}_{\text{if the } A_n\text{'s become sufficiently rare}} \Rightarrow \underbrace{P(A_n \text{ i.o.n}) = 0}_{\text{it is impossible for } A_n\text{'s to happen i.o.}}$$

Proof:

$$P(A_n \text{ i.o.n}) = P\Big(\bigcap_m \bigcup_{n \geq m} A_n\Big)$$

$$\leq P\Big(\bigcup_{n \geq m} A_n\Big), \quad \forall m$$

$$\leq \sum_{n=m}^{\infty} P(A_n), \quad \forall m$$

$$\longrightarrow 0 \quad \text{as } m \to \infty \text{ if } \sum_{n=1}^{\infty} P(A_n) < \infty$$

QED.

Warning: $P(A_n \text{ i.o.n}) = 0 \not\Rightarrow \sum_{n=1}^{\infty} P(A_n) < \infty$

e.g. $\Omega = (0,1]$

$A_n = (0, 1/n]$

P = uniform measure

$P(A_n \text{ i.o.n}) = 0$ but

$\sum P(A_n) = \infty$

If however the A_n's are independent then $P(A_n \text{ i.o.}) = 1$ or 0.

The contrapositive of the first Borel-Cantelli says

$$P(A_n \text{ i.o.n}) \neq 0 \Longrightarrow \sum P(A_n) = \infty$$

$$\Updownarrow \text{ indep}$$

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

The reverse implication is given by the next result.

Thm (Second Borel Cantelli lemma)

If $A_1, A_2, \ldots$ are independent then

$$\sum_{n=1}^{\infty} P(A_n) = \infty \iff P(A_n \text{ i.o.}_n) = 1$$

Proof: we just need to show $\Rightarrow$ by previous comments.

Suppose $\sum P(A_n) = \infty$.

Show $P(A_n^c \text{ a.a.}) = 0$.

$$P(A_n^c \text{ a.a.}_n) = P\Big(\bigcup_m \bigcap_{n \geq m} A_n^c\Big)$$

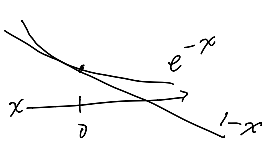
$$= P\Big(\underbrace{\Big(\bigcap_{n\geq 1} A_n^c\Big) \cup \Big(\bigcap_{n \geq 2} A_n^c\Big) \cup \cdots}_{\text{these grow}}\Big)$$

$$= P\Big(\lim_m \uparrow \bigcap_{n \geq m} A_n^c\Big)$$

$$= \lim_m P\Big(\bigcap_{n \geq m} A_n^c\Big)$$

$$= \lim_m \lim_p P\Big(\bigcap_{n \geq m}^{p} A_n^c\Big)$$

$$= \lim_m \lim_p \prod_{n \geq m}^{p} \underbrace{P(A_n^c)}_{= 1 - P(A_n) \leq e^{-P(A_n)}}$$



$$\leq \lim_m \lim_p \exp\Big(-\sum_{n \geq m}^{p} P(A_n)\Big)$$

$$= \lim_m \exp\Big(-\underbrace{\sum_{n \geq m}^{\infty} P(A_n)}_{-\infty}\Big)$$

$$= 0$$

QED

Restatement:

$$\sum P(A_n) < \infty \xRightarrow{\text{FBCL}} P(A_n \text{ i.o.}_n) = 0$$

If A_n's are indep then

$$\sum P(A_n) < \infty \xLeftrightarrow{\text{SBCL}} P(A_n \text{ i.o.}_n) = 0$$

Remark: Even though we haven't developed the notion of expected value yet it use full to understand that

$$\sum_n P(A_n) = \sum_n E(\underbrace{I_{A_n}(w)}_{\text{indicator of the event } A_n})$$

(expected value)

$$\overset{\text{Fubini}}{=} E\Big(\underbrace{\sum_n I_{A_n}(w)}_{\text{the number of times } A_n \text{ occurs for } w.}\Big)$$

Letting $N(w) = \sum_n I_{A_n}(w)$ we have

$$E(N) < \infty \implies P(A_n \text{ i.o}) = 0$$

$$E(N) = \infty \underset{\text{if the } A_n\text{'s are indep}}{\iff} P(A_n \text{ i.o.}) = 1$$

$\vdots$

19

Using the first Borel-Cantelli lemma for showing strong laws

The FBCL (first borell cantelli lemma) is useful for showing things like

$$P\left(\lim_n X_n = c \right) = 1$$

when you have bounds of the form

$$P\left(|X_n - c| \geq \varepsilon \right) \leq b(\varepsilon, n)$$

where $b(\varepsilon, n)$ has fast decay in n.

e.g. suppose $\exists\, \varepsilon_n \downarrow 0$ s.t. $\sum_{n=1}^{\infty} b(\varepsilon_n, n) < \infty$

$$\therefore \sum_{n=1}^{\infty} P(|X_n - c| \geq \varepsilon_n) < \infty$$

$$\therefore P\left(|X_n - c| \geq \varepsilon_n \text{ i.o.}_n \right) = 0 \quad \text{by FBCL}$$

$$\therefore P\Big(\underbrace{|X_n - c| < \varepsilon_n \text{ a.a.}_n}_{\substack{\text{imply that eventually} \\ |X_n - c| \to 0 \text{ at} \\ \text{rate } \leq \varepsilon_n}} \Big) = 1$$

$$\therefore 1 = P\left(|X_n - c| < \varepsilon_n \text{ a.a.}_n \right) \leq \underbrace{P\left(\lim_n X_n = c \right)}_{\text{so this is } 1.} \leq 1$$

20

Here is another way...

Suppose $\sum_{n=1}^{\infty} b(\varepsilon, n) < \infty \quad \forall \varepsilon > 0$

$$\therefore \sum_{n=1}^{\infty} P(|X_n - c| \geq \varepsilon) < \infty$$

$$\therefore P\left(|X_n - c| \geq \varepsilon \text{ i.o.}_n \right) = 0 \quad \forall \varepsilon$$

$$\therefore P\Big(\bigcup_{\varepsilon \in \mathbb{Q}^+} \{ |X_n - c| \geq \varepsilon \text{ i.o.}_n \} \Big) = 0$$

by subadditivity

$$\therefore P\Big(\underbrace{\bigcap_{\varepsilon \in \mathbb{Q}^+} \{ |X_n - c| < \varepsilon \text{ a.a.}_n \}}_{\text{equals the event } \{X_n \to c\}} \Big) = 1$$

Remark: the above two arguments do not require independence of the X_n's.

Thm Fatou's lemma

$$P(A_n \text{ a.a.n}) \le \liminf_n P(A_n)$$

$$\le \limsup_n P(A_n)$$

Note: for measures you don't have this inequality always

$$\le P(A_n \text{ i.o.n}).$$

Proof:

$$P(A_n \text{ a.a.n}) = P\Big(\lim_m \uparrow \bigcap_{n\ge m} A_n\Big)$$

$$= \lim_m P\Big(\underbrace{\bigcap_{n\ge m} A_n}_{\le P(A_n),\ \forall n \ge m}\Big)$$

$$\le \lim_m \inf_{n\ge m} P(A_n)$$

$$= \liminf_n P(A_n)$$

Now take complements for the other inequality.

$$\limsup_n P(A_n) \le P(A_n \text{ i.o.n})$$

$$\Updownarrow$$

$$\limsup_n (1 - P(A_n^c)) \le 1 - P(A_n^c \text{ a.a.n})$$

$$\Updownarrow$$

$$1 - \liminf_n P(A_n^c) \le 1 - P(A_n^c \text{ a.a.n})$$

holds by first inequality QED

e.g. $(\Omega, \mathcal{F}, P) \equiv$ uniform prob measure on $[0,1]$.

$$A_n = \begin{cases} [0, \tfrac{1}{3}] & \text{if } n \text{ is even} \\ (\tfrac{1}{3}, 1] & \text{if } n \text{ is odd} \end{cases}$$

Fatou gives:

$$\underset{0}{P(A_n \text{ a.a.})} \le \underset{\frac{1}{3}}{\liminf P(A_n)}$$

$$\le \underset{\frac{2}{3}}{\limsup P(A_n)}$$

$$\le \underset{1}{P(A_n \text{ i.o.})}$$

SLLN $\Rightarrow$ WLLN via Fatou

WLLN: $\frac{S_n}{n} \xrightarrow{P} 0$ means

$$\forall \varepsilon > 0, \ P\left(\left|\tfrac{S_n}{n}\right| \ge \varepsilon\right) \to 0 \text{ as } n \to \infty.$$

SLLN: $\frac{S_n}{n} \xrightarrow{a.e.} 0$ means

$$P\left(\tfrac{S_n}{n} \not\to 0\right) = 0$$

Fatou gives

$$\limsup_n P\left(\left|\tfrac{S_n}{n}\right| \ge \varepsilon\right) \overset{\text{Fatou}}{\le} P\left(\left|\tfrac{S_n}{n}\right| \ge \varepsilon \text{ i.o.n}\right)$$

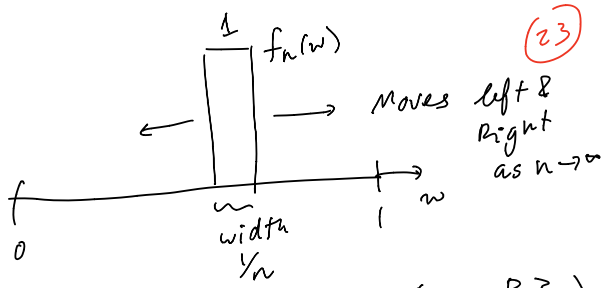
$$\le P\left(\tfrac{S_n}{n} \not\to 0\right)$$

$$= 0 \text{ by SLLN}$$

$\therefore$ SLLN $\Rightarrow$ WLLN

However WLLN $\not\Rightarrow$ SLLN

The classic counter example is the moving spike.



$$P(|f_n| \geq \varepsilon) = \frac{1}{n} \to 0 \quad \left(\text{where } P \text{ is our uniform measure}\right)$$

So the WLLN holds but

$$P(f_n \not\to 0) = 1$$

So SLLN does not.

Erdős & Renyi's extension of the SBCL

claim: If $\sum P(A_n) = \infty$ and

$$\liminf_{n\to\infty} \frac{\sum_{kj}^{n} P(A_k \cap A_j)}{\sum_{kj}^{n} P(A_k)P(A_j)} \leq 1$$

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

Proof in the special case that $(\Omega, \mathcal{F}, P)$ = uniform probability measure & $A_1, A_2, \ldots \in \mathcal{B}_0((0,1])$. ← because we haven't developed integration yet.

For any $Y(w) = \sum_{k=1}^{n} c_k I_{A_k}(w)$ define

$$EY = \int_0^1 Y(w)\,dw \quad \leftarrow \text{Rieman integral}$$

Set $X_n(w) := \sum_{k=1}^{n} I_{A_k}(w)$ & Notice

$$EX_n = \sum_{k=1}^{n} E(I_{A_k}) = \sum_{k=1}^{n} \int_{A_k} 1\,dw = \sum_{k=1}^{n} P(A_k)$$

Notice

$$\begin{aligned} E(X_n - EX_n)^2 &= \int (X_n(w) - EX_n)^2\,dw \\ &= E(X_n^2) - (EX_n)^2 \\ &= E\Big(\sum_{k,j=1}^{n} I_{A_k\cap A_j}\Big) - \Big(\sum_{k=1}^{n} P(A_k)\Big)^2 \\ &= \Bigg(\underbrace{\frac{\sum_{k,j=1}^{n} P(A_k\cap A_j)}{\big(\sum_{k=1}^{n} P(A_k)\big)^2}}_{:=\theta_n} - 1\Bigg)\Big(\underbrace{\sum_{k=1}^{n} P(A_k)}_{m_n}\Big)^2 \end{aligned}$$

Also

$$\begin{aligned} P(X_n \leq x) &= P(EX_n - X_n \geq EX_n - x) \\ &\leq P\big((EX_n - X_n)^2 \geq (EX_n - x)^2\big) \\ &= \int_{\{(EX_n - X_n)^2 \geq (EX_n - x)^2\}} 1\,dw \\ &\leq \int \frac{(EX_n - X_n(w))^2}{(EX_n - x)^2}\,dw \\ &= \underbrace{(\theta_n - 1)}_{\text{liminfs to} \leq 0 \text{ as } n\to\infty} \underbrace{\frac{m_n^2}{(m_n - x)^2}}_{\to 1} \quad \text{if } x < m_n \end{aligned}$$

If $x < EX_n$ which always happens for large enough n since $EX_n = \sum_{k=1}^{n} P(A_k) \to \infty$ by assumption

$\to 1$, since $m_n = \sum_{k=1}^{n} P(A_k) \to \infty$

$\therefore \liminf_{n\to\infty} P(X_n \leq x) = 0$

$\therefore P(X_n \leq x \text{ a.a.n}) = 0$ by Fatou

$\therefore P\Big(\bigcup_{x=1}^{\infty} \{X_n \leq x \text{ a.a.n}\}\Big) = 0$ by subadditivity

$\therefore P\Big(\bigcap_{x=1}^{\infty} \{X_n > x \text{ i.o.n}\}\Big) = 1$

$\therefore P\big(\limsup_{n\to\infty} X_n = \infty\big) = 1$ ↘ $\sum_{k}^{n} I_{A_k}$

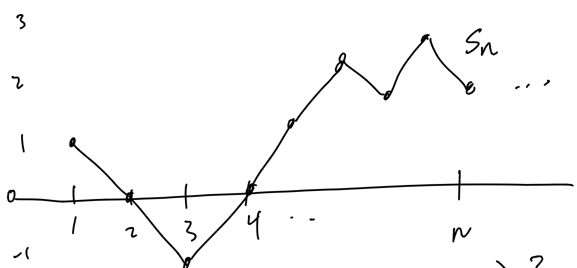
$\therefore P(A_n \text{ i.o.n}) = 1$ QED

Remark: there is a nice example of the use of this result for probing "runs" of coin flips in Billingsley p. 89

Hewitt-Savage 0-1 law for coin flips

when $R_1, R_2, \ldots$ represent the Rahdmacher R.V.s $R_k = \begin{cases} 1 & \text{w.p. } 1/2 \\ -1 & \text{w.p. } 1/2 \end{cases}$

from lecture 1, $S_n = \sum_{k=1}^{n} R_k$ represents a 1-d random walk:



Question: what is $P(S_n = 0 \text{ i.o.} n)$?

i.e. what is the chance the random walk returns to zero - infinitely often?

Note that Kolmogorov's 0-1 law doesn't apply here since technically checking $w \in \{S_n = 0 \text{ i.o.} n\}$ depends on the value of $X_1(w)$.

We will prove $P(S_n = 0 \text{ i.o.} n)$ is 0 or 1 by essentially proving a special case of Hewitt-Savage 0-1 law which applies to symmetric functions of exchangable random variables.

Suppose $\pi : \mathbb{N} \to \mathbb{N}$ denotes a permutation of the positive integers which permutes at most finitely many numbers.

e.g. $\pi(1) = 4$
$\pi(2) = 3$
$\pi(3) = 1$
$\pi(4) = 2$
$\pi(k) = k \quad \forall k > 4.$

Let $S_n^{\pi} = \sum_{k=1}^{n} R_{\pi(k)}$ which $= S_n$ when n is large enough

Notice 2 key facts:

(i) $\{S_n = 0 \text{ i.o.} n\} = \{S_n^{\pi} = 0 \text{ i.o.} n\}$

$S_n = S_n^{\pi}$ for large enough n.

(ii) Any probability calculated for $(S_1, S_2, \ldots)$ is the same as for $(S_1^{\pi}, S_2^{\pi}, \ldots)$.

Now fix $\varepsilon > 0$ & since

$$\{S_n = 0 \text{ i.o.} n\} = \bigcap_{m} \bigcup_{n \geq m} \{S_n = 0\}$$

$$= \lim_{m} \downarrow \bigcup_{n \geq m} \{S_n = 0\}$$

$\therefore \exists m_\varepsilon$ s.t.

$$P\Big(\bigcup_{n \geq m_\varepsilon} \{S_n = 0\} - \{S_n = 0 \text{ i.o.} n\}\Big) \leq \frac{\varepsilon}{2}$$

& $\exists n_\varepsilon$ s.t.

$$P\Big(\bigcup_{n \geq m_\varepsilon} \{S_n = 0\} - \bigcup_{n = m_\varepsilon}^{n_\varepsilon} \{S_n = 0\}\Big) \leq \frac{\varepsilon}{2}$$

$$\therefore P\Big(\{S_n = 0 \text{ i.o.} n\} \,\Delta \bigcup_{n = m_\varepsilon}^{n_\varepsilon} \{S_n = 0\}\Big) \leq \varepsilon$$

$$\therefore\ P\left(\{S_n=0 \text{ i.o.n}\} \,\Delta \bigcup_{n=m_\varepsilon}^{n_\varepsilon} \{S_n=0\}\right) \le \varepsilon$$

$$\| \text{ by (ii)}$$

$$P\left(\{S_n^\pi=0 \text{ i.o.n}\} \,\Delta \bigcup_{n=m_\varepsilon}^{n_\varepsilon} \{S_n^\pi=0\}\right)$$

$$\| \text{ by (i)}$$

$$P\left(\{S_n=0 \text{ i.o.n}\} \,\Delta \bigcup_{n=m_\varepsilon}^{n_\varepsilon} \{S_n^\pi=0\}\right)$$

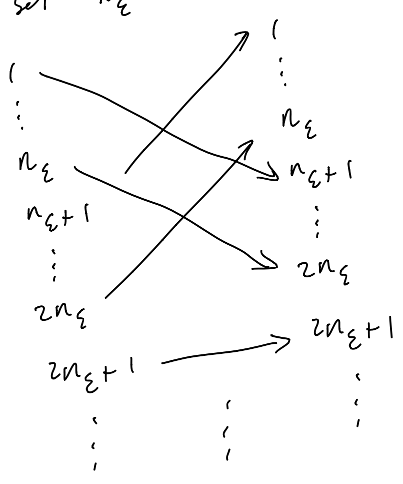
$\forall \pi$ that permutes finitely many indices.

Let $A = \{S_n = 0 \text{ i.o.n}\}$

$$A_\varepsilon = \bigcup_{n=m_\varepsilon}^{n_\varepsilon} \{S_n=0\}$$

$$A_\varepsilon^\pi = \bigcup_{n=m_\varepsilon}^{n_\varepsilon} \{S_n^\pi=0\}$$

If we set π_ε as



Then A_ε is indep of $A_\varepsilon^{\pi_\varepsilon}$

In summary

(iii) $P(A \,\Delta\, A_\varepsilon) = P(A \,\Delta\, A_\varepsilon^{\pi_\varepsilon}) \le \varepsilon$

(iv) $P(A_\varepsilon \cap A_\varepsilon^{\pi_\varepsilon}) = P(A_\varepsilon)P(A_\varepsilon^{\pi_\varepsilon})$

Now

$$P((A\cap A) \,\Delta\, (A_\varepsilon \cap A_\varepsilon^{\pi_\varepsilon}))$$
$$\le P(A \,\Delta\, A_\varepsilon) + P(A \,\Delta\, A_\varepsilon^{\pi_\varepsilon})$$
$$\longrightarrow 0 \quad \text{as } \varepsilon \to 0 \quad \text{from (iii)}$$

$\therefore$ an exercise shows

$$P(A_\varepsilon \cap A_\varepsilon^{\pi_\varepsilon}) \longrightarrow P(A\cap A) = P(A)$$
$$\|$$
$$P(A_\varepsilon)P(A_\varepsilon^{\pi_\varepsilon}) \longrightarrow P(A)P(A)$$

$\therefore P(A\cap A) = P(A)P(A)$.

i.e. $P(A) = 0$ or 1.

Remark: we will see later that for random walks in dimension 1 & 2, $P(S_n=0 \text{ i.o.n}) = 1$, but when dimension ≥ 3, $P(S_n = 0 \text{ i.o.n}) = 0$.

"A drunk man will find his way home eventually but a drunk bird may get lost forever."

Remark: The above argument can be extended to "exchangable" r.v.s $R_1, R_2, \ldots$