

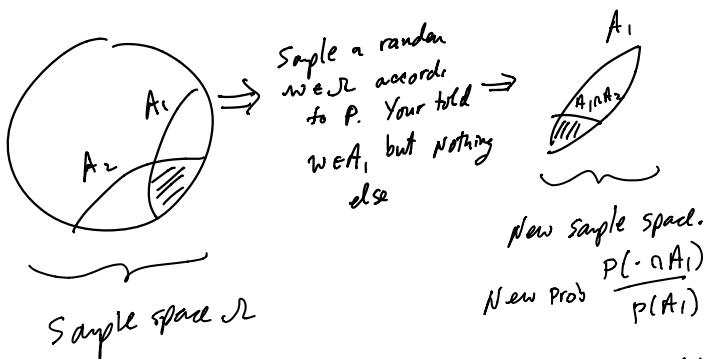
## Lecture 6: Independence

(1)

Let  $(\Omega, \mathcal{F}, P)$  be a probability space  
 sample space  $\Omega$   $\uparrow$   $\sigma$ -field  $\mathcal{F}$  prob measure

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

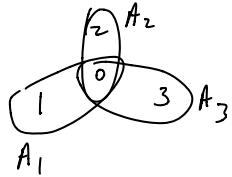
$$P(A_2 | A_1) = \frac{\text{prob of } A_2 \text{ given } A_1}{\text{given } A_1} := \frac{P(A_1 \cap A_2)}{P(A_1)}$$



$A_1$  is independent of  $A_2$  if  $P(A_2 | A_1) = P(A_2)$ .  
 i.e. if  $P(A_1 \cap A_2) = P(A_1)P(A_2)$ .

Question: How to make sense of independent among a collection of events (possibly uncountably many)? Is pairwise independent enough?

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



$$i \neq j \Rightarrow P(A_i \cap A_j) = P(A_i)P(A_j) \\ = \underbrace{\{0\}}_{Y_1} \underbrace{\{1\}}_{Y_2} \underbrace{\{2\}}_{Y_3} \\ = \frac{1}{4}$$

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

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

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

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

e.g. let  $A_1, \dots, A_n$  be events (i.e.  $A_i \in \mathcal{F}$ ) (2)

s.t.  $A_1 = \emptyset$ . Then

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

so the full factorization criterian will not work as a def of independent either

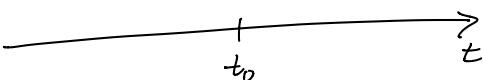
Here is the "right" def of indep for a collection of events.

Def: A collection of events  $\{A_h\}_{h \in K}$  are independent events iff  $\forall$  finite  $H \subset K$

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

Note:  $K$  is allowed to be any index set.

We will also need the notion of independent  $\sigma$ -fields to make sense of things like the strong markov property of Brownian motion  $B_t$ :



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

Def: Let  $K$  be an arbitrary index set.  $\forall k \in K$ , let  $\mathcal{A}_k$  be a collection of events.

The  $\mathcal{A}_k$ 's are independent collections if  $\{A_k\}_{k \in K}$  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  $k \in K$  (arb index set). Then (3)

(i) (subclasses):

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

(ii) (augmentation):

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

(iii) (simplified product):

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

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

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

Proof:

(i): trivial

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

For  $\Rightarrow$  choose  $A_k \in \mathcal{A}_k \cup \{\mathcal{D}\}$  & finite  $K \subset K$ . Let  $K_0 = \{k : A_k \in \mathcal{A}_k\}$ .

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

$$= P\left(\bigcap_{h \in K \setminus K_0} A_h\right), A_h = \mathcal{D} \text{ when } h \in K - K_0$$

$$= \prod_{h \in K \setminus K_0} P(A_h), \mathcal{A}_k \text{'s indep}$$

$$= \prod_{h \in K} P(A_h), P(A_h) = P(\mathcal{D}) = 1 \text{ when } h \in K - K_0$$

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

(iii)  $P\left(\bigcap_{k=1}^n A_k\right) = \prod_{k=1}^n P(A_k)$  (4)

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

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

QED

e.g. Coin flip Model from lecture 1:

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

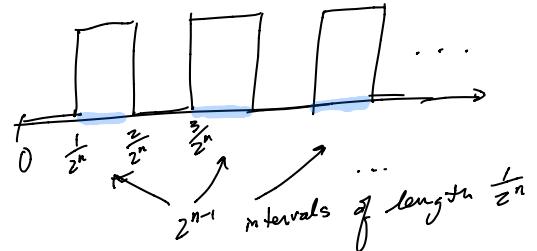
$X_k(w) := k^{\text{th}}$  binary digit of  $w = \underbrace{\square \square \square \square}_{\text{...}} \xrightarrow[w]{\text{---}} \underbrace{(\frac{i-1}{2^n}, \frac{i}{2^n}]}$

$$H_k := \{w \in \mathcal{D} : X_k(w) = 1\}$$

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

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

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



If  $m < n$  then  $H_m$  looks like



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

Let  $1 \leq i_1 < i_2 < \dots < i_m$  & show

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

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

$\uparrow$        $\uparrow$        $\uparrow$   
 $2^{in-1}$  intervals of length  $\frac{1}{2^{in}}$  reduce the # of intervals by  $\frac{1}{2}$  for each further intersection

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

as was to be shown QED.

$\pi$ -generators are enough & ANOVA

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

The following thm helps this.

Thm ( $\pi$ -generators are enough):

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

$\mathcal{Q}_k$ 's are indep  $\pi$ -systems

$\Rightarrow \sigma(\mathcal{Q}_k)$ 's are independent

Proof: Let  $B_k := \mathcal{Q}_k \cup \{\emptyset\}$

Suppose the  $\mathcal{Q}_k$ 's are indep  $\pi$ -sys

$\therefore$  the  $B_k$ 's are indep  $\pi$ -sys, by augmentation

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

the  $B_{k_1}, B_{k_2}, \dots, B_{k_n}$  are indep  $\pi$ -sys

Show  $\sigma(B_{k_1}), B_{k_2}, \dots, B_{k_n}$  are indep  $\pi$ -sys

and we will be done (by induction)

By the simplified product criterion (6)  
this is equivalent to showing

$$P(B_{i_1} \cap B_{i_{n-1}} \cap \dots \cap B_{i_1}) = P(B_{i_1}) \dots P(B_{i_n}) \quad (*)$$

$$\forall B_i \in \sigma(B_{k_i}), B_2 \in B_{k_2}, \dots, B_n \in B_{k_n}$$

Fixing  $B_2, \dots, B_n$  let

$$\mathcal{Y} := \{B_i \in \mathcal{F} : (*) \text{ holds}\}$$

& show  $\sigma(B_{k_1}) \subset \mathcal{Y}$ .

$\bullet B_{k_1} \in \mathcal{Y}$ : yes, since  $B_{k_1}$ 's are indep.

$\bullet \emptyset \in \mathcal{Y}$ : yes, since  $\emptyset \in B_{k_1} \quad \forall k \in K$ .

$\bullet B \in \mathcal{Y} \Rightarrow$

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

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



$$= P(\emptyset \cap B_2 \cap \dots) - P(B \cap B_2 \cap \dots)$$

$$= P(\emptyset) \cdot P(B_2) \dots - P(B) P(B_2) \dots$$

since  $\emptyset, B \in \mathcal{Y}$

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

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

$\bullet \underbrace{A_1, A_2, \dots}_{\text{disjoint}} \in \mathcal{Y}$

disjoint

$$\Rightarrow P((\bigcup_{k_1} A_{k_1}) \cap B_2 \cap \dots \cap B_n)$$

$$= \sum_k P(A_{k_1} \cap B_2 \cap \dots \cap B_n)$$

$$= \sum_k P(A_{k_1}) P(B_2) \dots P(B_n)$$

$$= P(B_2) \dots P(B_n) \left[ \sum_k P(A_{k_1}) \right]$$

$$\Rightarrow \bigcup_k A_{k_1} \in \mathcal{Y}$$

$$P(\bigcup_k A_{k_1})$$

$\mathcal{M}$  is a  $\lambda$ -sys &  $B_k$ , is a  $\pi$ -sys. ⑦

Thm (ANOVA): Matrix of  $\pi$ -systems (8)

$\sigma\langle B_k \rangle \subset \mathcal{M}$ . QED.

e.g. coin flip example showed  
 $H_1, H_2, \dots$  are indep

since  $\{H_p\}$  is  $\pi$ -sys for each  $p$ ,

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

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

$\therefore$  Any sequence  $H_1, H_2^c, H_3, H_4^c, H_5, \dots$   
 are indep.  
↑  
tails  
in the  
n-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  $B$ ?

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

Let  $R_i = \underbrace{\sigma\langle \mathcal{O}_{i1}, \mathcal{O}_{i2}, \dots \rangle}_{i\text{-th row}}$

Then

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

Proof:

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

$\therefore$  (ii) clearly holds

To show (i) note

$$R_p = \sigma\langle \mathcal{O}_{p1} \cup \mathcal{O}_{p2} \cup \dots \rangle = \sigma\langle P_p \rangle$$

would like to  
use  $\pi$ -generators  
but this isn't a  
 $\pi$ -sys

where  $P_p =$  the closure of  $\mathcal{O}_{p1} \cup \mathcal{O}_{p2} \cup \dots$   
 under finite intersection

Clearly  $P_p$ 's are  $\pi$ -systems.

Let's show the  $P_p$ 's are indep.

Select one  $P_{k_i}$  from  $P_k$  and note:

(9)

$$P_{k_1} \cap \dots \cap P_{k_n}$$

$\underbrace{\phantom{P_{k_1} \cap \dots \cap P_{k_n}}}_{\text{Row } k_1}$

$\underbrace{\phantom{P_{k_1} \cap \dots \cap P_{k_n}}}_{\text{Row } k_2}$

$\underbrace{\phantom{P_{k_1} \cap \dots \cap P_{k_n}}}_{\text{Row } k_3}$

$\dots$

Write this as  $(A_1, \dots) \cap (B_1, \dots) \cap (C_1, \dots) \cap \dots$

$P_{k_1}$

$\underbrace{\phantom{P_{k_1}}}_{\text{each event in here is from}}$

a unique  $\mathcal{A}_{k_{i,j}}$

merging (via "n") multiple sets from the same  $\mathcal{A}_k$ ,  
if necessary ... still a  $\mathcal{A}_k$  set by  $\pi$ -sys assumption

Now,

$$\begin{aligned} P(P_{k_1} \cap \dots \cap P_{k_n}) &= P(A_1) \dots P(B_1) \dots P(C_1) \dots \\ &= P(P_{k_1}) \dots P(P_{k_n}) \end{aligned}$$

e.g.

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

$\underbrace{\phantom{A_1 \cap A_2}}_{\in \mathcal{A}_{k_{1,1}}} \quad \underbrace{\phantom{B_1 \cap B_2}}_{\in \mathcal{A}_{k_{2,2}}}$

both in  $\mathcal{A}_{k_{1,1}} \mathcal{A}_{k_{2,2}}$

$\mathcal{A}_{k_{1,1}} \quad \mathcal{A}_{k_{2,2}}$

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

$\underbrace{\phantom{A_1 \cap A_2}}_{\in \mathcal{A}_{k_{1,1}} \text{ by}} \quad \text{since } \mathcal{A}_{k_{1,1}} \text{ &} \mathcal{A}_{k_{2,1}} \text{ & } \mathcal{A}_{k_{2,2}} \text{ are}$

$\pi\text{-sys}$

$\text{indep.}$

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

$\text{since } \mathcal{A}_{k_{2,1}} \text{ & } \mathcal{A}_{k_{2,2}}$

$\text{are indep.}$

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

( $\Leftarrow$ )

(10)

Suppose the row  $\sigma$ -fields  $R_k$  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  $(i, b) \in \mathcal{H}$  choose one  
 $A_{ik} \in \mathcal{A}_{ik} \subseteq R_i$

$$\therefore P\left(\bigcap_{(i, b) \in \mathcal{H}} A_{ik}\right) = P\left(\bigcap_{\substack{\text{rows } i \\ \text{in } \mathcal{H}}} \bigcap_{\substack{\text{cols } k \\ \text{s.t. } (i, b) \in \mathcal{H}}} A_{ik}\right)$$

$$\stackrel{R_i \text{ is indep}}{=} \prod_{\substack{\text{rows } i \\ \text{in } \mathcal{H}}} P\left(\bigcap_{\substack{\text{cols } k \\ \text{s.t. } (i, b) \in \mathcal{H}}} A_{ik}\right)$$

$$\stackrel{\text{w.r.t. rows indep}}{=} \prod_{\substack{\text{rows } i \\ \text{in } \mathcal{H}}} \prod_{\substack{\text{cols } k \\ \text{s.t. } (i, b) \in \mathcal{H}}} P(A_{ik})$$

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

QED

$\therefore P_k$ 's are indep  $\pi$ -sys.

$\therefore$  The  $\sigma$ -fields  $R_k := \sigma(P_k)$  are independent by  $\pi$ -generators.

## Kolmogorov's 0-1 law

(11)

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

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

$$\begin{aligned}\Sigma &:= \bigcap_{m=1}^{\infty} \sigma(\mathcal{Q}_m, \mathcal{Q}_{m+1}, \dots) \\ &= \left\{ A \in \mathcal{F} : A \in \sigma(\mathcal{Q}_m, \mathcal{Q}_{m+1}, \dots) \text{ for all } m \right\}\end{aligned}$$

( $\Sigma$  is a  $\sigma$ -field for the same reason  $\sigma(\mathcal{C})$  is)

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

$$\begin{aligned}\frac{S_n}{n} \rightarrow 0 &\iff X_1 + \dots + X_n \xrightarrow{n} \frac{1}{2} \\ &\iff \frac{X_1 + \dots + X_{m-1} + X_m + \dots + X_n}{n} \xrightarrow{n} \frac{1}{2}\end{aligned}$$

$$\iff \frac{X_m + \dots + X_n}{n} \xrightarrow{n} \frac{1}{2} \text{ s.t.}$$

where  $\left\{ \frac{X_m + \dots + X_n}{n} \xrightarrow{n \rightarrow \infty} \frac{1}{2} \right\} \in \sigma(H_m, H_{m+1}, \dots)$

$\therefore N = \left\{ \frac{S_n}{n} \rightarrow 0 \right\} \in \text{tail } \sigma\text{-field}$  generated by  $H_1, H_2, \dots$

Thm (Kolmogorov's 0-1 law)

If  $\mathcal{Q}_1, \mathcal{Q}_2, \dots$  are indep  $\pi$ -systems then  $\forall A \in \Sigma, P(A) = 0$  or  $P(A) = 1$ .

tail  $\sigma$ -field generated by the  $\mathcal{Q}_k$ 's

Proof: show  $A$  is independent of itself.

$\mathcal{Q}_1, \dots, \mathcal{Q}_{m-1}, \mathcal{Q}_m, \mathcal{Q}_{m+1}, \dots$  are indep  $\pi$ -sys.

$\therefore \sigma(\mathcal{Q}_1), \dots, \sigma(\mathcal{Q}_{m-1}), \sigma(\mathcal{Q}_m, \mathcal{Q}_{m+1}, \dots)$  are indep  $\pi$ -sys by above.

$\therefore \sigma(\mathcal{Q}_1), \dots, \sigma(\mathcal{Q}_{m-1}), \Sigma$  are indep  $\pi$ -sys holds for all  $m$  by subclasses.

$\therefore \sigma(\mathcal{Q}_1), \sigma(\mathcal{Q}_2), \dots, \Sigma$  are indep  $\pi$ -sys by the finite selection requirement of the def of indep.

$\therefore \sigma(\mathcal{Q}_1, \mathcal{Q}_2, \dots), \Sigma$  are indep  $\pi$ -sys by above.

$\therefore \Sigma, \Sigma$  are indep  $\pi$ -sys by subclasses

$\therefore \forall A \in \Sigma, P(A \cap A) = P(A)P(A)$

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

QED

## Borel-Cantelli and Fatou

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

Def:

$$\{A_n \text{ i.o.}\} := \left\{ w \in \Omega : w \in A_n \text{ infinitely often in } n \right\}$$

$$:= \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} A_n$$

$\forall m \exists n \geq m \text{ s.t. } w \in A_n$ .

$$\{A_n \text{ a.a.}\} := \left\{ w \in \Omega : w \in A_n \text{ for all but finitely many } n \right\}$$

$$:= \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} A_n$$

$\exists n \text{ s.t. } \forall n \geq m, w \in A_n$

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

(13)

Sometimes people write

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

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

since indicator of  $A_n$

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

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

Some Facts:

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

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

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

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

these grow since you're 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  $\sigma$ -field generated by  $\{A_1\}, \{A_2\}, \dots$

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

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

$= A_m$

(14)

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

$$= \left( \bigcup_{n \geq 1} A_n \right) \cap \left( \bigcup_{n \geq 2} A_n \right) \cap \dots$$

these decrease as sets

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

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

$\in$  tail  $\sigma$ -field generated  
by  $\{A_1\}, \{A_2\}, \dots$

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

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

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

Note: The 0-1 law already implies

$A_1, A_2, \dots \in \mathcal{F}$  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)

$$\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(A_n \text{ i.o.n}) = 0$$

$\curvearrowleft$  if the  $A_n$ 's become sufficiently rare

$\curvearrowleft$  if it is impossible for  $A_n$ 's to happen i.o.

Proof:

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

$$\leq P\left(\bigcup_{n \geq m} A_n\right), \text{ fm}$$

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

$$\rightarrow 0 \text{ as } m \rightarrow \infty$$

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

QED.

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

e.g.  $\mathcal{I}_2 = [0, 1]$

$A_n = [0, \frac{1}{n}]$

$P$  = 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

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

$\Updownarrow$  indep

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

The reverse implication is given by the  
next result.

## Thm (Second Borel Cantelli lemma) (17)

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

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

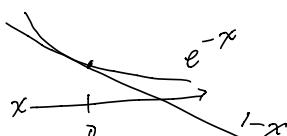
Proof: We just need to show  $\rightarrow$   
by previous comments.

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

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

$$\begin{aligned} P(A_n^c \text{ a.a.}) &= P\left(\bigcup_m \bigcap_{n \geq m} A_n^c\right) \\ &= P\left(\left(\bigcap_{n \geq 1} A_n^c\right) \cup \left(\bigcap_{n \geq 2} A_n^c\right) \cup \dots\right) \\ &\quad \xrightarrow{\text{these grow}} \\ &= P\left(\limsup_m \bigcap_{n \geq m} A_n^c\right) \end{aligned}$$

$$\begin{aligned} &= \lim_m P\left(\bigcap_{n \geq m} A_n^c\right) \\ &= \lim_m \lim_p P\left(\bigcap_{n \geq m} A_n^c\right) \\ &= \lim_m \lim_p \prod_{n \geq m}^p P(A_n^c) \\ &\quad \xrightarrow{\text{if the } A_n^c \text{ are indep}} \\ &= \lim_m \lim_p \prod_{n \geq m}^p (1 - P(A_n)) \\ &= \lim_m \lim_p e^{-\sum_{n \geq m}^p P(A_n)} \\ &\leq e^{-\sum_{n \geq m}^p P(A_n)} \end{aligned}$$



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

$$\begin{aligned} &= \lim_m \exp\left(-\sum_{n \geq m}^{\infty} P(A_n)\right) \\ &\xrightarrow{-\infty} 0 \quad \text{QED} \end{aligned}$$

Restatement:

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

If  $A_n$ 's are indep true

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

(18)

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

$$\begin{aligned} \sum_n P(A_n) &= \sum_n E(\underbrace{I_{A_n}(w)}) \\ &\quad \xrightarrow{\text{Indicator of the event } A_n} \\ &= E\left(\underbrace{\sum_n I_{A_n}(w)}_{\substack{\text{number of times} \\ A_n \text{ occurs for } w.}}\right) \end{aligned}$$

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

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

$$E(N) = \infty \iff P(A_n \text{ i.o.}) = 1$$

if the  $A_n$ 's  
are indep

:

(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

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

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

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

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

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

$$\therefore P\left(|X_n - c| < \varepsilon_n \text{ a.a.n}\right) = 1$$

*Imply that eventually*

$|X_n - c| \rightarrow 0$  at  
rate  $\leq \varepsilon_n$

$$\therefore 1 = P(|X_n - c| < \varepsilon_n \text{ a.a.n})$$

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

*So this is 1.*

Here is another way ...

(20)

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

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

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

$$\therefore P\left(\bigcup_{\varepsilon \in \mathbb{R}^+} \{|X_n - c| \geq \varepsilon \text{ i.o.n.}\}\right) = 0$$

by subadditivity

$$\therefore P\left(\bigcap_{\varepsilon \in \mathbb{R}^+} \{|X_n - c| < \varepsilon \text{ a.a.n}\}\right) = 1$$

*→*

equals the event  $\{X_n \rightarrow c\}$

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

## Thm Fatou's lemma (21)

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

$$\leq \limsup_n P(A_n)$$

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

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

Proof:

$$\begin{aligned} P(A_n \text{ a.a.n}) &= P\left(\lim_{m \rightarrow \infty} \bigcap_{n \geq m} A_n\right) \\ &= \lim_m P\left(\bigcap_{n \geq m} A_n\right) \\ &\leq P(A_n), \forall n \geq m \end{aligned}$$

$$\leq \lim_m \inf_{n \geq m} P(A_n)$$

$$= \liminf_n P(A_n)$$

Now take complements for the other inequality.

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

↑

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

↑

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

↑ holds by first inequality QED

e.g.  $(\Omega, \mathcal{F}, P) = \text{uniform}$

(22)

prob measure on  $[0,1]$ .

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

Fatou gives:

$$\begin{aligned} P(A_n \text{ a.a.}) &\leq \liminf P(A_n) \\ &\leq \limsup P(A_n) \\ &\leq P(A_n \text{ i.o.}) \end{aligned}$$

SLLN  $\Rightarrow$  WLLN via Fatou

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

$$\forall \varepsilon > 0, P\left(\left|\frac{s_n}{n}\right| \geq \varepsilon\right) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

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

$$P\left(\frac{s_n}{n} \not\rightarrow 0\right) = 0$$

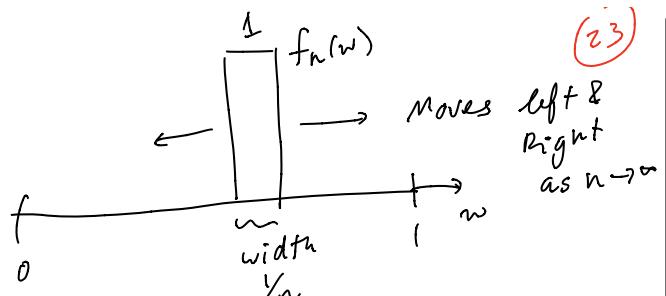
Fatou gives

$$\begin{aligned} \limsup_n P\left(\left|\frac{s_n}{n}\right| \geq \varepsilon\right) &\stackrel{\text{Fatou}}{\leq} P\left(\left|\frac{s_n}{n}\right| \geq \varepsilon \text{ i.o.n}\right) \\ &\leq P\left(\frac{s_n}{n} \not\rightarrow 0\right) \\ &= 0 \text{ by SLLN} \end{aligned}$$

$\therefore \text{SLLN} \Rightarrow \text{WLLN}$

However  $\text{WLLN} \not\Rightarrow \text{SLLN}$

The classic counter example is the moving spike.



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

So the WLLN holds but

$$P(f_n \rightarrow 0) = 1$$

So SLLN does not.

Erdős & Renyi's extension  
of the SLLN

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

$$\liminf_{n \rightarrow \infty} \frac{\sum_{k,j} P(A_k \cap A_j)}{\sum_{k,j} P(A_k)P(A_j)} \leq 1$$

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

Proof in the special case that

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

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

$$EY = \int_0^1 \gamma(w) dw \quad \text{Riemann integral}$$

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

$$\begin{aligned} E(X_n - EX_n)^2 &= \int (X_n(w) - EX_n)^2 dw \\ &= E(X_n^2) - (EX_n)^2 \\ &= E\left(\sum_{k,j=1}^n I_{A_k \cap A_j}\right) - \left(\sum_{k=1}^n P(A_k)\right)^2 \\ &= \left( \underbrace{\frac{\sum_{k,j=1}^n P(A_k \cap A_j)}{\left(\sum_{k=1}^n P(A_k)\right)^2}}_{-1} \right) \left( \underbrace{\sum_{k=1}^n P(A_k)}_{M_n} \right)^2 \end{aligned}$$

Also  $: \theta_n$

$$P(X_n \leq x) = P(EX_n - X_n \geq EX_n - x)$$

$$\leq P((EX_n - X_n)^2 \geq (EX_n - x)^2)$$

$$= \int 1 dw$$

If  $x < EX_n$   
which always happens for large enough  $n$  since

$$\begin{aligned} EX_n &= \sum_{n=1}^n P(A_n) \rightarrow \infty &\leq \int \frac{(EX_n - X_n(w))^2}{(EX_n - x)^2} dw \\ &\text{by assumption} &= \left(\theta_n^{-1}\right) \frac{M_n^2}{(M_n - x)^2} &\text{if } x < M_n \\ &&\text{limits} &\rightarrow 1, \text{ since } \\ &&\text{to } \leq 0 \text{ as } & M_n = \sum_{k=1}^n P(A_k) \rightarrow \infty \\ &&n \rightarrow \infty & \end{aligned}$$

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

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

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

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

$$\therefore P(\limsup_{n \rightarrow \infty} X_n = \infty) = 1$$

$$\therefore P(A_n \text{ i.o.n}) = 1 \quad \sum_{k=1}^n I_{A_k}$$

QED

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

(25)

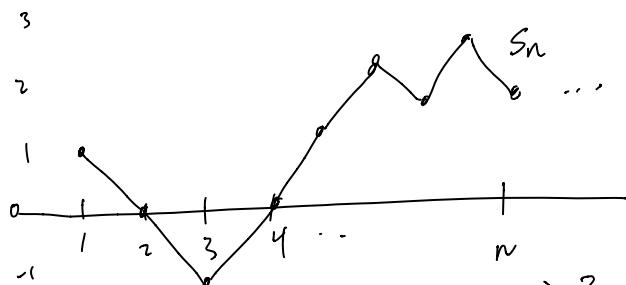
Hewitt-Savage 0-1 law  
for coin flips

when  $R_1, R_2, \dots$  represent the

Rademacher R.V.s  $R_k = \begin{cases} 1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{1}{2} \end{cases}$

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  $\{S_n = 0 \text{ i.o.n}\}$  depends on the value of  $X_i(w)$ .

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 exchangeable random variables.

Suppose  $\pi: \mathbb{N} \rightarrow \mathbb{N}$  denotes a permutation of the positive integers which permutes at most finitely many numbers. (26)

$$\text{e.g. } \pi(1) = 4$$

$$\pi(2) = 3$$

$$\pi(3) = 1$$

$$\pi(4) = 2$$

$$\pi(k) = k \quad \forall k > 4.$$

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

Notice 2 key facts:

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

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

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

Now fix  $\varepsilon > 0$  & since

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

$$= \liminf_m \bigcup_{n \geq m} \{S_n = 0\}$$

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

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

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

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

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

$$\therefore P\left(\{S_n=0 \text{ i.o.n}\} \Delta \bigcup_{n=m_\varepsilon}^{n_\varepsilon} \{S_n=0\}\right) \leq \varepsilon \quad (27)$$

|| by (ii)

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

|| 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

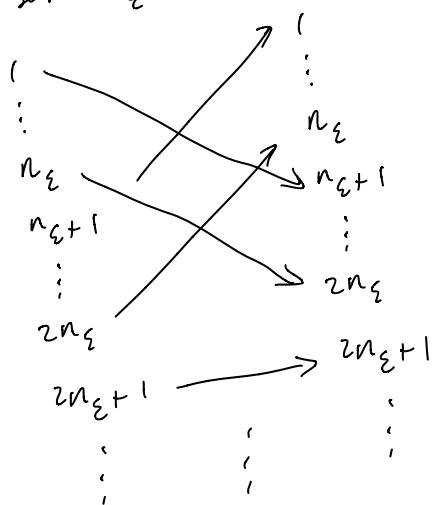
indices.

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

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

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

If we set  $\pi_\varepsilon$  as



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

In summary

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

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

Now

$$P((A \cap A) \Delta (A_\varepsilon \cap A_\varepsilon^{\pi_\varepsilon}))$$

$$\leq P(A \Delta A_\varepsilon) + P(A \Delta A_\varepsilon^{\pi_\varepsilon})$$

$\rightarrow 0$  as  $\varepsilon \rightarrow 0$  from (iii)

$\therefore$  an exercise shows

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

||

$$P(A_\varepsilon)P(A_\varepsilon^{\pi_\varepsilon}) \rightarrow P(A)P(A)$$

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

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  $\geq 3$ ,  $P(S_n=0 \text{ i.o.n})=0$ .

"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 "exchangeable" r.v.s  $R_1, R_2, \dots$

(28)