1.

Lecture 1

1. <u>Measure theory</u> (3)

2. LLN (3)

3. CLT (3)

4. Random Walks (2)

5. Martingales. (2)

ch. 1.
- probability space. $(\Omega, \mathcal{F}, P)$
- random variables and distributions. X, $\mathcal{L}(X)$.
- expected value. $E(X)$

~~Example~~: coin tossing space. $\Omega = \{H, T\}$.

$$\mathcal{F} = \{\phi, \Omega, \{H\}, \{T\}\}.$$

$$P(\phi) = 0, \quad P(\Omega) = 1$$

$$P(\{H\}) = \tfrac{1}{2}, \quad P(\{T\}) = \tfrac{1}{2}.$$

define X = number of Heads

$$P(X=0) = P(X=1) = \tfrac{1}{2}.$$

$$E(X) = \tfrac{1}{2}.$$

not so straightforward when Ω is not discrete. · this parts lays the foundation of probability

2.

~~Definitions~~. ~~[illegible]~~ $(\Omega, \mathcal{F}, P)$

Def: A sample space, denoted by Ω, is a set (of "outcomes")

Def: A ~~subset~~ ~~collection~~ collection of subsets of Ω, denoted by $\mathcal{F}$, is called a σ-field (or σ-algebra) if

(i). $\Omega \in \mathcal{F}$

(ii) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.

(iii) if $A_1, A_2, \cdots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

[equivalently $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$]

~~Remarks~~: Def: The above $(\Omega, \mathcal{F})$ is called a measurable space.

~~Rem~~:

Def. $\mu: \mathcal{F} \to \mathbb{R}$ is called a measure if

(1). $\mu(A) \geq 0. \quad \forall A \in \mathcal{F}$.

(2) $\mu(\phi) = 0$

(3) if $A_1, A_2 \cdots \in \mathcal{F}$ are disjoint, then,

$$\mu\Big(\underbrace{\bigcup_{i=1}^{\infty} A_i}_{}\Big) = \sum_{i=1}^{\infty} \mu(A_i) \quad (\in \mathcal{F})$$

If $\exists A_i \uparrow \Omega$, $\mu(A_i) < \infty$ [$A_1 \subset A_2 \subset \cdots$ & $\bigcup A_i = \Omega$], then μ is called a σ-finite measure

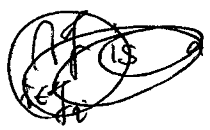
Def. If $\mu(\Omega) < \infty$, μ is called a finite measure.

If $\mu(\Omega) = 1$, μ is called a probability measure.

3. What can we say based on these definitions?

- smallest ~~σ-algebra~~ σ-field? $\{\phi, \Omega\}$.
- largest ~~σ-algebra~~ σ-field? $\{$All subsets of $\Omega\}$. 2^{Ω}

- If $\mathcal{F}_i, i \in I$ are all σ-fields, then



$\bigcap_{i \in I} \mathcal{F}_i$ is a σ-field

Def: Let $\mathcal{A}$ be a class of subsets of Ω,

$\sigma(\mathcal{A})$ denotes the smallest σ-field containing $\mathcal{A}$,

or equivalently $\sigma(\mathcal{A}) := \cancel{\bigcap \{\mathcal{F} : \mathcal{A} \subset \mathcal{F}\}}$ ~~all σ-fields~~ $\bigcap_{\mathcal{A} \subset \mathcal{F},\ \mathcal{F} \text{ is a } \sigma\text{-field}} \mathcal{F}$.

Example when $\mathcal{A} = \{A\}$.

$$\sigma(\mathcal{A}) = \{\phi, \Omega, A, A^c\}.$$

Def. A collection of subsets of Ω, $\mathcal{F}$ is called a field (or algebra)

if (i) $\Omega \in \mathcal{F}$

(ii) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$

(iii)' if $A_1, \dots, A_n \in \mathcal{F}$, then $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$.

- Any σ-field is a field but not vice versa.

Counter-example: $\Omega = \mathbb{Z}$, $\mathcal{F} = \{A \subset \mathbb{Z}$, either A or A^c is a finite set$\}$

4. **Def**: Borel σ-field on $\mathbb{R}$ ($\Omega = (-\infty, \infty)$), denoted by $\mathcal{B}$ or $\mathcal{R}$, is defined to be

$$\mathcal{B} = \sigma(\mathcal{A}) \quad \text{where}$$

$$\mathcal{A} = \{(a,b],\ a<b\}.$$

• $\sigma(\{(a,b],\ a<b\}) = \sigma(\{(a,b),\ a<b\})$
$= \sigma(\{\text{open sets in } \mathbb{R}\})$

Def: Borel σ-field on $\mathbb{R}^d$, denoted by $\mathcal{B}$ or $\mathcal{R}^d$, is defined to be

$$\mathcal{B} = \sigma(\mathcal{A}) \quad \text{where}$$

$$\mathcal{A} = \{(a_1,b_1] \times (a_2,b_2] \times \cdots \times (a_d,b_d] : a_i < b_i\}.$$

Thm: Let μ be a measure on a measurable space $(\Omega, \mathcal{F})$.

then (i). Monotonicity: if $A \subset B$ then $\mu(A) \le \mu(B)$

[proof: $\because$ A and $B \setminus A$ are disjoint.
$\therefore$ $\mu(B) = \mu(A \cup (B \setminus A)) \overset{(3)}{=} \mu(A) + \mu(B \setminus A) \overset{(1)}{\ge} \mu(A)$].

(ii). $\forall A, B$: $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$

(iii) Subadditivity: If $A \subset \bigcup_{m=1}^{\infty} A_m$, then $\mu(A) \le \sum_{m=1}^{\infty} \mu(A_m)$

[proof $\because$ $A = A \cap (\bigcup_{m=1}^{\infty} A_m) = (A \cap A_1) \cup (A \cap (A_2 \setminus A_1)) \cup (A \cap (A_3 \setminus (A_1 \cup A_2))) \cup \cdots$ disjoint

$\therefore$ $\mu(A) = \mu(A \cap A_1) + \mu(A \cap (A_2 \setminus A_1)) + \cdots$
$\le \mu(A_1) + \mu(A_2) + \cdots$]

(iv). Continuity from below. If $A_i \uparrow A$ [Recall this means $A_1 \subset A_2 \subset \cdots$ & $\bigcup_{i \ge 1} A_i = A$]

then $\mu(A_i) \uparrow \mu(A)$

[proof. Let $B_m = A_m - A_{m-1}$ the B_m are disjoint

$\mu(A) = \mu(\bigcup_{m \ge 1} B_m) \overset{(3)}{=} \sum_{m \ge 1} \mu(B_m) = \lim_{n \to \infty} \sum_{m=1}^{n} \mu(B_m) = \lim_{n \to \infty} \mu(A_n)$].

5. (v). Continuity from above: If $A_n \downarrow A$ with $\mu(A_1) < \infty$.

i.e., $[A_1 \supset A_2 \supset \cdots, \ \& \ \bigcap_{n \ge 1} A_n = A]$

then $\mu(A_n) \downarrow \mu(A)$

[proof: consider $A_1 - A_n \uparrow A_1 - A$.].

~~Recall Borel σ-field on the real line~~ Measures on $(\mathbb{R}, \mathcal{R})$:

Def: ~~F is called a~~ $F: \mathbb{R} \to \mathbb{R}$ is called a Stieltjes measure function if $\{ {}^{(i)}_{(ii)}$

Fact: Every ~~σ-finite~~ Measure μ on $(\mathbb{R}, \mathcal{R})$ s.t. $\mu((a, b]) < \infty$.
if ~~a b < ∞~~. $-\infty < a < b < \infty$

determines a Stieltjes measure function F (up to constants)

s.t. $F(b) - F(a) = \mu((a, b])$, $\forall \ -\infty < a < b < \infty$

~~with the properties~~

$$\boxed{\begin{array}{l} \text{(i). } F \text{ is nondecreasing.} \\ \text{(ii). } F \text{ is right continuous, i.e.,} \\ \qquad \lim_{y \downarrow x} F(y) = F(x) \end{array}}$$

[proof: Given μ, define. $F(0) = c$.

$$F(x) = \begin{cases} c + \mu((0, x]). & \text{if } x > 0 \\ c - \mu((x, 0]) & \text{if } x < 0 \end{cases}$$
]

Thm. Every Stieltjes measure function F determines a unique measure on $(\mathbb{R}, \mathcal{R})$. with.

$$\mu((a, b]) = F(b) - F(a). \quad \forall \ -\infty < a < b < \infty.$$

what's the point?

F is cdf for probability distributions

- below we only prove that such a measure is unique if $\mu((-\infty, \infty)) < \infty$.

6. ~~existence [proof by Carathéodory's Extension Theorem, omitted]~~

uniqueness [by Dynkin's $\pi-\lambda$ theorem below]

Theorem 2.1.2.

Def. $\mathcal{A}$ is ~~called~~ π-system if

$$A, B \in \mathcal{A} \Longrightarrow A \cap B \in \mathcal{A}.$$

Example: $\{(-a, b],\ -\infty < a \le b < \infty\}$ is a π-system

Def: $\mathcal{L}$ is a λ-system if

(i) $\Omega \in \mathcal{L}$

$[B \setminus A = (A \cup (B^c))^c] \rightarrow$ (ii) ~~$A \in \mathcal{L} \Rightarrow A^c \in \mathcal{L}$~~ if $A, B \in \mathcal{L}$ and $A \subset B$, then $B - A \in \mathcal{L}$

(iii) ~~If A_1, A_2 are disjoint and $\in \mathcal{L}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{L}$~~ If $A_1, A_2, \ldots \in \mathcal{L}$ and $A_n \uparrow A$, then $A \in \mathcal{L}$.

Fact: If $\mathcal{L}$ is both a π-system and a λ-system,

then $\mathcal{L}$ is a σ-field

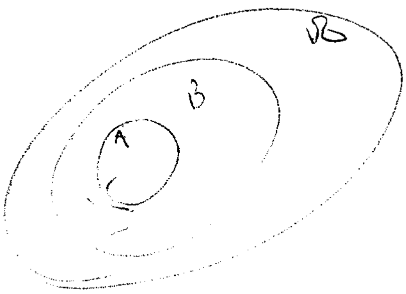
proof:

[If $A_1, A_2 \cdots \in \mathcal{L}$. (not necessarily disjoint)

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup (A_1 \cup A_2) \cup (A_1 \cup A_2 \cup A_3) \cup \cdots \in \mathcal{L} \ ]$$

Thm [Dynkin's $\pi-\lambda$ theorem]: If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system, and $\mathcal{P} \subset \mathcal{L}$.

then $\sigma(\mathcal{P}) \subset \mathcal{L}$



7. [Proof of uniqueness by the π-λ theorem :

~~Let~~ ~~Let $\mathcal{L}$ ... A~~

~~Let~~ Recall $\mathcal{R} = \sigma\{(-a, b] : -\infty < a < b < \infty\}$.

~~Let $\mathcal{L} = \{A \in \mathcal{R}$:~~

We want to show that if two measures μ_1 and μ_2

~~agree on~~ agree on $(-a, b]$ for all $-\infty < a < b < \infty$.

then they agree on $\mathcal{R}$.

Let $\mathcal{L} = \{A \in \mathcal{R} : \mu_1(A) = \mu_2(A)\}$.

$\because$ ~~Since~~ $\mu_1((-a, b]) = \mu_2((-a, b])$ $\left[= F(b) - F(a) \right]$, $\forall -\infty < a \le b < \infty$

$\therefore$ $\mathcal{L} \supset \{(-a, b] : -\infty < a \le b < \infty\}$.

[a π-system]

It can be verified that $\mathcal{L}$ is a λ-system.

By the π-λ theorem, $\mathcal{L} \supset \mathcal{R}$

$\therefore$ $\mathcal{L} = \mathcal{R}$]

8. [Proof of the π-λ theorem:

π-λ theorem follows from

(a): If $\lambda(\mathcal{P})$ is the smallest λ-system containing $\mathcal{P}$

then $\lambda(\mathcal{P})$ is a σ-field.

To prove (a), it is enough to show that

(b). $\lambda(\mathcal{P})$ is a closed under intersection ~~π-system~~.

xywang@se.cuhk.edu.cn.

To prove (b), we let

$\mathcal{G}_A = \{B \in \lambda(\mathcal{P}) : A \cap B \in \lambda(\mathcal{P})\}$ and

prove

(c). If $A \in \lambda(\mathcal{P})$, then $\mathcal{G}_A$ is a λ-system.

Check (c):

(i). $\Omega \in \mathcal{G}_A$? ✓

(ii) If $B, C \in \mathcal{G}_A$, $B \supset C$, then $B - C \in \mathcal{G}_A$? ✓

(iii) If $B_n \in \mathcal{G}_A$, $B_n \uparrow B$, then $B \in \mathcal{G}_A$? ✓]

- We can define similar measure functions F on $\mathbb{R}^d$, which must satisfy the additional condition:

$$\underline{\underline{\Delta_A F \geq 0}}$$

[measure on the multidimensional rectangle A].

1.

Lecture 2

$$(\Omega, \mathcal{F}, \mu)$$

$$X \downarrow \qquad \downarrow P$$

Example 1: $(\mathbb{R}, \mathcal{B}, \mu \; \mathcal{L}(X))$

$\mu \Leftrightarrow F$

If $F(x) = x$, then such determined μ is $\mu(a,b] = \cancel{F(b)-F(a)} = b-a$

the Lebesgue measure

If $F(-\infty)=0$, $F(\infty)=1$, then μ is a probability measure, denoted by P. $\mu(-\infty,\infty)=1$

Example 2: Ω is discrete (finite or countably infinite), $\Omega = \{e_1, e_2, \dots\}$

$\mathcal{F} = \{$all subsets of $\Omega\}$.

μ is determined by $\mu(e_1), \mu(e_2), \dots$

Today: X (random variables), $\mathcal{L}(X)$ (distributions).

Def: Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be measurable spaces.

$f: \Omega_1 \to \Omega_2$ is called <u>measurable</u> if

for any $A \in \mathcal{F}_2$, $f^{-1}(A) \in \mathcal{F}_1$

where $f^{-1}(A) = \{w_1 \in \Omega_1 : f(w_1) \in A\}$.

- by enlarging $\mathcal{F}_1$, we can make any f measurable.

Fact 1. $\{f^{-1}(A) : A \in \mathcal{F}_2\} =: \cancel{f^{-1}(\mathcal{F}_2)}$ $\sigma(f)$ $f^{-1}(\mathcal{F}_2)$ is a σ-field in Ω_1

[proof: ~~(i) $\Omega_1 = f^{-1}(\Omega_2) \in f^{-1}(\mathcal{F}_2)$~~

~~(ii) if $A_1 \in f^{-1}(\mathcal{F}_2)$, i.e. $A_1 = f^{-1}(A_2)$ for some $A_2 \in \mathcal{F}_2$~~

~~then $A_1^c = f^{-1}(A_2^c) \in f^{-1}(\mathcal{F}_2)$~~

e.g. if $f(e)$ is the same for all $e \in \Omega_1$, $\sigma(f) = \{\emptyset, \Omega_1\}$

$f^{-1}(A^c) = (f^{-1}(A))^c \qquad f^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} f^{-1}(A_i)$]

2.

Fact 2: $\{B \subset \Omega_2 : f^{-1}(B) \in \mathcal{F}_1\}$ is a σ-field in Ω_2

Consequence: If $\mathcal{F}_2 = \sigma(\mathcal{A}_2)$ then to check f is measurable

we only need to check that $\forall B \in \mathcal{A}_2$. $f^{-1}(B) \in \mathcal{F}_1$.

Fact 3: ~~If~~ Let $(\Omega_1, \mathcal{F}_1)$, $(\Omega_2, \mathcal{F}_2)$, $(\Omega_3, \mathcal{F}_3)$ be measurable spaces.

If $f_1 : \Omega_1 \to \Omega_2$ and $f_2 : \Omega_2 \to \Omega_3$ are both measurable

then $\underbrace{f_2 \circ f_1}_{\text{composition}} : \Omega_1 \to \Omega_3$ is measurable

- If there is a measure μ_1 on $(\Omega_1, \mathcal{F}_2)$.
  through a measurable function $f : \Omega_1 \to \Omega_2$
  μ_1 induces a measure μ_2 on $(\Omega_2, \mathcal{F}_2)$ such that
  $$\mu_2(A) = \mu_1(f^{-1}(A))$$

A special case (that we are most interested in):

Def. Let $(\Omega, \mathcal{F})$ be a measurable space. ~~Recall~~ Recall $(\mathbb{R}, \mathcal{R})$, $(\mathbb{R}^d, \mathcal{R}^d)$

~~Let~~ If $f : \Omega \to \mathbb{R}$ is measurable. then f is called
a real valued (or one-dimensional) random variable, denoted by X.

If $f : \Omega \to \mathbb{R}^d$ is measurable, $d \geq 2$, then f is called
a d-dimensional random variable (or a random vector), denoted by

$$X = (X_1, \cdots, X_d)^T$$

3.

<u>Fact</u>: $X=(X_1, \dots, X_d)^T$ is a random vector

iff X_i is a random variable for $1 \le i \le d$.

[proof: "$\Rightarrow$" ~~for $B \in \mathcal{R}$~~ $\cancel{X_i^{-1}(B) = X^{-1}(\mathbb{R}\times\dots\times\mathbb{R}\times B\times\mathbb{R}\times\dots\times\mathbb{R})}$

$$X_i^{-1}((a,b]) = X^{-1}(\overset{1}{\mathbb{R}}\times\dots\times\overset{i-1}{\mathbb{R}}\times \overset{i}{(a,b]}\times\overset{i+1}{\mathbb{R}}\times\dots\times\overset{d}{\mathbb{R}}) \in \mathcal{F}$$

"$\Leftarrow$" $X^{-1}\left((a_1,b_1]\times(a_2,b_2]\times\dots\times(a_d,b_d]\right)$

$$= \left[X_1^{-1}((a_1,b_1])\right]\cap\dots\cap\left[X_d^{-1}((a_d,b_d])\right] \in \mathcal{F}. \quad]$$

<u>Consequence</u>: If $X_1,\dots,X_n$ are random variables, ~~then~~ and $f: (\mathbb{R}^n,\mathcal{R}^n)\to(\mathbb{R},\mathcal{R})$ is a measurable function then $f(X_1,\dots,X_n)$ is a random variable.

~~e.g. $X_1+\dots+X_n$ is a random variable.~~

~~[proof.~~

- $X_1+\dots+X_n$ is a random variable.

[proof: $\{x: x_1+\dots+x_n < a\}$ is an open set in $\mathbb{R}^n$]

<u>Theorem 1.3.5</u>: If $X_1, X_2, \dots$ are random variables, then so are

$\inf_n X_n$ $\sup_n X_n$ $\limsup_n X_n$ $\liminf_n X_n$, regarded as functions from Ω to $([-\infty,\infty], \mathcal{R}^*)$

[proof: $\{\inf_n X_n < a\} = \bigcup_n \{X_n < a\} \in \mathcal{F}$.

$\{\sup_n X_n > a\} = \bigcup_n \{X_n > a\} \in \mathcal{F}$.

$\limsup_n X_n = \inf_n\left(\sup_{m\ge n} X_m\right)$

$\liminf_n X_n = \sup_n\left(\inf_{m\ge n} X_m\right)$.]

4. Note that $\Omega_0 := \{\omega \in \Omega : \lim_{n\to\infty} X_n \text{ exists}\}$

$= \{\omega \in \Omega : \limsup_{n\to\infty} X_n - \liminf_{n\to\infty} X_n = 0\}$ is measurable.

<u>Def</u>: If $\mu(\Omega_0) = \mu(\Omega)$ then we say X_n converges almost everywhere (a.e.)

If $\mu(\Omega_0) = \mu(\Omega) = 1$ — — — almost surely (a.s.)

$\mathcal{L}(X)$:

<u>Def</u>: Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Let $X : \Omega \to \underline{\mathbb{R}}$ be a real valued random variable (r.v.)

The induced measure $\mu(A) = P(\{\omega \in \Omega : X(\omega) \in A\})$
$=: P(X \in A)$

is called the probability measure (or probab. distribution) of X.

<u>Def</u>: Distribution function (d.f.) of X is

$F(x) = F_X(x) = P(X \le x)$

<u>Recall properties of F</u>:

1. $F \uparrow$
2. F is right continuous
3. $\lim_{x\to-\infty} F(x) = 0 \quad \lim_{x\to\infty} F(x) = 1$.

- it is a measurable function from $(\mathbb{R}, \mathcal{R})$ to $(\mathbb{R}, \mathcal{R})$.

<u>Def</u>: Let X be a $\mathbb{R}^d$-valued random vector.

The distribution function of X is defined by $F : \mathbb{R}^d \to [0,1]$

$$F(x) = P(X_1 \le x_1, \dots, X_d \le x_d)$$

~~Fact:~~

Example: If $F(x) = \begin{cases} 0 & x \le 0 \\ x & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$

then it is called the uniform distribution.

Fact. If X has d.f. F_X, and F_X is continuous

then $Y := F_X(X)$ has the uniform distribution.

[proof: for $0 < y < 1$:

$$\underbrace{P(Y \le y)}_{F_Y(y)} = P(F_X(X) \le y) = P(X \le \underbrace{F_X^{-1}(y)}_{\text{largest preimage}})$$

$$= F_X(F_X^{-1}(y)) = y \quad]$$

Given a d.f. F, define ~~$F^{-1}(y) :=$~~ and $w \in (0,1)$

$$F^{-1}(w) := \boxed{\inf \{y : F(y) \ge w\}}$$

~~$= \sup \{y : F(y) < w\}$~~

Thm: Let $\Omega = (0,1)$. $\mathcal{F}$ = Borel ~~σ-field~~ sets.

P = Lebesgue measure.

$X : \Omega \to \mathbb{R}$ is defined to be

$X(\omega) = F^{-1}(\omega)$ as above.

then the d.f. of X is F.

6.

[proof: Only need to show that

$$\cancel{P(\{w: X(w) \le x\}) = F(x)}$$

$$P(X \le x) = F(x)$$

$$\|$$

$$P(\{w: F^{-1}(w) \le x\})$$

$$\|? \checkmark$$

$$P(\{w: w \le F(x)\})$$

]

Def. X and Y are <u>equal in distribution</u> if $F_X(x) = F_Y(x)$, $\forall x \in \mathbb{R}$.

- note that X and Y may be defined on different probability spaces
  - maybe two different r.v.'s on the ~~samp~~ same probab. space, with the same distribution

Def. The <u>support</u> of X is defined to be

$$\{x \in \mathbb{R} : F(x+\varepsilon) - F(x-\varepsilon) > 0,\ \forall \varepsilon > 0\}.$$

Fact: The set of discontinuity of F is countable; denoted by

$$\{a_1, a_2, \cdots\}.$$

Let $b_j = F(a_j) - F(a_j-)$.

If $\sum_{j=1}^{\infty} b_j = 1$, then F is called a $\boxed{\text{discrete distribution}}$

- e.g.: $F(a) = 1, F(a-) = 0$ called a point mass at a.

If $\sum_{j=1}^{\infty} b_j = 0$, then F is called a continuous distribution.

7. If $F(x) = \int_{-\infty}^{x} f(y)\,dy$ $\forall x$

Then F is $\boxed{\text{absolutely continuous}}$ and has density function f.

Thm: Any distribution function F can be written as

$$F = c_1 \underset{\text{discrete}}{F_d} + c_2 \underset{\text{absolutely continuous}}{F_{ac}} + c_3 \underset{\text{singular}}{F_s}$$

where $c_1, c_2, c_3 \geq 0$. $c_1 + c_2 + c_3 = 1$

F_s' exists and equal to 0 almost everywhere

Examples of distributions:

- $N(\mu, \sigma^2)$: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x \in \mathbb{R}$
- $\exp(\lambda)$: $f(x) = \lambda e^{-\lambda x}$, $x > 0$
- Poisson(λ) : $P(k) = \frac{e^{-\lambda}\lambda^k}{k!}$ $k = 0, 1, 2, \ldots$

log normal, chi-squared dist., Gamma. Cauchy distr. Beta

check wikipedia

8. **Properties of $N(0,1)$:** p.d.f. $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

d.f.: $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$

We have: If $x>0$, then

$$\left(\frac{1}{x}-\frac{1}{x^3}\right)\varphi(x)\le 1-\Phi(x)\le\begin{cases}\frac{1}{x}\varphi(x)\\ \cancel{\frac{1}{2}\varphi(x)}\ \frac{1}{2}e^{-\frac{x^2}{2}}\end{cases}$$

[proof:
$$1-\Phi(x)=\int_x^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$$
$$=-\int_x^{\infty}\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{y}\,de^{-\frac{y^2}{2}}$$
$$=-\frac{1}{\sqrt{2\pi}}\frac{1}{y}e^{-\frac{y^2}{2}}\Big|_x^{\infty}-\int_x^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\frac{1}{y^2}dy$$
$$=\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{x}e^{-\frac{x^2}{2}}-\frac{1}{\sqrt{2\pi}}\frac{1}{x^3}e^{-\frac{x^2}{2}}+$$
]

Gaussian integration by parts: $EZf(Z)=Ef'(Z)$.

$0 \qquad x \qquad x+\delta$

$$P(x\le Z\le x+\delta)\approx\varphi(x)\cdot\delta\sim\delta x\cdot(1-\Phi(x))$$

1.

Lecture 3

$$X : \Omega \to \mathbb{R}.$$

$\mathcal{L}(X)$: probability measure on $(\mathbb{R}, \mathcal{R})$ $\Longleftrightarrow$ F_X

Correction, $F_X(X) \sim Unif(0,1)$ iff F_X is continuous

Expected value / Expectation of a random variable

[Recall from the elementary probability course: If F_X is a discrete distribution, then

$$E(X) = \sum_{i=1}^{\infty} a_i \cdot \underset{\Omega}{P}(X=a_i)$$

If F_X has density function f, then

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx \quad \text{change of variable}$$]

Now we ~~need to worry about~~ general F, existence of integrals, etc.
~~give a formal definition~~ taking care of
give a formal treatment

Recall notation

$$\begin{array}{c} (\Omega, \mathcal{F}, P) \\ \big\downarrow X \\ (\mathbb{R}, \mathcal{R}, \mathcal{L}(X)) \end{array}$$

2.

Def: $\boxed{\text{Indicator random variable}}$:

Given a ~~set~~ $A \in \mathcal{F}$.. $\cancel{1(A) := \{ \begin{smallmatrix} 1 \\ 0 \end{smallmatrix}}$

$$X(\omega) = 1_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}$$ (is a random variable)

Def 1 $E(1_A) := P(A)$

Def: $\boxed{\text{Simple random variable}}$

~~Given disjoint sets~~ where $A_1, A_2, \dots, A_n \in \mathcal{F}$ are disjoint and $a_1, a_2, \dots, a_n \in \mathbb{R}$

$$X = \sum_{i=1}^{n} a_i 1_{A_i} \left(+ 0 1_{(\bigcup_{i=1}^{n} A_i)^c} \right)$$

Def 2: $E(X) = \sum_{i=1}^{n} a_i P(A_i)$

Fact: If $X = \sum_{i=1}^{n} a_i 1_{A_i} = \sum_{j=1}^{m} b_j 1_{B_j}$

where $A_1, \dots, A_n$ are disjoint, $B_1, \dots, B_m$ are disjoint

then. $\sum_{i=1}^{n} a_i P(A_i) = \sum_{j=1}^{m} b_j P(B_j)$

[proof: Assume WLOG $\bigcup_{i=1}^{n} A_i = \bigcup_{j=1}^{m} B_j = \Omega$. (otherwise $+ 0 \cdot 1_{(\bigcup_{i=1}^{n} A_i)^c}$)

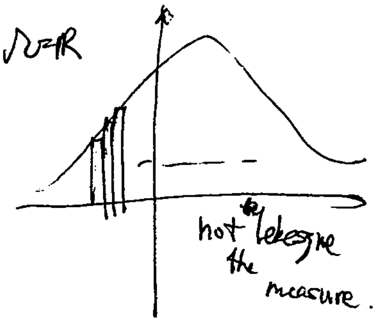
then $\sum_{i=1}^{n} a_i P(A_i) = \sum_{i=1}^{n} a_i P(A_i \cap (\bigcup_{j=1}^{m} B_j)) = \sum_{i=1}^{n} a_i P\left(\bigcup_{j=1}^{m} (A_i \cap B_j)\right)$

$= \sum_{i=1}^{n} a_i \sum_{j=1}^{m} P(A_i \cap B_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} E(a_i 1_{A_i \cap B_j})$

$= \sum_{i=1}^{n} \sum_{j=1}^{m} E(b_j 1_{A_i \cap B_j}) = \sum_{i=1}^{n} \sum_{j=1}^{m} b_j P(A_i \cap B_j)$

$= \dots = \sum_{j=1}^{m} b_j P(B_j)$].

3.



Def 3. If $X \geq 0$ define

$$E(X) := \sup_{\substack{0 \le Y \le X \\ Y \text{ is a simple r.v.}}} E(Y)$$

• ~~does not~~ gives previous definitions for simple random variables

~~$0 \times \infty = 0$.~~

Write a general r.v.

Def 4. $X = X^+ - X^-$, where $X^+ = \max\{X, 0\}$.
$X^- = \max\{-X, 0\}$

$E(X^+)$ and $E(X^-)$ are defined as above.

If $E(X^+) = E(X^-) = \infty$. then the expected value of X does not exist.

Otherwise, define $E(X) = E(X^+) - E(X^-)$

If both $E(X^+)$ and $E(X^-) < \infty$. then $E(X)$ is finite

$E(|X|) = E(X^+) + E(X^-)$ is also finite

• Def 3 and Def 4 can be defined similarly for generalized r.v.'s taking values on $[-\infty, \infty]$

- In particular ~~$0 \times \infty = 0$~~

set with measure 0 can be neglected in the expectation.

$$X = \begin{cases} 0 & \text{in } \Omega_0 \\ \infty & \text{in } \Omega_0^c \end{cases} \quad \text{and} \quad P(\Omega_0) = 1$$

$E(X) = 0$.

$0 \cdot \infty = 0$

4. Properties of $E(\cdot)$. : Suppose $X, Y \geq 0$ or $E|X|, E|Y| < \infty$.

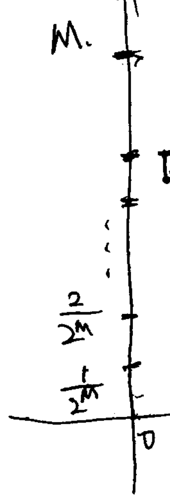
(a). $E(X+Y) = E(X) + E(Y)$

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

(c). If $X \geq Y$ a.s. then $EX \geq EY$.

(a) in the case $X, Y \geq 0$:

$M \geq n$, M is an integer



$X_M^{(l)} \leq X \leq X_M^{(u)}$

$\frac{1}{2^M}$

$$E(X+Y) \leq E(X+Y)_M^{(u)} \leq E(X_M^{(u)}) + E(Y_M^{(u)}) \leq E(X) + \frac{1}{2^M} + E(Y) + \frac{1}{2^M}$$

[proof: of (a) in the case $E|X|, E|Y| < \infty$: ~~$|X| = X^+ - X^-$~~

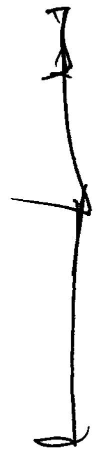
$$X = X^+ - X^-, \quad Y = Y^+ - Y^-.$$

$$E(X+Y) = E\left((X^+ + Y^+) - (X^- + Y^-)\right)$$

$$\xlongequal{\text{by (a) for the case } X, Y \geq 0} E(X^+ + Y^+) - E(X^- + Y^-)$$

$$\xlongequal{\text{again}} E(X^+) + E(Y^+) - E(X^-) - E(Y^-)$$

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



(c): $X = Y + Y'$, $Y' \geq 0$ a.s.

$$(a) \Rightarrow E(X-Y) = E(X) - E(Y)$$

$$\parallel$$

$$E(Y') \geq 0$$

Next: $\lim_{n\to\infty} E(X_n) \overset{?}{=} E(\lim_{n\to\infty} X_n)$

5. <u>Thm</u>: (Monotone convergence theorem):

$X_n \geq 0, \forall n.$

If $X_n \uparrow X$. (i.e., $0 \leq X_1 \leq X_2 \leq \cdots$, $\lim_{n\to\infty} X_n = X$).

then. $E(X_n) \uparrow E(X)$.

[proof: ~~Case 1: $E(X) = \infty$~~.

$E(X_n) \uparrow a$.

If $a = \infty$. ✓.

If $a < \infty$. then we need to prove $E(X) \leq a$.

Recall $E(X) = \sup_{0 \leq Y \leq X} E(Y)$, only need to show $E(Y) \leq a$

or $E(Y) \leq a + \varepsilon$, $\forall \varepsilon$.

Y is a simple r.v.

Suppose $Y = \sum_{j=1}^{m} b_j 1_{B_j}$ B_j disjoint.

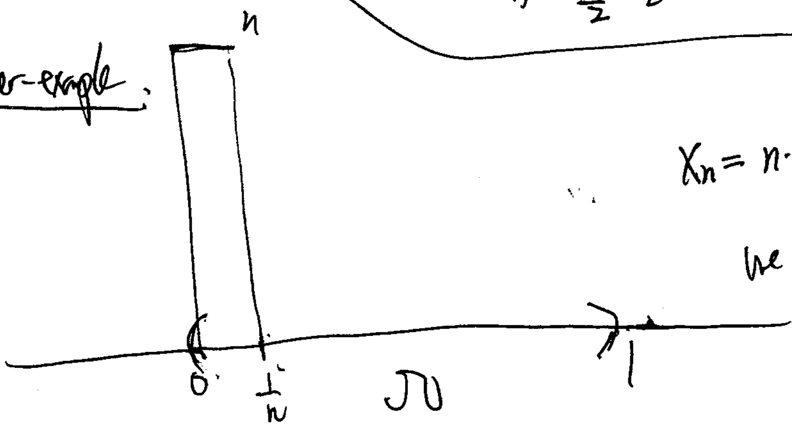
define $Y_\varepsilon = \sum_{j=1}^{m} (b_j - \frac{\varepsilon}{2}) 1_{B_j}$ then ~~$Y_\varepsilon \leq X_n$ for large enough n.~~ for large enough n,

~~which means~~ $E(Y_\varepsilon) \overset{-\frac{\varepsilon}{2}}{\leq} E(X_n) \leq a$

$E(Y_\varepsilon) \geq E(Y) - \frac{\varepsilon}{2} - \frac{\varepsilon}{2}$, $E(X_n) \leq a$

■]

<u>Counter-example</u>:



$X_n = n \cdot 1_{(0, \frac{1}{n})}$.

we have $X_n \geq 0$.

$X_n \to 0$.

but $E(X_n) = 1 \not\to 0$.

- later we will see that we only need a weaker condition instead of $X_n \geq 0$ $\forall n$

6. <u>Fact</u> Given $X \geq 0$.

We can find a sequence of simple r.v.'s $X_n \geq 0$. s.t $X_n \uparrow X$.

[Proof: related to the Lebesgue integral technique].

<u>Thm</u> (Fatou's Lemma).

If $X_n \geq 0 \quad \forall n$. then.

$$\liminf_{n\to\infty} E[X_n] \geq E[\liminf_{n\to\infty} X_n].$$

[proof:
$$
\begin{aligned}
\liminf_{n\to\infty} E[X_n] &\geq \liminf_{n\to\infty} E[\inf_{k\geq n} X_k] \\
&= \lim_{n\to\infty} E[\inf_{k\geq n} X_k] \\
&\overset{MCT.}{=} E[\lim_{n\to\infty} \inf_{k\geq n} X_k] \\
&= E[\liminf_{n\to\infty} X_n]
\end{aligned}
$$
].

counter-examples

7. Thm (Dominated Convergence Theorem).

If $X_n \to X$ a.s. and $|X_n| \le Y$, for some Y s.t. $E[Y] < \infty$.

then $E[X_n] \to E[X]$

[proof: ~~Only need to show (N.T.S.)~~ (Need To Show) N.T.S.: $\liminf_{n\to\infty} E(X_n) \ge E(X)$ ✓.

& $\limsup_{n\to\infty} E(X_n) \le E(X)$

$\cancel{|E[X_n - X]|} \cancel{\le E|X_n - X| \to 0}$

$\cancel{\limsup_{n\to\infty} E|X_n - X|}$

$\cancel{= -\liminf_{n\to\infty} [-E|X_n - X|]}$

• $\liminf_{n\to\infty} E(X_n + Y) \overset{\text{Fatou}}{\ge} E(X + Y)$

• $\limsup_{n\to\infty} E(X_n - Y) = -\liminf_{n\to\infty} E(-X_n + Y)$

$\le - E(-X_n + Y)$

$= E(X_n - Y)$]

• We can relax the conditions in MCT & Fatou's lemma using DCT.

$$(\Omega, \mathcal{F}, P) \xrightarrow{X} (\mathbb{R}, \mathcal{R}, \underset{\text{call it } \mu}{\mathcal{L}(X)}) \xrightarrow{f} (\mathbb{R}, \mathcal{R}, \mathcal{L}(f(X)))$$

with the composite $f(X)$.

Change of variables formula: If either $f \ge 0$ or $E|f(X)| < \infty$:

$$E_P[f(X)] = E_\mu[f]$$

Theorem 1.6.9.

• $\int_{\varphi(a)}^{\varphi(b)} f(u)\,du = \int_a^b f(\varphi(x))\varphi'(x)\,dx$

[Proof: $f = 1_B$, $B \in \mathcal{R}$.

LHS $= P(X^{-1}(B))$

RHS $= \mu(B)$ ✓

...]

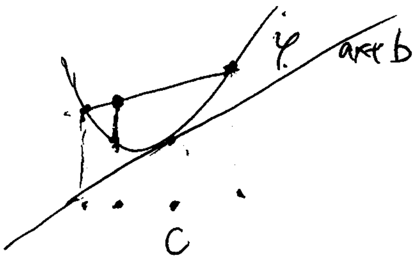
8. Useful inequalities:

Jensen's inequality: X is r.v.

φ is convex function.

$E|X|<\infty$ and $E|\varphi(X)|<\infty$.

then. $E(\varphi(X)) \geq \varphi(E(X))$.



[proof: Let $c=E(X)$.

let $ax+b$ be s.t. $\varphi(c)=ac+b$.
$\varphi(x)\geq ax+b$

$E(\varphi(X))\geq aE(X)+b=\varphi(c)=\varphi(E(X))$.]

· for concave functions ... $\log(x)$

Examples, $E[|X|^p]\geq (E|X|)^p$. for $p\geq 1$. e^x

Hölder's inequality: If $p,q\geq 1$. $\frac{1}{p}+\frac{1}{q}=1$. then.

$$E[|XY|] \leq (E|X|^p)^{\frac{1}{p}} \cdot (E|X|^q)^{\frac{1}{q}}.$$

- $\|X\|_p := (E|X|^p)^{\frac{1}{p}}$. $\|X\|_\infty := \cancel{\sup|X|}\ \text{esssup}|X|$
$:= \inf\{a: P(|X|>a)=0\}$

If $p=q=2$, it is called the Cauchy-Schwarz inequality.

· If $r=1$. reduces to above example.

9.

[proof: Without loss of generality (W.L.O.G.):

We may assume $\|X\|_p = \|Y\|_q = 1$. and prove

$$E|XY| \le 1.$$

$\because$ $$xy \le \frac{x^p}{p} + \frac{y^q}{q} \quad \text{for } x, y \ge 0$$

$\therefore$ $$E|XY| \le \frac{1}{p} + \frac{1}{q} = 1 \quad]$$

<u>Minkowski's Inequality</u>: $p \ge 1$

$$\|X+Y\|_p \le \|X\|_p + \|Y\|_p$$

[proof: $$\left(E|X+Y|^p\right)^{\frac{1}{p}} = \left(E\, |X||X+Y|^{p-1} + E|Y||X+Y|^{p-1}\right)^{\frac{1}{p}}$$

$$\le \left[(E|X|^p)^{\frac{1}{p}} \cdot (E|X+Y|^{(p-1)q})^{\frac{1}{q}} + (E|Y|^p)^{\frac{1}{p}} \cdot (E|X+Y|^{(p-1)q})^{\frac{1}{q}} \right]^{\frac{1}{p}}$$

$\frac{1}{q} = 1 - \frac{1}{p}$ $\quad q = \frac{p}{p-1}$ $$= (\|X\|_p + \|Y\|_p)^{\frac{1}{p}} \cdot (E|X+Y|^p)^{\frac{1}{p}\cdot\frac{1}{q}} \quad]$$

<u>Chebyshev's Inequality</u>: $\forall a > 0$. $$P(X \ge a) \le \frac{E|X|}{a}.$$

[proof: $$P(X \le a) = E[I(X \ge a)] \le E\left[\frac{X}{a} \cdot 1(X \ge a)\right] \le \frac{E|X|}{a} \quad].$$

<u>Markov's inequality</u>: $$P(|X| \ge a) \le \frac{E|X|^2}{a^2}.$$

$$\parallel$$

$$P(|X|^2 \ge a^2)$$

$$\forall a > 0, t > 0: P(X > a) \le \frac{Ee^{tX}}{e^{ta}},$$

$$\parallel$$

$$P(e^{tX} > e^{ta})$$

For any r.v. $X: \Omega \to \mathbb{R}$

$$P(X > a) \to 0$$

① Independence. Lecture 4

Def. ~~sets~~ events A_1, A_2 are independent if $P(A_1 \cap A_2) = P(A_1)P(A_2)$

$A_1, \ldots, A_n$ are independent if $P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$

for any $I \subset \{1, \ldots, n\}$.

or equivalently. $P(B_1 \cdots B_n) = P(B_1) \cdot P(B_2) \cdots P(B_n)$

where. $B_i = A_i$ or Ω.

[$A_1^c \perp\!\!\!\perp A_2$; $\phi \perp\!\!\!\perp A$; $\Omega \perp\!\!\!\perp A$: $\{\phi, A_1, A_1^c, \Omega\} \perp\!\!\!\perp \{\phi, A_2, A_2^c, \Omega\}$ $1_{A_1} \perp\!\!\!\perp 1_{A_2}$]

Def. ~~σ-fields~~ $\mathcal{F}_1, \ldots, \mathcal{F}_n$ $\mathcal{A}_1 \cdots \mathcal{A}_n$ are independent if

[~~A collection~~ collections of sets

$P(A_1 \cdots A_n) = P(A_1) \cdots P(A_n)$ ~~where~~ $\forall A_i \in \mathcal{F}_i$ ($\mathcal{A}_i$ or $A_i = \Omega$), $i = 1, \ldots, n$.

Def. random variables. $X_1, \ldots, X_n$ are independent if

$\sigma(X_1), \ldots, \sigma(X_n)$ are independent.

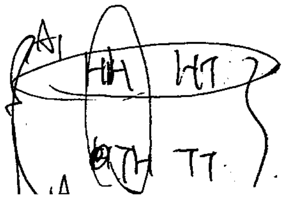
[Recall $\sigma(X_i) = \{X_i^{-1}(A), A \in \mathcal{B}\}$.] [$X_1 \equiv a$, then $a \perp\!\!\!\perp X_2$ for any X_2]

Def. Pairwise independence. if $\mathcal{F}_i$ is independent of $\mathcal{F}_j$ if $i \neq j$.

Def. ~~Independence of~~ An infinite collection of $\mathcal{F}$'s is independent

if every finite subcollection is.

[Example: 2-coin tossing



$A_1 = \{HH, HT\}$

$A_2 = \{HH, TH\}$].

② **Thm**: If $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_n$ are independent, and each $\mathcal{A}_i$ is a π-system

then. $\sigma(\mathcal{A}_1), \cdots, \sigma(\mathcal{A}_n)$ are independent.

[proof: N.T.S. $P(B_1 \cdots B_n) = P(B_1) \cdots P(B_n)$, $\forall B_i \in \sigma(\mathcal{A}_i)$.

Fix, $B_i \in \mathcal{A}_i$ or $B_i = \Omega$, $i = 2, \cdots, n$.

Let $\mathcal{L}_1 = \{B_1 \in \sigma(\mathcal{A}_1) : P(B_1 \cdots B_n) = P(B_1) \cdots P(B_n)\}$

$\because$ $\mathcal{L}_1 \supset \mathcal{A}_1$, $\mathcal{A}_1$ is a π-system,

$\Omega \in \mathcal{L}_1$; $B, A \in \mathcal{L}_1$ & $B \supset A \Rightarrow B \setminus A \in \mathcal{L}_1$; $\Rightarrow$ $\mathcal{L}_1$ is a λ-system

$A_i \in \mathcal{L}_1$ & $A_i \uparrow A \Rightarrow A \in \mathcal{L}_1$

$\therefore$ $\mathcal{L}_1 \supset \sigma(\mathcal{A}_1)$

We have proved that $\sigma(\mathcal{A}_1), \mathcal{A}_2, \cdots, \mathcal{A}_n$ are independent.

Repeat this argument to $\mathcal{A}_2, \cdots, \mathcal{A}_n$ completes the proof. ▪]

Thm: random variables $X_1, \cdots, X_n$ are independent iff

$$P(X_1 \le x_1, \cdots, X_n \le x_n) = \prod_{i=1}^{n} P(X_i \le x_i) \quad \forall x_i \in \mathbb{R}$$

Exercise: Suppose $(X_1, \cdots, X_n)$ has density $f(x_1, \cdots, x_n)$

and $f(x_1, \cdots, x_n) = g_1(x_1) \cdots g_n(x_n)$. $g_i \ge 0$ measurable

then. $X_1, \cdots, X_n$ are independent.

[proof: use above Thm and Fubini's theorem]

1.7.2

(3) **Exercise.** Suppose $\{X_1, \ldots, X_n\}$ are discrete random variables

then $X_1, \ldots, X_n$ are independent $\iff P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$.

Thm. If $X_1, \ldots, X_n$ are independent. Then.

$g(X_i, i \in I)$ and $h(X_j, j \in I^c)$ are independent.

$[\ g_1, g_2, \ldots, g_m\ ]$

[proof: $\{\bigcap_{i \in I} A_i : A_i \in \mathcal{F}_i\}$

$\{\bigcap_{j \in I^c} B_j : B_j \in \mathcal{F}_j\}$]

Fact. If $\mathcal{A}$ and $\mathcal{A}$ are indep. then $P(A) = 0$ or 1 $\forall A \in \mathcal{A}$.

[proof. $P(A) = P(A \cap A) = P(A) \cdot P(A)$]

Kolmogorov's 0-1 law: Theorem 2.5.1

$X_1, X_2, \ldots$, are independent, Let $\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots)$ be the tail σ-field. Then. $P(A) = 0$ or 1 $\forall A \in \mathcal{T}$.

[proof: (a). $\underline{\sigma(X_1, \ldots, X_k)} \perp\!\!\!\perp \sigma(X_{k+1}, X_{k+2}, \ldots)$

$$\sigma(X_{k+1}, X_{k+2}, \ldots) = \sigma\Big(\bigcup_{j=1}^{\infty} \underline{\sigma(X_{k+1}, \ldots, X_{k+j})}\Big)$$

(b) $\sigma(X_1, \ldots, \ldots) \perp\!\!\!\perp \mathcal{T}$.

$$\sigma(X_1, \ldots) = \sigma\Big(\bigcup_{j=1}^{\infty} \sigma(X_1, \ldots, X_j)\Big)$$

(c) $\mathcal{T} \subset \sigma(X_1, X_2, \ldots)$.]

④ Examples: ① If $X_1, \cdots$ indep.

$$\left\{ \lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} \text{ exists} \right\} \in \mathcal{T}.$$

$$\left\{ \lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = a \right\} \in \mathcal{T}.$$

②. ~~To prove~~ If $A \in \mathcal{T}$. to prove $P(A)=1$ only N.T.S. $P(A)>0$.

~~used in law of the iterated lo~~.

Thm: If X and Y are independent r.v.'s $E|X|, E|Y| < \infty$ or $X, Y \geq 0$

then $E(XY) = E(X)E(Y)$

[proof, 1. If $X = 1_A$ $Y = 1_B$. then $X \perp\!\!\!\perp Y \Leftrightarrow A \perp\!\!\!\perp B$.

$E(XY) = P(AB) = P(A)P(B) = E(X)E(Y)$

2. If $X = \sum_{i=1}^n a_i 1_{A_i}$ $Y = \sum_{j=1}^m b_j 1_{B_j}$

then. $E(XY) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j P(A_i B_j)$

$= \sum_{i=1}^n \sum_{j=1}^m a_i b_j P(A_i) P(B_j)$ [a_i are distinct b_j are distinct use independence]

$= E(X) E(Y)$

3. $X, Y \geq 0$. $X_n \uparrow X$ $Y_n \uparrow Y$.

4. $X = X^+ - X^-$.].

[X, Y, Z; $X \cdot Y$;]

⑤. Corollary: If $X, Y \in \mathbb{R}^d$ are independent.

$g, h : \mathbb{R}^d \to \mathbb{R}$.

$g, h \geq 0$ or $E|g(X)|, E|h(Y)| < \infty$.

Then $E\, g(X) h(Y) = E g(X) \cdot E h(Y)$

$[g_1, g_2, \ldots, g_n]$

Applications: If $X_1, \ldots, X_n$ are independent then

① $$E e^{t(X_1 + \cdots + X_n)} = \prod_{i=1}^{n} E e^{t X_i}$$

② $$E e^{it(X_1 + \cdots + X_n)} = \prod_{i=1}^{n} E e^{it X_i}.$$

Under suitable moment conditions,

③ $$E[(X_1 + \cdots + X_n)^2] = \sum_{i=1}^{n} \sum_{j=1}^{n} (EX_i)(EX_j)$$

Preparation for WLLN:

Def: $\mathrm{Cov}(X, Y) := E(X - EX)(Y - EY) = EXY - EX \cdot EY$.

- $\mathrm{Cov}(X, Y) = 0$ if $X \perp\!\!\!\perp Y$. not vise versa unless $X, Y \sim$ Normal.

Def. $$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

- $\mathrm{Corr}(aX + b, cY + d) = \mathrm{Corr}(X, Y) \cdot \mathrm{sign}(ac)$

(b).

Recall Def. $Y_n \to Y$ a.s.

Def. $Y_n \to Y$ in pr., denoted by $Y_n \xrightarrow{P} Y$

if $\forall \varepsilon>0$, $P(|Y_n - Y| > \varepsilon) \to 0$.

Def. Convergence in L_p: $p>0$.

$Y_n \to Y$ in L_p if $E|Y_n|^p<\infty$, $E|Y|^p<\infty$

& $E(|Y_n - Y|^p) \to 0$.

Fact. If $Y_n \to Y$ in L_p for some $p>0$

then $Y_n \to Y$ in prob.

[proof: $P(|Y_n - Y| > \varepsilon) \le \frac{E|Y_n - Y|^p}{\varepsilon^p} \to 0$. $\forall \varepsilon>0$.]

Thm (Weak Law of Large Numbers)

$X_1, X_2, \ldots$ uncorrelated, i.e. $Cov(X_i, X_j)=0$ for $i \neq j$.

$Var(X_i) \le C < \infty$, $E(X_i) = \mu_i$

Let $S_n = \sum_{i=1}^n X_i$

then. $\frac{S_n}{n} - \frac{\sum_{i=1}^n \mu_i}{n} \to 0$ in L_2, in prob.

[proof: $E\left(\frac{S_n}{n} - \frac{\sum_{i=1}^n \mu_i}{n}\right)^2 = Var\left(\frac{S_n}{n}\right) = \frac{\sum_{i=1}^n Var(X_i)}{n^2} \le \frac{C}{n} \to 0$.]

- Special case. $X_1, X_2, \ldots$ are independent and identically distributed (i.i.d.) with $EX_i^2<\infty$.

(7). <u>Thm</u>. S_n ~~as are independent~~ with $\sigma_n^2 = Var(S_n)$ and $\frac{\sigma_n^2}{b_n^2} \to 0$

then $\frac{S_n - E(S_n)}{b_n} \to 0$ in L_2 in prob.

<u>Example 2.2.3</u> Coupon collector's problem,

Let $X_1, X_2, \cdots$ be indep. and $Unif\{1, \cdots, n\}$.

Let $\tau_k^n = \inf\{n : |\{X_1, \cdots, X_m\}| = k\}$.

Let $T_n = \tau_n^n$.

Then. $\frac{T_n}{n \log n} \to 1$ in prob.

[proof: Let $X_k = \tau_k^n - \tau_{k-1}^n$, $k = 1, 2, \cdots n$

Then $X_k \sim Geom(\underbrace{1 - \frac{k-1}{n}}_{p})$. with $EX_k = \frac{n}{n-k+1}$ $Var(X_k) = (1-p)\frac{n^2}{(n-k+1)^2}$

$T_n = \sum_{k=1}^{n} X_k$ indep. sum.

$\therefore \sum_{k=1}^{n} EX_k = n \cdot (1 + \frac{1}{2} + \cdots + \frac{1}{n}) \sim n \cdot \log n$.

$Var(T_n) = \sum_{k=1}^{n} Var(X_k) \le \sum_{k=1}^{n} \frac{n^2}{(n-k+1)^2} \le n^2 \cdot (1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots) \le C \cdot n^2$

Choose $b_n = n \cdot \log n \gg Var(T_n)$

$\therefore \frac{T_n - n\log n}{n \log n} \to 0$.]

⑧. Truncation: When X_i does not have finite moments.

Thm 2.2.6: Weak LLN for triangular arrays

Assume $X_{n,k}$, $1 \le k \le n$ are independent.

Let $b_n > 0$. & $\bar{X}_{n,k} = X_{n,k} 1(|X_{n,k}| \le b_n)$

If (i). $\sum_{k=1}^{n} P(|X_{n,k}| > b_n) \to 0$ and

(ii). $b_n^{-2} \cdot \sum_{k=1}^{n} E\bar{X}_{n,k}^2 \to 0$.

Then, with $S_n = X_{n,1} + \dots + X_{n,n}$ and $a_n = \sum_{k=1}^{n} E\bar{X}_{n,k}$.

$\frac{S_n - a_n}{b_n}$ in prob.

[proof: Let $\bar{S}_n = \bar{X}_{n,1} + \dots + \bar{X}_{n,n}$

$$P\left(\left|\frac{S_n - a_n}{b_n}\right| > \varepsilon\right) \le P(S_n \ne \bar{S}_n) + P\left(\left|\frac{\bar{S}_n - a_n}{b_n}\right| > \varepsilon\right)$$

$\cancel{P(S_n \ne \bar{S}_n) \le P(}$ $\sum_{k=1}^{n} P(|X_{n,k}| > b_n) \xrightarrow{(i)} 0$ $\qquad$ $\xrightarrow{\text{(ii) + Chebyshev}} 0$]

⑨ Thm 2.2.7. WLLN: Let $X_1, \dots, X_n$ be i.i.d. with

$$x P(|X_i| > x) \to 0 \quad \text{as} \quad x \to \infty.$$

Let $S_n = X_1 + \dots + X_n$, $\mu_n = E(X_1 I_{(|X_1| \le n)})$

Then $\frac{S_n}{n} - \mu_n \to 0$ in prob.

[proof: Choose $b_n = n$ in previous theorem.

(i) ✓

(ii). $n^{-2} \sum_{k=1}^{n} \cancel{E(X_k^2 I(|X_k| \le n))}\ E(\bar{X}_k^2)$

$= \cancel{\frac{1}{n} E X_1^2 I(|X_1| \le n)}$

$= n^{-1} E(\bar{X}_1^2) \underset{\uparrow}{=} n^{-1} \int_0^\infty 2y\, P(\bar{X}_1 > y)\, dy$

Lemma 2.2.8

$= n^{-1} \int_0^n 2y\, P(\bar{X}_1 > y)\, dy$

$\le n^{-1} \int_0^n \underline{2y\, P(X_1 > y)}\, dy \to 0$ ✓]

Thm 2.2.9. Let $X_1, \dots, X_n$ be i.i.d. with $E|X_i| < \infty$.

Let $S_n = X_1 + \dots + X_n$, $\mu = EX_1$. Then $S_n/n \to \mu$ in prob.

[proof: $\mu_n = E(X_1 I_{(|X_1| \le n)}) \xrightarrow{DCT} \mu$.

$x P(|X_1| > x) \le E(|X_1| I_{(|X_1| > x)}) \xrightarrow{DCT} 0$]

(10). Example 2.2.7. The "St. Petersburg paradox".

Let $X_1, X_2, \ldots$ be independent with

$$P(X_i = 2^j) = 2^{-j}, \quad j \geq 1$$

$$EX_i = \infty.$$

- $$\frac{S_n}{n \log_2 n} \to 1. \quad \text{in prob.}$$

[proof: use Theorem 2.2.6:

$b_n = 2^{\log_2 n + k(n)}$, $k(n) \to \infty$. and $\log_2 n + k(n)$ ~~is~~ are integers.

(i) $$\sum_{k=1}^{n} P(X_{n,k} > b_n) = n \cdot 2^{-(\log_2 n + k(n))} \cdot 2 \to 0.$$

(ii). $$b_n^{-2} \cdot \sum_{k=1}^{n} E\bar{X}_{n,k}^2 = b_n^{-2} \cdot n \cdot \sum_{j=1}^{\log_2 n + k(n)} 2^{2j} \cdot \cancel{2^{-j}}$$

$$= b_n^{-2} \cdot n \cdot 2 \cdot 2^{\log_2 n + k(n)} = \frac{2n}{2^{\log_2 n + k(n)}} \to 0.$$

$$a_n = n \cdot E\bar{X}_n = n \cdot \sum_{j=1}^{\log_2 n + k(n)} 2^j \cdot 2^{-j} = n \cdot (\log_2 n + k(n))$$

$$\frac{S_n - n(\log_2 n + k(n))}{n \cdot 2^{k(n)}} \Rightarrow \frac{S_n}{n \log_2 n} - 1 \to 0 \quad \text{in prob} \;]$$

let $k(n) \leq \log_2 \log_2 n$

Lecture 5

WLLN: Let $X_1, \dots, X_n$ be i.i.d. with $E|X_1| < \infty$.

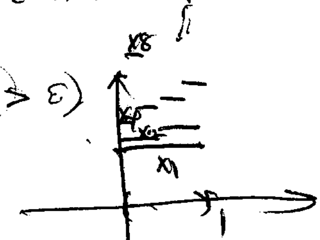
Then $\frac{X_1 + \cdots + X_n}{n} \to \mu := EX_1$ in prob.

02 OCT 2018

SLLN: (a.s.)

~~proof~~ $\cancel{P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right) \le}$

- $X_i = \pm 1$ $[-1, 1]$
- $P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right)$
- ~~SLLN~~: counterexample



Recall Def: Y_n converges to Y almost surely (a.s.) if

$$P\left(\{\omega : \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\}\right) = 1.$$

Def: Infinitely often (i.o.).

$$\{A_n \text{ i.o.}\} := \{\omega \in \Omega : \omega \text{ is in infinitely many } A_n\text{'s}\}$$

$$= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \limsup A_n$$

$$= \{\omega : \exists\, n_k \ \ \omega \in \{A_{n_k}\}\}$$

$$\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Fact: $Y_n \to Y$ a.s. $\iff \forall \varepsilon > 0,\ P(|Y_n - Y| > \varepsilon \text{ i.o.}) = 0$

$\Downarrow$

$$P\left(\limsup_{n\to\infty} |Y_n - Y| \le \varepsilon\right) = 1$$

$\Downarrow$

$$P\left(\limsup_{n\to\infty} |Y_n - Y| = 0\right) = \lim_{\varepsilon \downarrow 0} P\left(\limsup_{n\to\infty} |Y_n - Y| \le \varepsilon\right) = 1$$

② .

Theorem   Borel - Cantelli Lemmas

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

(ii). If A_n are independent then $\sum_{n=1}^{\infty} P(A_n) = \infty$ implies $P(A_n \text{ i.o.}) = 1$

[proof: (i) $P(A_n \text{ i.o.}) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} A_k\right) = 0.$

(ii) $P\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c\right) = \lim_{n\to\infty} P\left(\bigcap_{k=n}^{\infty} A_k^c\right) = \lim_{n\to\infty} \prod_{k=n}^{\infty} P(A_k^c) = 0.$

$$\prod_{k=n}^{\infty} P(A_k^c) = \prod_{k=n}^{\infty} (1 - P(A_k)) \le \prod_{k=n}^{\infty} e^{-P(A_k)} = e^{-\sum_{k=n}^{\infty} P(A_k)} = 0$$

Theorem: If $Y_n \xrightarrow{a.s.} Y$ then $Y_n \xrightarrow{P} Y$ ✓.

If $Y_n \xrightarrow{P} Y$ then $\exists n_k$ $Y_{n_k} \xrightarrow{a.s.} Y$ . .

[proof. ~~N.T.S $\forall \varepsilon > 0$.~~

Choose $n_1 < n_2 < n_3 < \cdots$ s.t.

$$P\left(|Y_{n_k} - Y| > \frac{1}{2^k}\right) < \frac{1}{2^k}$$

$\because \sum_{k=1}^{\infty} P\left(|Y_{n_k} - Y| > \frac{1}{2^k}\right) < \infty$. ~~By B-C (i)~~

$\therefore$ By B-C (i): $P\left(|Y_{n_k} - Y| > \frac{1}{2^k} \text{ i.o.}\right) = 0$]

③. Theorem 2.3.8: If $A_1, A_2 \dots$ are pairwise independent, and $\sum_{i=1}^{\infty} P(A_i) = \infty$ then. as $n \to \infty$

$$\frac{\sum_{i=1}^{n} 1_{A_i}}{\sum_{i=1}^{n} P(A_i)} \to 1 \quad \text{a.s.}$$

[Generalize B-C (ii)] ~~[special case of SLLN]~~

[proof]: Let $S_n = \sum_{i=1}^{n} 1_{A_i}$. $b_n = \sum_{i=1}^{n} P(A_i)$

$$P\left(\left|\frac{S_n}{b_n} - 1\right| > \varepsilon\right) \le \frac{Var\left(\frac{S_n}{b_n}\right)}{\varepsilon^2} = \frac{\sum_{i=1}^{n} P(A_i)(1-P(A_i))}{\varepsilon^2 b_n^2}$$

$$\le \frac{\sum_{i=1}^{n} P(A_i)}{\varepsilon^2 \left(\sum_{i=1}^{n} P(A_i)\right)^2} = \frac{1}{\varepsilon^2 \sum_{i=1}^{n} P(A_i)} = \frac{1}{\varepsilon^2 b_n}$$

~~Choose $n_1 < n_2 < \dots$ s.t.~~ Choose $n_k = \min\{m : b_m \ge k^2\}$.

then $\sum_{k \ge 1}^{\infty} P\left(\left|\frac{S_{n_k}}{b_{n_k}} - 1\right| > \varepsilon\right) < \infty. \overset{B\text{-}C\,(i)}{\Longrightarrow} \frac{S_{n_k}}{b_{n_k}} \to 1$ a.s.

$\forall m.\ \exists k$ s.t. $n_{k-1} \le m \le n_k$.

$$\underset{\downarrow\,1}{\frac{S_{n_{k-1}}}{b_{n_{k-1}}}} \cdot \underset{\downarrow\,1}{\frac{b_{n_{k-1}}}{b_{n_k}}} = \frac{S_{n_{k-1}}}{b_{n_k}} \le \underset{\downarrow\,1\ \text{a.s.}}{\frac{S_m}{b_m}} \le \frac{S_{n_k}}{b_{n_{k-1}}} = \underset{\downarrow\,1}{\frac{S_{n_k}}{b_{n_k}}} \cdot \underset{\downarrow\,1}{\frac{b_{n_k}}{b_{n_{k-1}}}}$$

].

[subsequence method]

[Example 2.3.2. Record values Example 2.3.3. Head runs.]

B

④ Thm 2.3.5 Let $X_1, X_2, \ldots$ be i.i.d. with $EX_i = \mu$, $EX_i^4 < \infty$.

若 $S_n = X_1 + \cdots + X_n$, then $\frac{S_n}{n} \to \mu$ a.s

[proof: ~~$P(|\frac{S_n - n\mu}{n}|$~~ WLOG, assume $\mu = 0$.

Otherwise, consider $X_i' = X_i - \mu$.

$$P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \le \frac{E[S_n^4]}{\varepsilon^4 n^4} \le \frac{C\left(\sum_{i=1}^{n} E[X_i^4] + \sum_{i=1}^{n}\sum_{j=1}^{n} EX_i^2 X_j^2\right)}{\varepsilon^4 n^4}$$

$$\le \frac{C}{\varepsilon^4 n^2} \quad \text{summable.}$$]

- Rosenthal's inequality:

X_i indep. $EX_i = 0$. $S_n = \sum_{i=1}^{n} X_i$. $p \ge 2$. then.

$$E(|S_n|^p) \le C_p\left[(ES_n^2)^{p/2} + \sum_{i=1}^{n} E|X_i|^p\right].$$

Thm 2.4.1 SLLN.

Let $X_1, X_2, \ldots$ be pairwise indep. identically distributed. $E|X_i| < \infty$.

Let $EX_i = \mu$. and $S_n = X_1 + \cdots + X_n$. Then $\frac{S_n}{n} \to \mu$ a.s., as $n \to \infty$.

Remark. Suppose $\frac{S_n}{n} \to \mu$ as then

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{S_{n-1}}{n-1} \cdot \frac{n-1}{n} \to 0 \quad \text{as}$$

$\Leftrightarrow P\left(\left|\frac{X_n}{n}\right| > \varepsilon \text{ i.o.}\right) = 0 \underset{B\text{-}C}{\Leftrightarrow} \sum_{n=1}^{\infty} P\left(\frac{|X_n|}{n} > \varepsilon\right) < \infty \Rightarrow \sum_{n=1}^{\infty} P(|X_1| > n) < \infty.$

$\Rightarrow E|X_1| < \infty.$

~~Exercise~~ (consequence of the next theorem) [proof]: WLOG. assume $X_i \geq 0$ (o.w. $X_i = X_i^+ - X_i^-$).

idea: truncation ; second moment calculation : subsequence method.

Let $Y_k = X_k 1_{(|X_k| \leq k)}$. $T_n = Y_1 + \cdots + Y_n$.

$\because$ $\sum_{k=1}^{\infty} P(|X_k| > k) \leq \int_0^{\infty} P(|X_1| > t)\, dt = E|X_1| < \infty$

$\therefore$ $P(X_k \neq Y_k \text{ i.o.}) = 0$

$\therefore$ ~~$P(\lim \frac{S_n}{n}$~~ Suffices to prove $\boxed{\lim_{n\to\infty} \frac{T_n}{n} \to \mu \text{ a.s.}}$

Fix $\alpha > 1$. choose $k(n) = [\alpha^n]$, integer part $\geq \frac{\alpha^n}{2}$

$$\sum_{n=1}^{\infty} P\left(\left| \frac{T_{k(n)} - E T_{k(n)}}{k(n)} \right| > \varepsilon \right) \leq \varepsilon^{-2} \cdot \sum_{n=1}^{\infty} \frac{\operatorname{Var}(T_{k(n)})}{(k(n))^2} < \infty \,!!$$

~~$\operatorname{Var}(T_{k(n)})$~~

$$= \varepsilon^{-2} \sum_{n=1}^{\infty} \frac{1}{(k(n))^2} \cdot \sum_{m=1}^{k(n)} \operatorname{Var}(Y_m)$$

$$= \varepsilon^{-2} \sum_{m=1}^{\infty} \operatorname{Var}(Y_m) \cdot \sum_{n: k(n) \geq m} k(n)^{-2}$$

$$\leq \varepsilon^{-2} \sum_{m=1}^{\infty} \operatorname{Var}(Y_m) \cdot 4 \sum_{n: \alpha^n \geq m} \alpha^{-2n}$$

$$\leq 4(1-\alpha^{-2})^{-1} \cdot \varepsilon^{-2} \boxed{\sum_{m=1}^{\infty} \frac{\operatorname{Var}(Y_m)}{m^2}}$$

Note that
$$\sum_{m=1}^{\infty} \frac{\operatorname{Var}(Y_m)}{m^2} \leq \sum_{m=1}^{\infty} \frac{E Y_m^2}{m^2} = \sum_{m=1}^{\infty} \frac{\int_0^m 2y \cdot P(|Y_m| > y)\, dy}{m^2} \leq \sum_{m=1}^{\infty} \frac{\int_0^{\infty} 2y \, 1(y<m)\, P(|X_1| > y)\, dy}{m^2}$$

$$= \int_0^{\infty} \left\{ \sum_{\substack{m=1 \\ m > y}}^{\infty} \frac{1_{(y<m)}}{m^2} \right\} 2y \cdot P(|X_1| > y)\, dy$$

$$\leq C \int_0^{\infty} P(|X_1| > y)\, dy \leq C E|X_1| < \infty.$$

⑥

$$\Rightarrow \quad \left.\begin{array}{l} \dfrac{T_{k(n)} - ET_{k(n)}}{k(n)} \to 0 \quad a.s. \\[2ex] \dfrac{ET_{k(n)}}{k(n)} = \dfrac{\sum_{m=1}^{k(n)} EY_m}{k(n)} \to u \end{array}\right\} \Rightarrow \frac{T_{k(n)}}{k(n)} \to u \quad a.s.$$

$$EY_m = EX_1 I(|X_1| \le m) \to u.$$

For $k(n) \le m < k(n+1)$:

$$\frac{k(n)}{k(n+1)} \frac{T_{k(n)}}{k(n)} \le \frac{T_m}{m} \le \frac{T_{k(n+1)}}{k(n+1)} \cdot \frac{k(n+1)}{k(n)}$$

$\approx \frac{1}{\alpha}$ $\quad \downarrow u \quad\quad \downarrow u \quad\quad \approx \alpha$

$$\Rightarrow \quad \limsup_{n\to\infty} \frac{T_m}{m} \le u \cdot \alpha.$$

$$\liminf_{n\to\infty} \frac{T_m}{m} \ge u \cdot \frac{1}{\alpha}$$

Let $\alpha \to 1$.]

Thm: $X_1, X_2 \cdots$ i.i.d. $EX_i^+ = \infty$. $EX_i^- < \infty$.

If $S_n = X_1 + \cdots + X_n$, then $S_n/n \to \infty$ a.s.

[proof. Let $X_i^M = X_i \wedge M$. $E|X_i^M| < \infty$.

By SLLN, If $S_n^M = X_1^M + \cdots + X_n^M$, then $\frac{S_n^M}{n} \to EX_i^M$.

$$\liminf_{n\to\infty} \frac{S_n}{n} \ge \lim_{n\to\infty} \frac{S_n^M}{n} = EX_i^M \uparrow \infty \quad \text{as } M \to \infty$$

$$\therefore \liminf_{n\to\infty} \frac{S_n}{n} \ge \infty \Rightarrow \frac{S_n}{n} \to \infty. \quad]$$

⑦ Applications (Empirical distribution functions).

$X_1, X_2, \ldots$ i.i.d. from a population distribution $F(x)$.

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1(X_i \le x)$$

by SLLN, $F_n(x) \to F(x)$ a.s. $\forall x$.

Thm 2.4.7 The Glivenko–Cantelli theorem. As $n \to \infty$

$$\sup_x |F_n(x) - F(x)| \to 0 \quad \text{a.s.}$$

1. $\boxed{\text{Lecture 6}}$

Ch2-4: sums of indep. r.v.'s
Ch5: martingales
~~next sem.:~~ [Shao:] self-normalized r.v.'s - eg. t-statistics   results on rate of convergence

Thm 2.5.2   Kolmogorov's maximal inequality

$X_1, X_2, \dots$ indep. $EX_i = 0$, $EX_i^2 < \infty$, $\forall i$. $S_n = \sum_{i=1}^n X_i$.

then: $\forall x > 0$, $P(\max_{1\le k\le n} |S_k| \ge x) \le \frac{E(S_n^2)}{x^2}$

· similar techniques in future

{proof: Let $A_k := \{|S_i| < x,\ i < k,\ |S_k| \ge x\}$, disjoint for k.

Then. $\{\max_{1\le k\le n} |S_k| \ge x\} = \bigcup_{k=1}^n A_k$

$P(\max_{1\le k\le n} |S_k| \ge x) = \sum_{k=1}^n P(A_k) \le \sum_{k=1}^n \frac{E(S_k^2 1_{A_k})}{x^2} \overset{?}{\le} \sum_{k=1}^n \frac{E(S_n^2 1_{A_k})}{x^2}$

$E(S_n^2 1_{A_k}) = E((S_k + S_n - S_k)^2 1_{A_k})$

$= E(S_k^2 1_{A_k}) + \underbrace{2E[S_k (S_n - S_k) 1_{A_k}]}_{=0} + \underbrace{E[(S_n - S_k)^2 1_{A_k}]}_{\ge 0}$

$\ge E(S_k^2 1_{A_k})$

2. Theorem 2.5.3: $X_1, X_2, \ldots$ independent, $EX_i = 0$ $\forall i$.

If $\sum_{i=1}^{\infty} EX_i^2 < \infty$ then $\sum_{i=1}^{\infty} X_i$ converges a.s. [i.e., $\sum_{i=1}^{\infty} X_i(\omega)$ converges a.s.]

[proof: let $S_n = \sum_{i=1}^{n} X_i$.

~~Idea~~ same as: prove that a.s., S_n is a Cauchy sequence, i.e.,

$\cancel{\forall \varepsilon > 0, \exists M, \text{ s.t. }}$ $\underbrace{\sup_{m,n \geq M} |S_m - S_n|}_{w_M :=} \downarrow 0$ a.s. as $M \uparrow \infty$.

$\forall \varepsilon > 0$,

$$P(w_M > 2\varepsilon \text{ i.o.}) = P(w_M > 2\varepsilon, \forall M) = P\left(\bigcap_M \{w_M > 2\varepsilon\}\right) = \lim_{M\to\infty} P(w_M > 2\varepsilon)$$

$$P\left(\sup_{m,n \geq M} |S_m - S_n| > 2\varepsilon\right)$$

$$\leq P\left(\sup_{m \geq M} |S_m - S_M| > \varepsilon\right) + P\left(\sup_{n \geq M} |S_n - S_M| > \varepsilon\right)$$

$$= 2P\left(\sup_{m \geq M} |S_m - S_M| > \varepsilon\right)$$

$$P\left(\sup_{m \geq M} |S_m - S_M| > \varepsilon\right) = \lim_{N\to\infty} P\left(\sup_{M \leq m \leq N} |S_m - S_M| > \varepsilon\right)$$

$$\overset{\text{Theorem 2.5.2}}{\leq} \limsup_{N\to\infty} \frac{Var(S_N - S_M)}{\varepsilon^2} = \frac{1}{\varepsilon^2} \sum_{i=M+1}^{\infty} EX_i^2$$

$\longrightarrow 0$ as $M \to \infty$.]

3

Theorem. 2.5.4 Kolmogorov's three-series theorem

$X_1, X_2 \ldots$ indep. Let $A>0$. ~~$Y_i = X_i I_{(|X_i| \le A)}$~~.

If (i). $\sum_{n=1}^{\infty} P(|X_n|>A)<\infty$.

(ii) $\sum_{n=1}^{\infty} EY_n$ converges

(iii) $\sum_{n=1}^{\infty} Var(Y_n) < \infty$.

Then. $\sum_{n=1}^{\infty} X_n$ converges a.s.

- (i), (ii), (iii) are ~~actually~~ also necessary conditions

[proof: let $\mu_n = EY_n$.

From Thm 2.5.3 & (iii): $\sum_{n=1}^{\infty} (Y_n - \mu_n)$ converges a.s.

+ (ii). : $\sum_{n=1}^{\infty} Y_n$ converges a.s.

(i) + B.C : $P(X_n \neq Y_n \text{ i.o.}) = 0$.

$\Rightarrow \sum_{n=1}^{\infty} X_n$ converges a.s.

Next Example first

Theorem 2.5.5. Kronecker's lemma

If $a_n \uparrow \infty$, $\sum_{n=1}^{\infty} \frac{x_n}{a_n}$ converges. then $\frac{\sum_{i=1}^{n} x_n}{a_n} \to 0$.

[proof omitted]

Example 2.5.3: $\{X_1, X_2, \ldots\}$ i.i.d. $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$.

Thm 2.5.3 $\Rightarrow$ If $\alpha > \frac{1}{2}$, then $\sum_{i=1}^{\infty} \frac{X_i}{i^\alpha}$ converges a.s.

Thm 2.5.4. (necessity part; (iii)) $\Rightarrow$ vice versa.

Another proof for SLLN: $X_1, X_2, \ldots$ i.i.d. $E|X_i| < \infty$ $\forall i$.

$EX_i = u$ $S_n = X_1 + \cdots + X_n$. then $\frac{S_n}{n} \to u$ a.s. $n \to \infty$.

Proof: Let $Y_k = X_k 1(|X_k| \le k)$. $T_n = Y_1 + \cdots + Y_n$.

~~By P(We have shown that~~

Recall $P(X_i \neq Y_i \text{ a.i.o.}) = 0$.

Only N.T.S. $\frac{T_n}{n} \to u$ a.s.

Recall $\frac{ET_n}{n} \to u$.

Only N.T.S. $\frac{T_n - ET_n}{n} \to 0$ a.s., i.e.,

$$\sum_{k=1}^{n} \frac{Y_k - EY_k}{n} \to 0 \quad \text{a.s.}$$

By Kronecker's lemma: Only N.T.S. $\sum_{k=1}^{\infty} \frac{Y_k - EY_k}{k}$ converges

By Theorem 2.5.4: Recall $\sum_{k=1}^{\infty} \mathrm{Var}\left(\frac{Y_k - EY_k}{k}\right) = \sum_{k=1}^{\infty} \frac{EY_k^2}{k^2} < \infty$. ■

5. Thm 2.5.8 Marcinkiewicz–Zygmund SLLN.

$X_1, X_2, \ldots$ i.i.d. $EX_1 = 0$ (μ). $E|X_1|^p < \infty$ ($E|X_1-\mu|^p<\infty$) for some $1<p<2$.

then with $S_n = X_1 + \cdots + X_n$, $\dfrac{S_n}{n^{1/p}} \to 0$ a.s. ($\dfrac{S_n - n\mu}{n^{1/p}}$, Exercise)

~~proof~~ • $\boxed{\text{Exercise:}}$ prove that $\dfrac{S_n}{n^{1/p}} \to 0$ a.s. $\Rightarrow$ $E|X_1|^p<\infty$.

$\boxed{\text{Exercise:}}$ i.i.d. If $E|X_1|^p<\infty$, $0<p<1$ (mean may not exist) then $\dfrac{S_n}{n^{1/p}} \to 0$ a.s.

proof: ~~let $Y_k = X_k I(|X_k| \le k^{1/p})$.~~

$$\cancel{\sum_{k=1}^{\infty} P(Y_k \ne X_k) = \sum_{k=1}^{\infty} P(|X_k|^p > k) \le E|X_k|^p < \infty.}$$

$$\cancel{\therefore P(Y_k \ne X_k \text{ i.o.}) = 0.}$$

Suffices to prove $\displaystyle\sum_{n=1}^{\infty} \frac{X_n}{n^{1/p}}$ converges a.s.

Choose $A=1$ in Thm 2.5.4 : $Y_i = \dfrac{X_i}{i^{1/p}} I(|X_i| \le i^{1/p})$

(i) $\displaystyle\sum_{n=1}^{\infty} P\left(\left|\frac{X_n}{n^{1/p}}\right| > 1\right) = \sum_{n=1}^{\infty} P(|X_n|^p > n) \le E|X_1|^p < \infty$

(ii) $\displaystyle\sum_{i=1}^{\infty} |EY_i| = \sum_{i=1}^{\infty} \frac{1}{i^{1/p}} \left|EX_i I(|X_i| \le i^{1/p})\right|$

From $EX_i=0$

$$= \sum_{i=1}^{\infty} \frac{1}{i^{1/p}} \left|EX_i I(|X_i|^p > i)\right| \le \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} \frac{1}{i^{1/p}} E|X_1| I(j \le |X_1|^p < j+1)$$

$$= \sum_{j=1}^{\infty} \sum_{i=1}^{j} \frac{1}{i^{1/p}} E|X_1| I(j \le |X_1|^p < j+1) \le C_p \sum_{j=1}^{\infty} \boxed{j^{1-\frac{1}{p}}} E|X_1| I(j \le |X_1|^p \le j+1)$$

$\approx (|X_1|^p)^{1-\frac{1}{p}}$

$$\le C_p E|X_1|^p < \infty$$

6.

(iii) $\sum_{i=1}^{\infty} E Y_i^2 = \sum_{i=1}^{\infty} \frac{1}{i^{2/p}} E\left(|X_i|^2 \cdot I\left(|X_i| \le i^{1/p}\right)\right)$

$$= \sum_{i=1}^{\infty} \frac{1}{i^{2/p}} \sum_{j=1}^{i} E\left(|X_1|^2 I\left((j-1)^{1/p} < |X_1| \le j^{1/p}\right)\right)$$

$$= \sum_{j=1}^{\infty} \sum_{i=j}^{\infty} \frac{1}{i^{2/p}} E\left(|X_1|^2 I\left((j-1)^{1/p} < |X_1| \le j^{1/p}\right)\right)$$

$$\le C_p \sum_{j \ge 1} j^{1-\frac{2}{p}} \cdot E\left(|X_1|^2 I\left((j-1)^{1/p} < |X_1| \le j^{1/p}\right)\right)$$

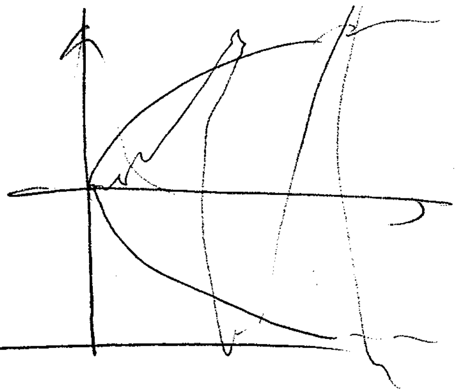
$|X_1|^{p-2+2}$

$$\le C_p E|X_1|^p < \infty. \qquad \blacksquare \;]$$

LIL

Thm 8.8.3 $X_1, X_2, \ldots$ i.i.d. $EX_1 = 0$. $EX_1^2 = \sigma^2 > 0$.

then
$$\limsup_{n\to\infty} \frac{S_n}{\sigma\sqrt{n}\cdot\sqrt{2\log\log n}} = 1 \quad \text{a.s.}$$
$$\liminf_{n\to\infty} \frac{S_n}{\sigma\sqrt{n}\cdot\sqrt{2\log\log n}} = -1 \quad \text{a.s.}$$



Let $W_n = \frac{S_n}{\sigma\sqrt{n}}$, F_n be the d.f. of W_n

then $F_n \to \Phi$. (CLT)

$n \cdot o(\frac{1}{n}) \to 0$

$W_n = \sum_{i=1}^{n} \xi_i$, where $\xi_i = \frac{X_i}{\sigma\sqrt{n}}$, $E\xi_i = 0$, $E\xi_i^2 = \frac{1}{n}$

$N(0,1) \sim Z = \sum_{i=1}^{n} \eta_i$ $\quad \eta_i \sim N(0, \frac{1}{n})$. $E\eta_i = 0$, $E\eta_i^2 = \frac{1}{n}$

$\mathcal{L}(W_n) \Rightarrow \mathcal{L}(Z)$ as $n \to \infty$. ?

7.

Ch3 — converge in distribution; characteristic function; CLT.

$\downarrow$

Special case: Thm 3.1.3 De Moivre–Laplace theorem.

$X_1, X_2, \cdots$ i.i.d. $P(X_1=1)=P(X_1=-1)=\frac{1}{2}$.

$S_n = X_1 + \cdots + X_n$.

$$P\left(a \le \frac{S_n}{\sqrt{n}} \le b\right) \to \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx.$$

Def: F_n is a sequence of d.f.'s, F is a d.f.

F_n is said to converge weakly to F, denoted by $F_n \Rightarrow F$,

if $F_n(y) \to F(y)$ for every continuity point of F

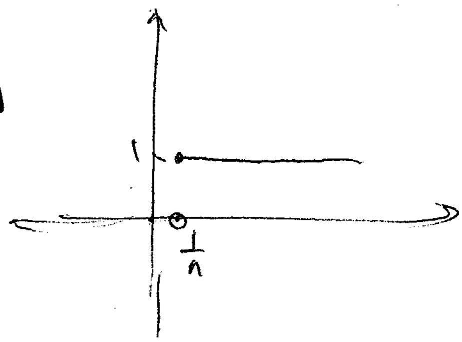
Def: X_n is a sequence of r.v.'s, X is a r.v.

X_n converges to X weakly (or in distribution), denoted by $X_n \Rightarrow X$ or $X_n \xrightarrow{d} X$,

if $F_{X_n} \Rightarrow F_X$.

- $X_n \Rightarrow X$. the r.v.'s may not be defined on the same probability space
- $X_n \to X$ in prob $\Rightarrow$ $X_n \Rightarrow X$

  $[\ F(x-\varepsilon) \le \liminf_{n\to\infty} F_n(x) \le \limsup_{n\to\infty} F_n(x) < F(x+\varepsilon)\ ]$
- $X + \frac{1}{n} \Rightarrow X$. $\leftarrow$



- Glivenko–Cantelli theorem $F_n \Rightarrow F$.

8. Thm 3.2.2. If $F_n \Rightarrow F_\infty$, then there are r.v.'s Y_n, $1 \le n \le \infty$.

Skorokhod's theorem: with d.f. F_n s.t. $Y_n \to Y_\infty$ a.s.

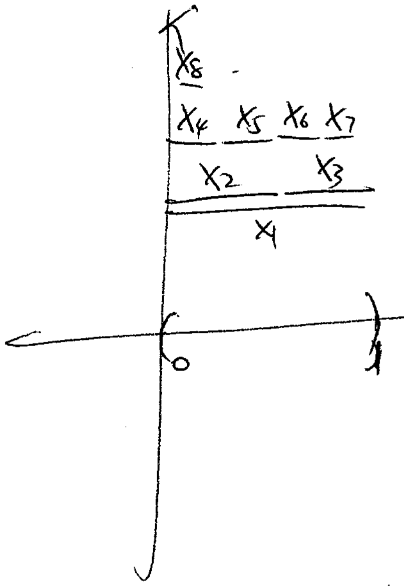
[proof: ~~recall~~ let $\Omega = (0,1)$, $\mathcal{F}$ Borel sets, P Lebesgue measure.

~~recall $F_n^{-1}(\omega)$ is a r.v. with distri~~

let $F_n^{-1}(x) = \sup_y \{y : F_n(y) < x\}$.

recall. $F_n^{-1}(\omega)$ follows d.f. F_n.

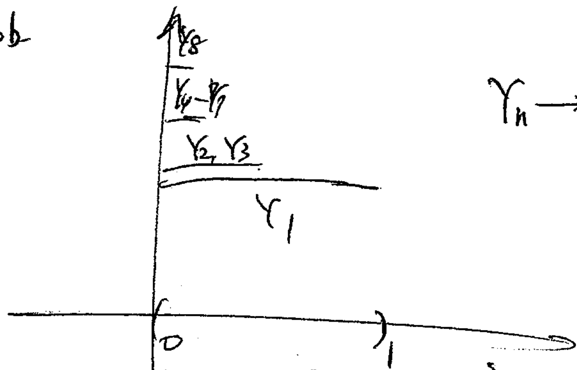
......]



$X_n \not\to 0$ a.s.

$X_n \to 0$ in prob.

$X_n \Rightarrow 0$



$Y_n \to 0$ a.s.

Exercise. • Fatou's Lemma: let $g \ge 0$ be continuous, If $X_n \Rightarrow X_\infty$, then $\liminf_{n\to\infty} Eg(X_n) \ge Eg(X_\infty)$.

Slutsky's theorem:

1. If $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} c$. (c is a constant)
   then $X_n + Y_n \xrightarrow{d} X + c$, $X_n \cdot Y_n \xrightarrow{d} cX_n$.

[statistical inference with restricted ~~parameter~~ nuisance parameter].

2. If $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{d} X$.

3. If $X_n \xrightarrow{d} c$ then $X_n \xrightarrow{P} c$.

• $\dfrac{\sum X_i}{\hat{\sigma}\sqrt{n}} \Rightarrow N(0,1)$

Sample standard deviation

9. [proof:

2. ✓.

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

$P(X_n - c \leq -\varepsilon) = F_n(c-\varepsilon) \to 0.$

1. Let $x-c$ be a continuity point of F_X, choose $\varepsilon \downarrow 0$ s.t. $x-c+\varepsilon$ be continuity points of F_X

we can do this because continuity points is a dense set in R

- $P(X_n + Y_n \leq x) \leq P(|Y_n - c| \geq \varepsilon) + P(X_n + Y_n \leq x, |Y_n - c| < \varepsilon)$

$\leq P(|Y_n - c| \geq \varepsilon) + P(X_n \leq x - c + \varepsilon)$

$\xrightarrow{n\to\infty} P(X \leq x - c + \varepsilon) = P(X + c \leq x + \varepsilon) \downarrow_{\varepsilon \downarrow 0} P(X + c \leq x)$

$P(X_n + Y_n \leq x) \geq P(X_n + c \leq x + \varepsilon) - P(|Y_n - c| > \varepsilon)$

....

- : $c > 0$, $0 < \varepsilon < c$

$P(X_n Y_n \leq x) \leq P(|Y_n - c| > \varepsilon) + P(X_n \leq \frac{x}{c-\varepsilon})$

$c < 0$, $c = 0$, ..

]

1. $\boxed{\text{Lecture 7}}$

$F_n \Rightarrow F \quad (F_n(y) \to F(y) \quad \text{if} \quad F(y-) = F(y))$

$X_n \Rightarrow X$

- $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X.$

- $F_n \Rightarrow F$

$\quad \sim \quad \sim$

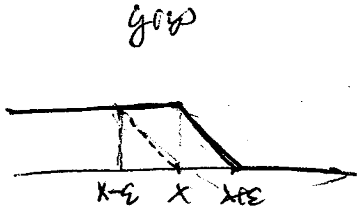
$\exists \; Y_n \to Y$ a.s.

test function $\quad C_b$

<u>Thm 3.23</u> $X_n \Rightarrow X$ iff $\forall \; g: \mathbb{R} \to \mathbb{R}$ bounded continuous

$$E g(X_n) \to E g(X).$$

[proof. "$\Leftarrow$" let $g(y) = \begin{cases} 1 & \text{if } y \le x \\ 0 & \text{if } y \ge x+\varepsilon \\ \text{linear} & \text{if } x \le y \le x+\varepsilon \end{cases}$



$\forall \; x$ being continuity point of F:

$\because F_n(x) = E\,1(X_n \le x) \le E g(X_n) \xrightarrow{n\to\infty} E g(X) \le E\,1(X \le x+\varepsilon) = F(x+\varepsilon)$

$\downarrow \varepsilon \downarrow 0$

$F(x)$

$\therefore \limsup_{n\to\infty} F_n(x) = F(x)$

lower bound is proved similarly.

"$\Rightarrow$" Let $Y_n \stackrel{d}{=} X_n$, $Y \stackrel{d}{=} X$. and $Y_n \to Y$ a.s

(by Skorokhod's ~~embedding~~ theorem)

$\cancel{\text{Only}}$ s $E g(Y_n) \to E g(Y) = E g(X).$

$E g(X_n) \quad \uparrow \text{BCT}$]

2\. ~~Thm 3.2.4~~

- $C_b(\mathbb{R})$ bounded ~~convex~~ continuous functions

$$\begin{array}{l} C_b(\mathbb{R}) \\ \cup \\ C_b^1 : \\ \cup \\ C_b^k : \\ \cup \\ C_b^\infty ; \end{array}$$

C_b^∞: ~~all de~~ $\forall k$, kth order derivative exists and bounded by $M_k < \infty$.

$$X_n \Rightarrow X \iff Eg(X_n) \to Eg(X). \quad \forall g \in C_b^\infty \quad ?$$

~~Thm 3.2.4: Continuous mapping theorem~~

CLT: $X_1, X_2, \dots$ i.i.d.

$EX_i = \mu$, $EX_i^2 = \sigma^2$, $E|X_i|^3 < \infty$. (can be removed by truncation?)

$W_n = \sum_{i=1}^n \frac{X_i - \mu}{\sigma\sqrt{n}}$. Then $W_n \xrightarrow{d} Z \sim N(0,1)$.

[proof: N.T.S $\forall g \in C_b^\infty$, $Eg(W_n) \longrightarrow Eg(Z)$ $n \to \infty$.

~~Write~~ Let $\xi_i := \frac{X_i - \mu}{\sigma\sqrt{n}}$. $\left(E\xi_i = 0.\ E\xi_i^2 = \frac{1}{n}.\ E|\xi_i|^3 \le \frac{C}{n^{3/2}}\right)$

Let $\eta_i \overset{i.i.d.}{\sim} N(0, \frac{1}{n})$ $\left(E\eta_i = 0.\ E\eta_i^2 = \frac{1}{n}.\ E|\eta_i|^3 \le \frac{C}{n^{3/2}}\right)$

$$Eg(W_n) - Eg(Z) = Eg\left(\sum_{i=1}^n \xi_i\right) - Eg\left(\sum_{i=1}^n \eta_i\right)$$

$$= Eg(\xi_1 + \dots + \xi_n) - Eg(\xi_1 + \dots + \xi_{n-1} + \eta_n) + Eg(\xi_1 + \dots + \xi_{n-1} + \eta_n)$$

$$- Eg(\xi_1 + \dots + \xi_{n-2} + \eta_{n-1} + \eta_n) + \dots + Eg(\xi_1 + \eta_2 + \dots + \eta_n) - Eg(\eta_1 + \dots + \eta_n)$$

$\therefore 3.$

$$= \sum_{k=1}^{n} \left[Eg(\xi_1 + \cdots + \xi_{k-1} + \xi_k + \eta_{k+1} + \cdots + \eta_n) - Eg(\xi_1 + \cdots + \xi_{k-1} + \eta_k + \eta_{k+1} + \cdots + \eta_n) \right]$$

$$\cancel{= E \sum_{k=1}^{n} \left[(\xi_k - \eta_k) \cdot g'(\xi_1 + \cdots + \xi_{k-1} + \eta_{k+1} + \cdots + \eta_k) + \frac{(\xi_k - \eta_k)^2}{2} g''(\xi_1 + \cdots + \xi_{k-1} + \eta_{k+1} + \cdots + \eta_k) + O(|\xi_k - \eta_k|^3) \right]}$$

$$= E \sum_{k=1}^{n} \Big[g(\underbrace{\xi_1 + \cdots + \xi_{k-1} + \eta_{k+1} + \cdots + \eta_k}_{V_k}) + \xi_k \cdot g'(V_k) + \frac{\xi_k^2}{2} g''(V_k) + O(|\xi_k|^3)$$

$$- g(V_k) - \eta_k g'(V_k) - \frac{\eta_k^2}{2} g''(V_k) + O(|\eta_k|^3) \Big]$$

$\underline{\underline{V_k \perp\!\!\!\perp \xi_k,\ V_k \perp\!\!\!\perp \eta_k}}$ $\quad O\left(\frac{1}{\sqrt{n}}\right) \longrightarrow 0$ as $n \to \infty$. ∎]

~~example~~: RAP ?

Thm 3.2.4. Continuous mapping theorem

Let g be a measurable function and $D_g = \{x : g \text{ is discontinuous at } x\}$.

If $X_n \Rightarrow X$ and $P(X \in D_g) = 0$, then

$g(X_n) \Rightarrow g(X)$

[proof: by Skorokhod's theorem, construct $Y_n \to Y$ a.s. and $Y_n \overset{d}{=} X_n$, $Y \overset{d}{=} X$

then $g(Y_n) \to g(Y)$ a.s.

which implies

$$\begin{array}{ccc} g(Y_n) & \Rightarrow & g(Y) \\ \overset{d}{\parallel} & & \overset{d}{\parallel} \\ g(X_n) & & g(X) \end{array}$$

∎]

4. **Thm 3.2.5** Portmanteau theorem

The following are equivalent

(i) $X_n \Rightarrow X_\infty$

(ii) $\forall$ open sets G, $\liminf_{n\to\infty} P(X_n \in G) \geq P(X_\infty \in G)$

(iii) $\forall$ closed sets K, $\limsup_{n\to\infty} P(X_n \in K) \leq P(X_\infty \in K)$

(iv) $\forall A$ with $P(X_\infty \in \partial A) = 0$, $\lim_{n\to\infty} P(X_n \in A) = P(X_\infty = A)$

[proof: (i) $\Rightarrow$ (ii): $Y_n \overset{d}{=} X_n$, $Y_n \to Y_\infty$ a.s.

$$\cancel{1(Y_n \in G) \Rightarrow 1(Y_\infty \in G)}$$

$$\liminf_{n\to\infty} P(X_n \in G) = \liminf_{n\to\infty} P(Y_n \in G) = \liminf_{n\to\infty} E\,1(Y_n \in G) \overset{\text{Fatou}}{\geq} E\,\boxed{\liminf 1(Y_n \in G)}$$

$$\cancel{\geq P(Y_\infty \in G) = P(X_\infty \in G)}$$

$$P(X_\infty \in G) = P(Y_\infty \in G) = E\,\boxed{1(Y_\infty \in G)}$$

(ii) $\Leftrightarrow$ (iii). $K = G^c$

(ii), (iii) $\Rightarrow$ (iv): $A^\circ \subset A \subset \bar{A}$

(iv) $\Rightarrow$ (i): $A = (-\infty, x]$]

5. Thm 3.2.6. Helly's selection theorem

every sequence of d.f.'s F_n has a subsequence $F_{n(k)}$

converging vaguely to a right continuous nondecreasing function F.

$$\lim_{k\to\infty} F_{n(k)}(y) = F(y) \quad \forall\, y : F(-y) = F(y)$$

Thm 3.2.7. Every subsequential limit is a d.f

iff $\{F_n\}$ is tight, i.e., $\forall \varepsilon > 0 \quad \exists M_\varepsilon$.
$\{X_n\}$

$$P(|X_n| \geq M_\varepsilon) \leq \varepsilon. \quad \forall n.$$

~~Thm 3.2.8~~ example. if $E|X_n| \leq C < \infty$. then $\{X_n\}$ is tight

---

Distances between distribution functions

~~Thm~~ Def: Total variation distance between two probability measures μ_1, μ_2 is defined as

$$\sup_{A\in\mathcal{F}} |\mu_1(A) - \mu_2(A)| =: d_{TV}(\mu_1, \mu_2)$$

- ~~$d_{TV}(\mu_1,\mu_2)$~~ $d_{TV}(\mu_n, \mu_\infty) \to 0 \Rightarrow \mu_n \xrightarrow{d} \mu_\infty.$
  $\not\Leftarrow$

Thm: f_n is a sequence of pdfs. f is a pdf $X_n \sim f_n$, $X \sim f$.

if $f_n(x) \to f(x) \quad \forall x \in \mathbb{R}$.

~~then $X_n \to X$~~ then. $d_{TV}(\mathcal{L}(X_n), \mathcal{L}(X)) \to 0.$

6. [proof:

MTS. $\left|\int_A f_n(x)\,dx - \int_A f(x)\,dx\right| \to 0 \quad \forall A \in \mathcal{B}$

$\wedge\!\!|$

$\int_{-\infty}^{\infty} |f_n(x) - f(x)|\,dx \xrightarrow{DCT} 0$

$\|$

$2\int_{-\infty}^{\infty} (f(x) - f_n(x))^+\,dx$

.]

Def. Kolmogorov distance between two df's F_1, F_2 is

$$d_K(F_1, F_2) := \sup_x |F_1(x) - F_2(x)|.$$

$$d_K(\mathcal{L}(X_1), \mathcal{L}(X_2)) = \sup_x |P(X_1 \le x) - P(X_2 \le x)|$$

- $d_{TV} \to 0 \Rightarrow d_K \to 0$
  $\not\Leftarrow$
- if F is continuous, $d_K \to 0 \Rightarrow F_n \Rightarrow F$.

rate of convergence provides an explicit upper bound on $d(F_n, F)$.

Thm 3.4.9. (Berry-Esseen theorem)

$X_1, X_2, \ldots$ i.i.d. $EX_i = 0$. $EX_i^2 = \sigma^2$. $E|X_i^3| = \gamma < \infty$.

$W_n = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$. Φ standard normal distribution function

Then $\sup_{x\in\mathbb{R}} |P(W_n \le x) - \Phi(x)| \le 1 \frac{\gamma}{\sigma^3\sqrt{n}}$

7. Characteristic functions (ch.f.)

Def. ch.f. of a r.v. X is $\varphi_X(t) = E e^{itX} = E\cos tX + iE\sin tX$

well defined for any X.

properties:

1. $\varphi_X(0) = 1$ $\quad |\varphi_X(t)| \le 1$

2. $\varphi_X(-t) = \overline{\varphi_X(t)}$

3. $|\varphi_X(t+h) - \varphi_X(t)| \le E|e^{ihX} - 1| \to 0$ as $h \to 0$

   $\therefore$ $\varphi(t)$ is uniformly continuous on $(-\infty, \infty)$.

4. $\varphi_{aX+b}(t) = e^{itb} \cdot \varphi_X(at)$.

5. if $X_1 \perp\!\!\!\perp X_2$,

   then. $\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t) \cdot \varphi_{X_2}(t)$

6. If $E(X^2) < \infty$. then. $\varphi_X(t) = 1 + itE(X) - \frac{t^2}{2}E(X^2) + o(t^2)$ as $t \to 0$.

Examples:

- $X \sim \text{Poi}(\lambda)$, $\varphi_X(t) = e^{\lambda(e^{it}-1)}$
- $X \sim N(0,1)$, $\varphi_X(t) = e^{-\frac{t^2}{2}}$
- $X \sim N(\mu, \sigma^2)$. $\varphi_X(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$
- X with pdf $(1-\cos x)/\pi x^2$ $\quad \varphi_X(t) = (1-|t|)^+$
- X with pdf $1/\pi(1+x^2)$, $\varphi_X(t) = e^{-|t|}$
- $X \sim \alpha$ stable distribution, $\varphi_X(t) = e^{-|t|^\alpha}$. $0 < \alpha \le 2$.

8. Thm 3.3.4. The inversion formula.

Let $\varphi_X(t) = E(e^{itX})$

then for $a<b$.

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

[proof:

$$\frac{e^{-ita}-e^{-itb}}{it} = \int_a^b e^{-ity}\,dy \overset{|\cdot|}{\le} b-a$$

$$\frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\cdot\varphi(t)\,dt = \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it} E e^{itX}\,dt$$

$$\overset{\text{Fubini}}{=} E\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{it(X-a)}-e^{it(X-b)}}{it}\,dt$$

$$= E\frac{1}{2\pi}\int_{-T}^{T}\left[\cancel{\frac{\cos(t(X-a))-\cos(t(X-b))}{it}} + \frac{\sin(t(X-a))-\sin(t(X-b))}{t}\right]dt$$

$$= E\frac{1}{\pi}\int_0^T \frac{\sin(t(X-a))-\sin(t(X-b))}{t}\,dt$$

Facts: Exercise 1.7.5.
$$\lim_{T\to\infty}\int_0^T \frac{\sin(t\cdot c)}{t}\,dt = \begin{cases} \frac{\pi}{2} & c>0 \\ 0 & c=0 \\ -\frac{\pi}{2} & c<0 \end{cases}, \quad \left|\int_0^T \frac{\sin(t\cdot c)}{t}\,dt\right| < 4$$

$1(a<X<b)$: 1 $\quad$ $1(X=a)$: $\frac{1}{2}$

$1(X<a)$: 0 $\quad$ $1(X=b)$: $\frac{1}{2}$

$1(X>b)$: 0]

9. ~~Thm~~ :

~~For~~ Exercise 3.3.2, If X is integer valued, then.

$$P(X=x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itx}\varphi(t)\,dt \qquad \forall x \in \mathbb{Z}.$$

Thm 3.3.5 If $\int |\varphi_X(t)|dt < \infty$, then X has bounded ~~0~~ continuous density $f(x) = \frac{1}{2\pi}\int e^{-itx}\varphi(t)\,dt$.

Thm 3.3.6. (i). If $X_n \xrightarrow{d} X$ then. $\varphi_{X_n}(t) \to \varphi_X(t)$ $\forall t \in \mathbb{R}$

[e^{itx} is bdd continuous]

(ii). If $\varphi_{X_n}(t) \to \varphi(t)$ $t \in \mathbb{R}$

and φ is continuous at 0.

then φ is a ch.f. of some r.v. X, and $X_n \xrightarrow{d} X$.

~~[(i)]~~ [proof omitted]

~~proof sketch : (ii) $\{X_n\}$ is tight~~

10. 2nd proof of CLT: $X_1, X_2, \ldots$ i.i.d.

$EX_i = 0$. $EX_i^2 = \sigma^2$. $W_n = \dfrac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$.

$W_n \xrightarrow{d} N(0,1)$.

[proof: ch.f. of W_n, denoted by φ_n is [recall. $\varphi_X(t) = 1 + it\,EX - \frac{t^2}{2} EX^2 + o(t^2)$] as $t \to 0$

$$\varphi_n(t) = \prod_{i=1}^{n} E e^{it\frac{X_i}{\sigma\sqrt{n}}} = \prod_{i=1}^{n} \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n^2}\right)\right)$$

$$= \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n^2}\right)\right)^n$$

$$\to e^{-\frac{t^2}{2}}$$]

Sums of independent r.v.'s ?

not necessarily identically distributed

$\boxed{\text{Lecture 8}}$

1. Thm 3.4.5 The Lindeberg-Feller theorem.

For each n, $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nn}$ are independent.

$$E\xi_{ni} = 0 \quad \text{for } m=1, \ldots, n.$$

$$E\sum_{i=1}^{n} \xi_{ni}^2 = 1.$$

If $\forall \varepsilon > 0$, $\sum_{i=1}^{n} E\xi_{ni}^2 1_{(|\xi_{ni}| > \varepsilon)} \to 0$ (Lindeberg's condition)

then. $\sum_{i=1}^{n} \xi_{ni} \xrightarrow{d} N(0,1)$.

- CLT for standardized sums of an i.i.d. sequence of r.v.'s is a special case:

$X_1, X_2, \ldots$ i.i.d. $EX_i = \mu$ $EX_i^2 = \sigma^2$

$$W_n = \sum_{i=1}^{n} \underbrace{\frac{X_i - \mu}{\sigma\sqrt{n}}}_{=\xi_{ni}}$$

Verify Lindeberg's condition: $\sum_{i=1}^{n} E\xi_{ni}^2 1(|\xi_{ni}| > \varepsilon)$

$$= n \cdot E\left(\frac{X_1 - \mu}{\sigma\sqrt{n}}\right)^2 1\left(\left|\frac{X_1-\mu}{\sigma\sqrt{n}}\right| > \varepsilon\right) = E\frac{1}{\sigma^2}\cdot (X_1-\mu)^2 \cdot 1(|X_1 - \mu| > \varepsilon\cdot\sigma\cdot\sqrt{n})$$

$$\to 0$$

- $\sum_{i=1}^{n} E|\xi_{ni}|^p \to 0$, for some $p > 0$

implies Lindeberg's condition

2. ~~proof:~~

- Lindeberg's condition implies

$$\max_{1\le i\le n} E\xi_{ni}^2 \to 0 \quad \text{as} \quad n\to\infty$$

[proof: $\cancel{\max_{1\le i\le n} E\xi_{ni}^2 1(|\xi_{ni}|>\varepsilon) \to 0}$
by contradiction

$$\sum_{i=1}^{n} E\xi_{ni}^2 1(|\xi_{ni}|>\varepsilon) \ge \max_{1\le i\le n} E\xi_{ni}^2 1(|\xi_{ni}|>\varepsilon) \qquad (*)$$

$$\ge \max_{1\le i\le n} (E\xi_{ni}^2) - \varepsilon^2$$

If $\max_{1\le i\le n} E\xi_{ni}^2 \not\to 0$, $\exists\ \varepsilon>0$ s.t.

$$\max_{1\le i\le n} E\xi_{ni}^2 > 2\varepsilon^2 \quad \text{i.o.}$$

then. from $(*)$ $\sum_{i=1}^{n} E\xi_{ni}^2 1(|\xi_{ni}|>\varepsilon) > \varepsilon^2$ i.o.

this contradicts with Lindeberg's condition,]

- $\max_{1\le i\le n} E\xi_{ni}^2 \to 0$ is a necessary condition

for $\cancel{\sum_{i=1}^{n}\xi_{ni}}$ CLT to hold.

3. [proof of CLT : by chf.

Let $\varphi(t) = E e^{it \sum_{i=1}^{n} \xi_{ni}}$, N.T.S. $\varphi(t) \xrightarrow{n\to\infty} e^{-\frac{t^2}{2}}$

$$\varphi(t) = \prod_{i=1}^{n} E e^{it\xi_{ni}} = \prod_{i=1}^{n} E\Big(1 + it\xi_{ni} - \frac{t^2}{2}\xi_{ni}^2 + O\big(t^2 \xi_{ni}^2 I(|\xi_{ni}| > \varepsilon)\big) + O\big(t^3 |\xi_{ni}|^3 I(|\xi_{ni}| \le \varepsilon)\big)\Big).$$

$$= \prod_{i=1}^{n} \Big(\underbrace{1 - \frac{t^2}{2} E\xi_{ni}^2 + O\big(t^2 E\xi_{ni}^2 I(|\xi_{ni}| > \varepsilon)\big) + O\big(t^3 \varepsilon E\xi_{ni}^2\big)}\Big)$$

$1+x \approx e^x$

$$\xrightarrow{n\to\infty} e^{-\frac{t^2}{2} + \cancel{O\left[t^2 \sum_{i=1}^{n} E\xi_{ni}^2 I(|\xi_{ni}| > \varepsilon)\right]} + O[t^3 \varepsilon]}$$

$$\xrightarrow{\varepsilon \to \infty} e^{-\frac{t^2}{2}} \quad \cancel{(\text{let } n \to \infty.\ \text{then } \varepsilon \downarrow 0)}$$

]

Berry-Essen thm: ~~Under the~~ $\xi_{n1}, \dots, \xi_{nn}$ indep

$E\xi_{ni} = 0$. $\sum_{i=1}^{n} E\xi_{ni}^2 = 1$

Then $d_K\big(\mathcal{L}(\sum_{i=1}^{n} \xi_{ni}), N(0,1)\big) \le \sum_{i=1}^{n} E|\xi_{ni}|^3$

- ~~Red~~ i.i.d. case reduces to the ~~prev~~ result mentioned last week.

Example

4. ~~Theorem~~ 3.4.7. $X_1, X_2, \text{———}$ indep. $A>0$.

$Y_m = I_{(|X_m| \le A)}$ If $\sum_{n=1}^{\infty} X_n$ converges a.s,

then (i) $\sum_{n=1}^{\infty} P(|X_n|>A) < \infty$.

(ii). $\sum_{n=1}^{\infty} EY_n$ converges

(iii) $\sum_{n=1}^{\infty} Var(Y_n) < \infty$.

[proof: (i) ✓ o.w. by BC, $P(|X_n|>A \text{ i.o.}) = 1$

$\boxed{\text{If (iii) is not true}}$ ~~To prove (iii)~~, note that $\dfrac{\sum_{m=1}^{n}(Y_m - EY_m)}{\sqrt{\sum_{m=1}^{n} Var(Y_m)}} \xrightarrow{d} N(0,1)$

check Lindeberg's condition?

(i) $\Rightarrow$ $P(X_m \neq Y_m \text{ i.o.}) = 0$

$\sum_{n=1}^{\infty} X_n$ converges

$\Bigg\} \Rightarrow \sum_{n=1}^{\infty} Y_n$ converges a.s

$$\Rightarrow \frac{\sum_{m=1}^{n} Y_m}{\sqrt{\sum_{m=1}^{n} Var(Y_m)}} \to 0 \text{ a.s}$$

By Slutsky's theorem, $\dfrac{-\sum_{m=1}^{n} EY_m}{\sqrt{\sum_{m=1}^{n} Var(Y_m)}} \xrightarrow{d} N(0,1)$ contradiction!

(iii) $\overset{\text{Thm 2.5.3}}{\Rightarrow}$ $\sum_{m=1}^{\infty}(Y_m - EY_m)$ converges a.s

(i) $\Rightarrow$ $\sum_{m=1}^{\infty} Y_n$ converges a.s.

$\Bigg\} \Rightarrow \sum_{m=1}^{\infty} EY_m$ converges

]

5. Self-normalized CLT.

$X_1, X_2, \ldots$ i.i.d. $EX_i = 0$. $\overset{\sigma^2}{\overset{\shortparallel}{EX_i^2}} < \infty$

$V_n^2 = \sum_{i=1}^{n} X_i^2$, then $\frac{X_1+\cdots+X_n}{V_n} \xrightarrow{d} N(0,1)$

[proof: $\frac{X_1+\cdots+X_n}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$ by CLT.

$\frac{V_n}{\sigma\sqrt{n}} \xrightarrow{p} \cancel{\sigma^2}\ 1$ by LLN

$\frac{X_1+\cdots+X_n}{V_n} \xrightarrow{d} N(0,1)$ by Slutsky]

Thm 3.4.6 $X_1, X_2, \ldots$ i.i.d. $S_n = X_1+\cdots+X_n$.

$\exists\, a_n, \& \ b_n$ with $b_n > 0$ s.t. $\frac{S_n - a_n}{b_n} \xrightarrow{d} N(0,1)$

$\iff$ $\frac{y^2 P(|X_1| > y)}{E(|X_1|^2 I(|X_1| \le y))} \to 0$. as $y \to \infty$

Ex

Normal approximation to the binomial. $S_n \sim Bin(n, p)$, $ES_n = np$ $Var(S_n) = npq$.

$q = 1-p$

for $0 \le k \le n$. $P(S_n = k) = P(k - \frac{1}{2} \le S_n \le k + \frac{1}{2})$

$= P(S_n \le k+\frac{1}{2}) - P(S_n \le k - \frac{1}{2})$

6. $Z \sim N(0,1)$.

$$\approx P\left(np + \sqrt{npq}\cdot Z \le k+\tfrac{1}{2}\right) - P\left(np+\sqrt{npq}\cdot Z \le k-\tfrac{1}{2}\right)$$

by CLT

$$= P\left(k-\tfrac{1}{2} < np+\sqrt{npq}\cdot Z \le k+\tfrac{1}{2}\right).$$

$$=: P\left(Z^{(d)}_{np,npq} = k \right)$$

Section 3.5. applied to this example, gives

Local limit theorem:

$$\sup_{k\in\mathbb{Z}} \left| P(S_n=k) - P(Z^{(d)}_{np,npq}=k) \right| \le C_p \cdot \frac{1}{n}. \quad \text{(What's new?)}$$

Cramér's moderate deviation (1938)

$X_1, X_2 \ldots$ i.i.d. $EX_1=0$. $EX_1^2=\sigma^2$ $Ee^{t_0\sqrt{|X_1|}}<\infty$. for some $t_0>0$

then
$$\frac{P\left(\frac{X_1+\cdots+X_n}{\sigma\sqrt{n}} \ge x\right)}{1-\Phi(x)} \to 1 \quad \text{uniformly in } x\in(0, o(n^{\frac{1}{6}}))$$

Compare to B–E?

7. <u>Method of moments</u>:

Recall $\varphi_{X_n}(t) \to \varphi_X(t)$ $\Rightarrow$ $X_n \xrightarrow{d} X$

If all the moments of $\{X_n\}$ and X exist and,

$$EX_n^k \to EX^k \quad \forall k \in \mathbb{Z}^+ \quad \Rightarrow \quad X_n \xrightarrow{d} X.$$

useful in the random matrix theory, random graphs, ---.

<u>Edgeworth expansion</u> X_i's i.i.d. with finite kth moment

$$P\left(\frac{S_n}{\sqrt{n}\,\sigma} \le x\right) \approx \Phi(x) + \frac{Q_3(x)}{\sqrt{n}} + \cdots + \frac{Q_k(x)}{n^{(k-2)/2}}$$

where $Q_3, \ldots, Q_k$ are functions depending on the 3rd - kth moments of X_1

8. Poisson limit theorem (used to approximate sums of "rare" events)

$\varphi_Z(t) = e^{\lambda(e^{it}-1)}$

$Z \sim Poi(\lambda)$ pmf $P(Z=k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$

Thm 3.6.1 For each n, $X_{n1}, \ldots, X_{nn}$ are indep.

$P(X_{nm}=1) = p_{ni}$ $\quad P(X_{ni}=0) = 1-p_{ni}$

If (i). $\sum_{i=1}^{n} p_{ni} \to \lambda$

(ii). $\max_{1\le i\le n} p_{ni} \to 0$.

Then. $\sum_{i=1}^{n} X_{ni} \xrightarrow{d} Poi(\lambda)$

[proof: $Ee^{it\sum_{i=1}^{n} X_{ni}} = \prod_{i=1}^{n} \left(e^{it \cdot 1} \cdot p_{ni} + (1-p_{ni})\right)$

$= \prod_{i=1}^{n} \left(1 + p_{ni}(e^{it}-1)\right)$ $\qquad 1+x \approx e^x$

$\xrightarrow[?]{n\to\infty} e^{\lambda(e^{it}-1)}$

Lemma 3.4.4. b is a complex number with $|b| \le 1$. then. $\left|e^b - (1+b)\right| \le |b|^2$

Lemma 3.4.3. let $z_1, \ldots, z_n$ and $w_1, \ldots, w_n$ be complex numbers $|z_i|, |w_i| \le \theta$.

Then. $\left|\prod_{m=1}^{n} z_m - \prod_{m=1}^{n} w_m\right| \le \theta^{n-1} \cdot \sum_{m=1}^{n} |z_m - w_m|$.

]

· 9.

Rate of convergence.

<u>Thm</u> ~~Under the same se~~ $X_1, \dots, X_n$ indep.

$$P(X_i=1)=p_i \qquad P(X_i=0)=1-p_i.$$

$$\lambda=\sum_{i=1}^{n} p_i$$

then $$d_{TV}\left(\mathcal{L}(X_1+\cdots+X_n),\ \mathrm{Poi}(\lambda)\right)$$

$$\le \min\left(1, \frac{1}{\lambda}\right) \cdot \sum_{i=1}^{n} p_i^2 .$$

~~Results are available for m-dependent sums. Arratia, Goldstein, Gordon (1990)~~

<u>Roughly</u>. $$\mathcal{L}(X_1+\cdots+X_n) \overset{d}{\approx} \mathrm{Poi}(\lambda)$$

If $X_1, \dots, X_n$ are <u>weakly</u> dependent.

<u>Example 3.6.5</u> Occupancy problem.

r balls thrown randomly into

n boxes

$$\#\{\text{empty boxes}\} = X_1+\cdots+X_n.$$

where $X_i = I(i\text{th box is empty})$

10. $$P(X_i = 1) = \underbrace{\left(1-\frac{1}{n}\right)^r}_{p_i}.$$

~~Suppose $\sum_{i=1}^{n} p_i = n(1-\frac{1}{n})^r \to \lambda$.~~

~~Thm 3.6.5~~ Let $\lambda \approx n(1-\frac{1}{n})^r$ for large n,

We expect $\mathcal{L}(X_1+\cdots+X_n) \overset{d}{\approx} \text{Poi}(\lambda)$.

Poisson process

Def: $\{N(t), t \geq 0\}$ is called a Poisson process with rate λ if

(i). $N(0) = 0$

(ii) independent increments

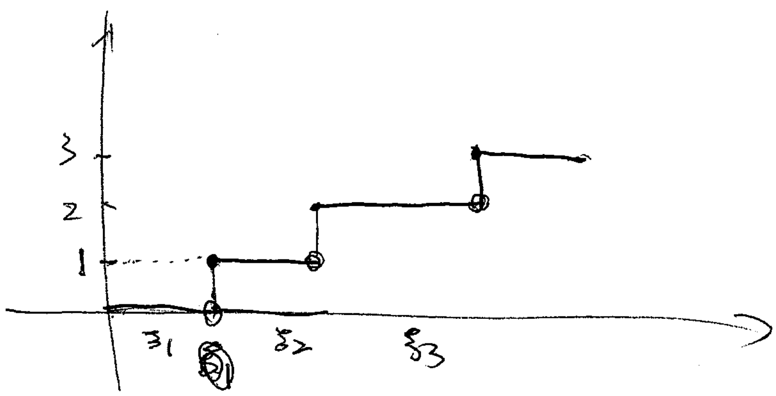
$\forall\, 0 \leq t_0 < t_1 < \cdots < t_k$

$N(t_0), N(t_1)-N(t_0), \cdots, N(t_k)-N(t_{k-1})$ are independent

(iii) $N(t)-N(s) \sim \text{Poisson}(\lambda(t-s))$

To construct a Poisson process. generate $\xi_1, \xi_2, \cdots$ i.i.d $\sim \exp(\lambda)$

$S_n = \xi_1 + \cdots + \xi_n$



$P(\xi_i > t+s \mid \xi_i > t)$

$= P(\xi_i > s)$

11. § 3.9 Limit theorems in $\mathbb{R}^d$

$X = (X_1, \longrightarrow X_d)$ random vector

df. $F_X(x) = P(X_1 \le x_1, - , X_d \le x_d)$.

Def: F_n converges weakly to F. $F_n \Rightarrow F$ if

$F_n(x) \to F(x)$ at all continuity points of F

Thm: $\underset{X_n}{\underset{\wr}{F_n}} \Rightarrow \underset{X}{\underset{\wr}{F}} \iff Ef(X_n) \to Ef(X)$, $\forall f \in C_b(\mathbb{R}^d)$

Def. Ch.f. $\varphi_X(t) = Ee^{it \cdot X}$ $\quad t \cdot X = t_1X_1 + \cdots + t_dX_d$.

Thm: $X_n \Rightarrow X_\infty \iff \varphi_{X_n}(t) \to \varphi_{X_\infty}(t) \quad \forall t \in \mathbb{R}^d$.

Thm. Cramér–Wold device:

$$X_n \Rightarrow X_\infty \iff \theta \cdot X_n \Rightarrow \theta \cdot X_\infty, \ \forall \theta \in \mathbb{R}^d.$$

Thm 3.9.6 CLT in $\mathbb{R}^d$.

$X_1, X_2 \text{ --- }$ i.i.d. $EX_n = \mu$. $Var(X_n) = \Sigma$.

then $\dfrac{X_1 + \cdots + X_n - n\mu}{\sqrt{n}} \xrightarrow{d} N(0, \Sigma)$

Lecture 9 上

1. ch4: Random Walks

def;

0-1 law;

stopping time; Wald's identity;

Recurrence vs transience; arcsine laws; BM

Def. $X_1, X_2, \ldots$ i.i.d. r.v.'s in $\mathbb{R}^d$,

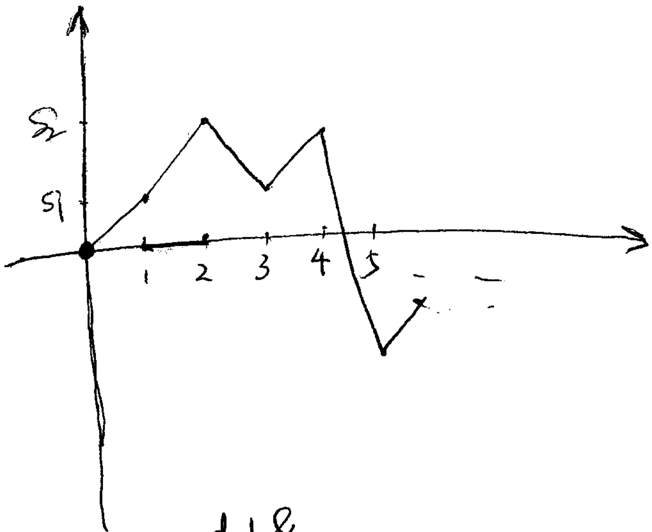
$\boxed{S_n} = X_1 + \cdots + X_n$

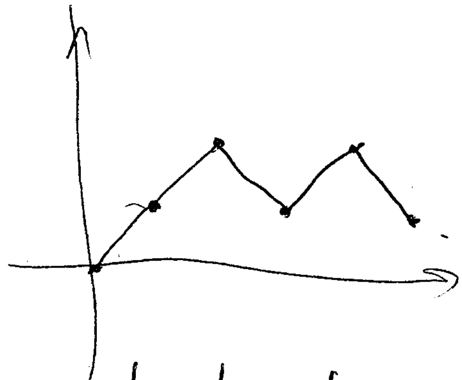
$\boxed{\{S_n : n \geq 1\}}$ is a random walk $\quad S_0 = 0$

Def. Simple Random Walk (SRW) in $\mathbb{R}^1$

if $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$

d=1:





- example: play a fair game

d=1 &

- focus on the properties of the sequence $\{S_n : n \geq 1\}$.

Def. An event is <u>permutable</u> (or <u>exchangeable</u>) if it does not change under finite permutation of $\{X_1, X_2, \ldots\}$

- Tail events are permutable.
- $\{\limsup_{n\to\infty} S_n > c\}$ is permutable, but not a tail event.

Thm 4.1.1. A is permutable $\Longrightarrow P(A) = 0$ or 1

for random walk on $\mathbb{R}$,

2. Thm 4.1.2. One of the following has probability 1.

(i) $S_n = 0$ for all n

(ii) $S_n \to \infty$

(iii) $S_n \to -\infty$.

(iv) $-\infty = \liminf S_n < \limsup S_n = \infty$

[proof: $\limsup_{n\to\infty} S_n$ equals some value in $\mathbb{R} \cup \{-\infty, \infty\}$ with prob. 1 by the 0-1 law.]

- SRW ?

Let $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$

- $\mathcal{F}_n \uparrow$ (filtration)
- $A \in \mathcal{F}_n \Leftrightarrow A = \{(X_1, \dots, X_n) \in B\}$ for some $B \in \mathcal{B}^d$

Whether A occurs or not is determined by the values of $X_1, \dots, X_n$.

<u>Def.</u> $\mathcal{F}_n \uparrow$ σ-fields, τ is a stopping time (optional random variable, optional time, Markov time) $\in \{1, 2, \dots\} \cup \{\infty\}$

with respect to $\{\mathcal{F}_n\}$ if $\{\tau = n\} \in \mathcal{F}_n$, $\forall n = 1, 2, \dots$

- $\{\tau = n\} \in \mathcal{F}_n$, $\forall n = 1, 2, \dots$

$\Leftrightarrow \{\tau \le n\} \in \mathcal{F}_n$ $\forall n = 1, 2, \dots$

$\Leftrightarrow \{\tau \ge n+1\} \in \mathcal{F}_n$ $\forall n = 1, 2, \dots$

- τ_1, τ_2 are stopping times w.r.t. $\{\mathcal{F}_n\}$.

then $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$, $\tau_1 + \tau_2$ are all stopping times w.r.t. $\{\mathcal{F}_n\}$

3. • ~~a constant is a constant n is a stopping time~~ $\cancel{A \subset \mathbb{R}}$ $\cancel{A \subset \mathbb{R}^d}$ $A \in \mathcal{B}^d$.

• Example: $\tau = \inf\{n : S_n \in A\}$. is a stopping time

$[\ \{\tau = n\} = \{S_1 \in A^c, S_2 \in A^c, \cdots, S_{n-1} \in A^c, S_n \in A\}.\ ]$

is a function of $\{X_1, \cdots, X_n\}$

$\in \sigma(X_1, \cdots, X_n)$

• n is a stopping time.

• ~~Example~~ • $\mathcal{F}_\tau$? $\quad \mathcal{F}_\tau \perp\!\!\!\perp \{X_{\tau+1}, \cdots\}$.

Thm 4.1.5 Wald's equation

$X_1, X_2, \cdots$ i.i.d. τ is a stopping time. (with respect to $\{\sigma(X_1, \cdots, X_n)\}$)

If $E|X_1| < \infty$, $E\tau < \infty$.

then. $E(S_\tau) = E(X_1) \cdot E(\tau)$

$\cdot\ E(S_n) = n \cdot E(X_1)$

[proof: First suppose $X_i \geq 0$

$$E(S_\tau) = E\left(\sum_{i=1}^{\tau} X_i\right)$$

$$= E\left(\sum_{i=1}^{\infty} X_i I_{(i \leq \tau)}\right)$$

$$= \sum_{i=1}^{\infty} E\Big(X_i \underbrace{I_{(i \leq \tau)}}_{\in \mathcal{F}_{i-1}}\Big)$$

$$= \sum_{i=1}^{\infty} EX_i \cdot P(\tau \geq i)$$

$$= (EX_1) \cdot \sum_{i=1}^{\infty} P(\tau \geq i) = (EX_1) \cdot E\tau)$$

Write $X_i = X_i^+ - X_i^-$ (some details)].

4. Thm 4.1.6 Wald's second equation

$\cancel{X_1, \dots, X} X_1, X_2 \dots$ i.i.d. $EX_1 = 0$. $EX_1^2 = \sigma^2 < \infty$.

τ is a stopping time & $E\tau < \infty$

then $E\boxed{S_\tau^2} = \sigma^2 \cdot E\tau$.

$E|X_n S_{n-1} I_{(\tau \ge n)}| \le \sqrt{EX_n^2 \cdot E(S_{n-1}^2)} < \infty$

[proof: $\because$ $S_{\tau\wedge n}^2 = \cancel{S^2_{\tau\wedge(n-1)}} + S^2_{\tau\wedge(n-1)} + (2X_nS_{n-1} + X_n^2) I_{(\tau \ge n)}$

$\therefore$ $E(S_{\tau\wedge n}^2) = E(S^2_{\tau\wedge(n-1)}) + 0 + \sigma^2 \cdot P(\tau \ge n)$

$\therefore$ $E(S_{\tau\wedge n}^2) = \underbrace{\sigma^2 \sum_{i=1}^{n} P(\tau \ge i)}$ $< \infty$ [thus means $S_{\tau\wedge n} \in L^2(\Omega)$]

$\uparrow \sigma^2 E\tau$ as $n \to \infty$

$L^2(\Omega) := \{X : E(X^2) < \infty\}$. metric

closed

complete metric space with $\|\cdot\|_2$

as for (*)
similarly one can prove $E(S_{\tau\wedge n} - S_{\tau\wedge m})^2 = \sigma^2 \sum_{k=m+1}^{n} P(\tau \ge k)$ for $n \ge m$

$\therefore$ $\{S_{\tau\wedge n}\}$ is a Cauchy sequence in L^2]

$E(S_{\tau\wedge n} - S_\tau)^2 \to 0$

$ES_\tau^2 = E(S_{\tau\wedge n} + (S_\tau - S_{\tau\wedge n}))^2$

$= ES_{\tau\wedge n}^2 + \underbrace{2\, ES_{\tau\wedge n} \cdot (S_\tau - S_{\tau\wedge n})} + \underline{E(S_\tau - S_{\tau\wedge n})^2}$ $\downarrow 0$

$\overset{\text{H}}{\le} 2\sqrt{ES_{\tau\wedge n}^2 \cdot E(S_\tau - S_{\tau\wedge n})^2}$

$\downarrow 0$

5.

<u>Exercise 4.1.12</u> $X_1, X_2, \cdots$ Unif $(0,1)$

$S_n = X_1 + \cdots + X_n$

$\tau = \inf\{n : S_n > 1\}$.

then (1) $P(\tau > n) = \frac{1}{n!}$

$$\Big[\ P(\tau>n) = P(S_n \le 1) = \int \cdots \int I(x_1 + \cdots + x_n \le 1)\, dx_1 \cdots dx_n$$

$$= \int_{0 \le y_1 \le y_2 \le \cdots \le y_n \le 1} \cdots \int dy_1 \cdots dy_n$$

$$= \int_0^1 \int_0^{y_n} \cdots \int_0^{y_2} dy_1 \cdots dy_{n-1}\, dy_n$$

$$= \int_0^1 \int_0^{y_n} \cdots \int_0^{y_3} y_2\, dy_2 \cdots dy_{n-1}\, dy_n$$

$$= \int_0^1 \int_0^{y_n} \cdots \frac{y_3^2}{2} \cdots dy_{n-1}\, dy_n$$

$$= \int_0^1 \frac{y_n^{n-1}}{(n-1)!}\, dy_n = \frac{1}{n!}\ \Big]$$

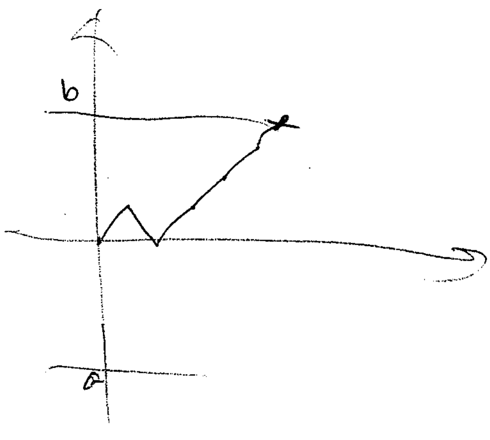
(2) $E(\tau) = \sum_{n=1}^{\infty} P(\tau > n) = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} = e$

(3) By Wald's equation : $E(S_\tau) = E(X_1) \cdot e = \frac{e}{2}$.

<u>Example 4.1.5</u>: $X_1, X_2, \cdots$ i.i.d. $P(X_1 = 1) = P(X_1 = -1) = \frac{1}{2}$

$S_n = X_1 + \cdots + X_n$ (SRW)

a, b integers, $a < 0 < b$



$N = \inf\{n : S_n \notin (a,b)\}$

then ① $E(N) < \infty$

② $S_N = a$ or $S_N = b$.

6. ③. $P(S_N=a)=\frac{b}{b-a}$ $P(S_N=b)=\frac{-a}{b-a}$

④. $E(N)=E(S_N^2)=(-a)\cdot b$

[proof: ① to be proved later

③: $E(S_N) = \underline{\underline{E(X_1)}}\cdot E(N)=0$

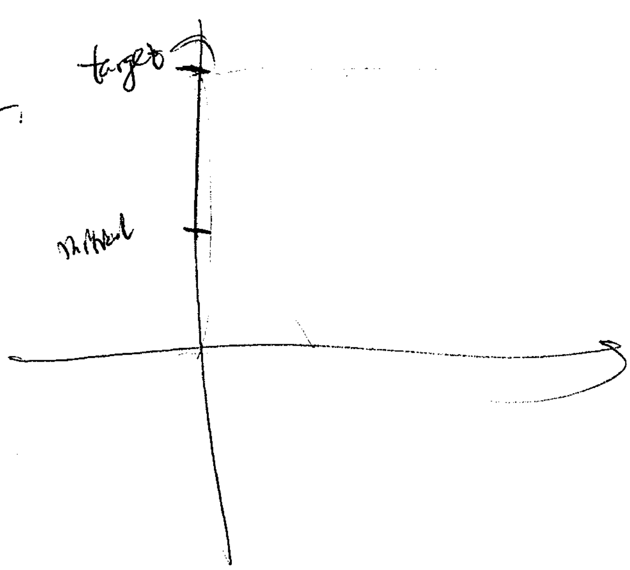
$E(S_N) = P(S_N=a)\cdot a+P(S_N=b)\cdot b$, $E(X_1)=0$

④. $E(S_N^2)=E(X_1^2)\cdot E(N)=E(N)$

$E(S_N^2) \overset{③}{=} \frac{a^2b}{b-a}+\frac{b^2(-a)}{b-a}=(-a)\cdot b$

①: $P(N>k(b-a))\leq(1-\frac{1}{2^{b-a}})^k$.]

Application in gambling:



7.

Thm 4.1.7. $X_1, X_2, \dots$ i.i.d. $EX_1 = 0$, $EX_1^2 = 1$.

Let $T_c = \inf\{n \geq 1 : |S_n| > c n^{\frac{1}{2}}\}$.

then $ET_c \begin{cases} < \infty & \text{for } c < 1 \\ = \infty & \text{for } c \geq 1. \end{cases}$

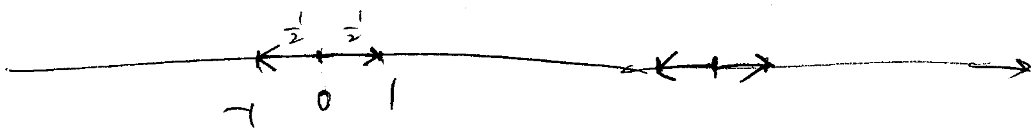
Recall LIL:

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{n}\sqrt{2\log\log n}} = 1.$$

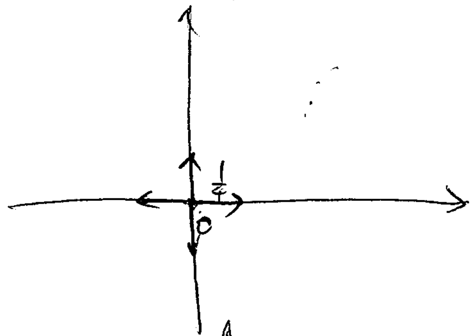
[proof omitted]

Section 4.2 Recurrence vs. Transience.

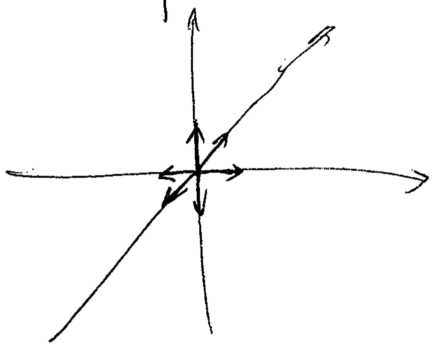
SRW in $\mathbb{R}^1$



in $\mathbb{R}^2$



in $\mathbb{R}^3$



$\frac{1}{2d}$.

define $\tau_1 = \inf\{m \geq 1 ; S_m = 0\}$

$\tau_n = \inf\{m > \tau_{n-1} ; S_m = 0\}$

Thm The following are equivalent

(i). $P(\tau_1 < \infty) = 1$

(ii). $P(\tau_n < \infty) = 1 \quad \forall n \in \mathbb{Z}^+$

(iii). $P(S_m = 0 \text{ i.o.}) = 1$.

(iv) $\sum_{m=0}^{\infty} P(S_m = 0) = \infty$

[proof:
$$P(\tau_n<\infty) = P(\tau_1<\infty, \tau_2-\tau_1<\infty, \cdots, \tau_n-\tau_{n-1}<\infty)$$
$$= \left(P(\tau_1<\infty)\right)^n$$

$$\{S_m=0 \text{ i.o.}\} = \{\tau_n<\infty, \ \forall n\in\mathbb{Z}^+\}.$$

$$\sum_{m=0}^{\infty} P(S_m=0) = \sum_{m=0}^{\infty} E I_{(S_m=0)} = E\sum_{m=0}^{\infty} I_{(S_m=0)}$$
$$= E\sum_{n=0}^{\infty} I(\tau_n<\infty) = \sum_{n=0}^{\infty} P(\tau_n<\infty)$$
$$= \sum_{n=0}^{\infty}\left(P(\tau_1<\infty)\right)^n = \frac{1}{1-P(\tau_1<\infty)} \quad \blacksquare \;]$$

Def. If $P(\tau_1<\infty)=1$ then the RW is ~~said~~ called recurrent
If $P(\tau_1<\infty)<1$, then — — — transient.

Thm 4.2.3 SRW is recurrent in $\mathbb{R}^1$ & $\mathbb{R}^2$
is transient in $\mathbb{R}^d$. $d\geq 3$

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

Thm. $\tau_1 =: \tau_{0\to 0}$ define $\tau_{a\to b}$ similarly for $a, b\in\mathbb{Z}^d$.

$$P(\tau_1<\infty)=1 \iff P(\tau_{a\to b}<\infty)=1 \quad \forall a, b\in\mathbb{Z}^d$$

- no class on November 15 (homework 5 due on Nov. 22)

9. • Final exam on December 6 3:30–6:30 pm ~~Venue~~ LSB C5 ? Closed book.

~~~~ 1 A4 helpsheet, double sided ~~hand written ?~~ (handwritten or printed)

[proof: $d=1$: N.T.S. $\sum_{n=0}^{\infty} P(S_n=0)=\infty$.

$\boxed{\text{Lecture 10}}$

$$P(S_{2n}=0)=\binom{2n}{n}\cdot\left(\frac{1}{2}\right)^{2n}\sim\frac{\sqrt{2\pi\cdot 2n}\cdot\left(\frac{2n}{e}\right)^{2n}}{\left(\sqrt{2\pi n}\cdot\left(\frac{n}{e}\right)^{n}\right)^{2}}\cdot\frac{1}{2^{2n}}\sim\frac{2\sqrt{\pi}\cdot\sqrt{n}}{2\pi n}$$

n left steps

n right steps

Stirling's formula: $n!\sim\sqrt{2\pi n}\cdot\left(\frac{n}{e}\right)^{n}$.

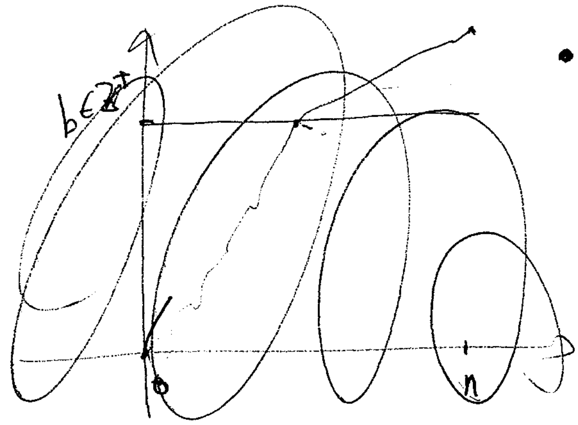
~~d=2~~ $d\geq 2$ proof omitted. ■]

$\boxed{\text{Lecture 9 finished here}}$

Section 4.3. ~~Arcsine Laws~~ Classical results on 1-d SRW.

$P(X_i=\pm 1)=\frac{1}{2}$.

$S_n=X_1+\cdots+X_n$.



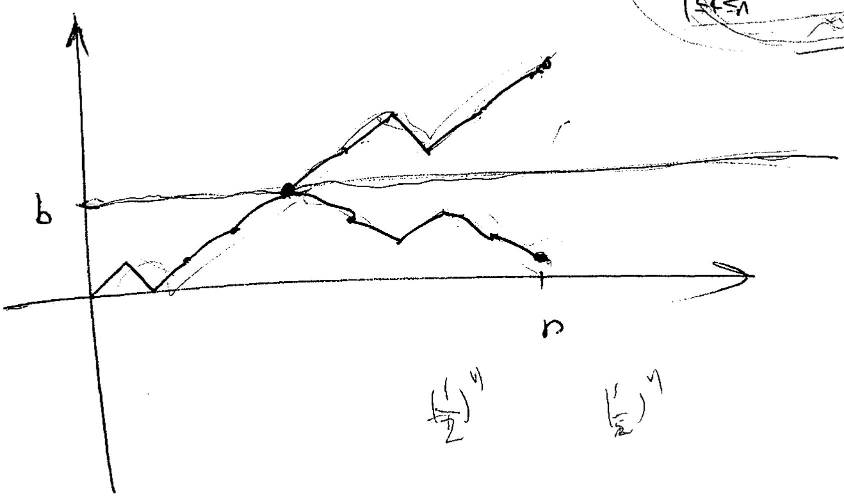
- Reflection Principle:

$\forall b\in\mathbb{Z}^+$

$$P\left(\max_{1\le k\le n} S_k\geq b\right)=2P(S_n>b)+P(S_n=b)$$

$$\parallel$$

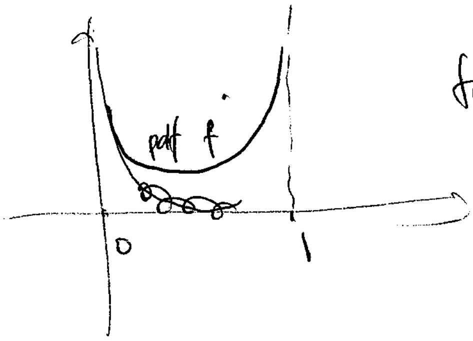
$$P\left(\max_{1\le k\le n} S_k\geq b,\ S_n>b\right)+P\left(\max_{1\le k\le n} S_k\geq b,\ S_n<b\right)+P\left(\max_{1\le k\le n} S_k\geq b,\ S_n=b\right)$$



$\left(\frac{1}{2}\right)^{n}$ $\left(\frac{1}{2}\right)^{n}$
~~~~

10. ~~Thm 4.3. Ballot theorem~~

Arcsine ~~laws~~ distribution.



$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

Define $L_{2n} := \sup\{m \le 2n : S_m = 0\}$.

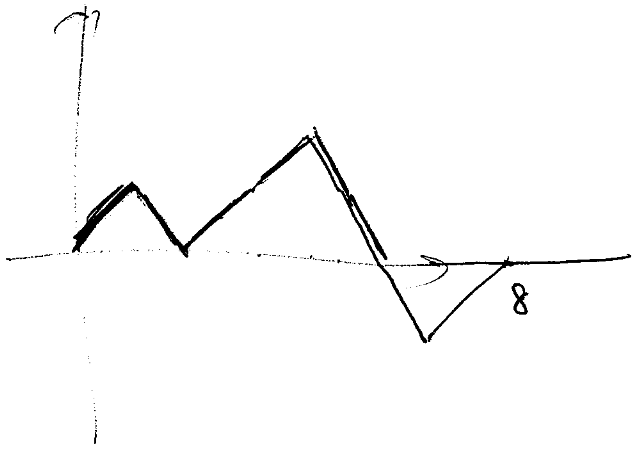
Thm 4.3.5 $P\left(a < \frac{L_{2n}}{2n} < b\right) \xrightarrow{n\to\infty} \int_a^b \frac{1}{\pi\sqrt{x(1-x)}}\,dx \quad \forall\, 0 < a < b < 1$

$\frac{L_{2n}}{2n}$.

$F_n := \inf\{0 \le m \le n : S_m = \max_{0\le k\le n} S_k\}$

$\frac{F_n}{n}$

$\pi_{2n} := \#\{1 \le k \le 2n : (k-1, S_{k-1}) \to (k, S_k)$ is above the x-axis$\}$ (number of)



$\frac{\pi_{2n}}{2n}$.

Thm: all these $\xrightarrow{d}$ arcsine distribution.

11. Ch 8 Brownian Motion. $\{B_t : t \geq 0\}$.

To construct a standard BM: (1) Generate a SRW. $\{S_i : i \in \mathbb{Z}^+\}$.

(2) Extend it to $\{S_i : i \in \mathbb{R}^+\}$ by linear interpolation

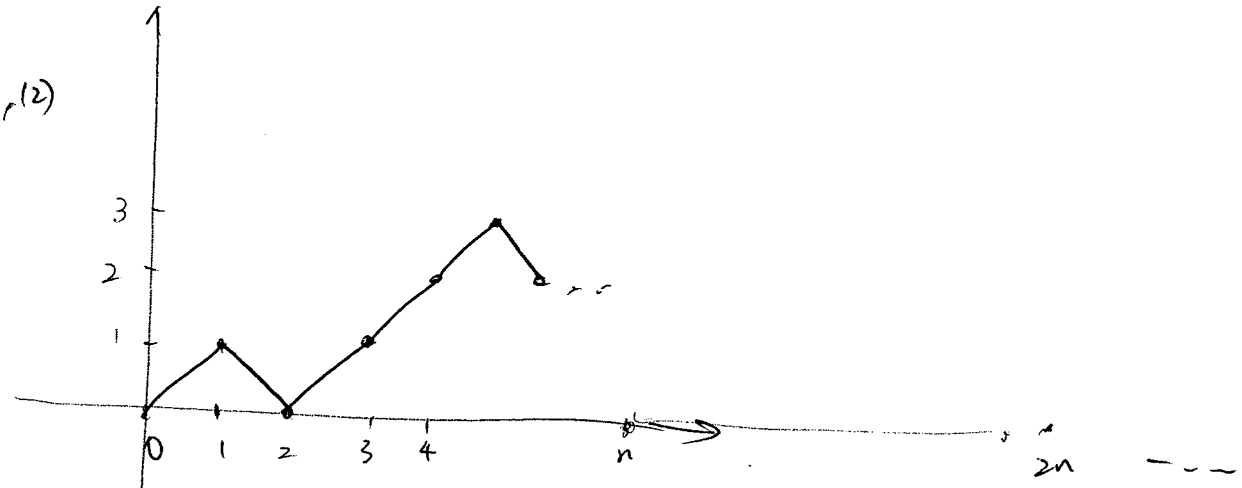
~~define $B_t =$~~

(3) Scale it by n in time, by $\sqrt{n}$ in space

$$B_t^{(n)} := \frac{S_{nt}}{\sqrt{n}} \quad t \in \mathbb{R}^+$$

(4) Let $n \to \infty$.

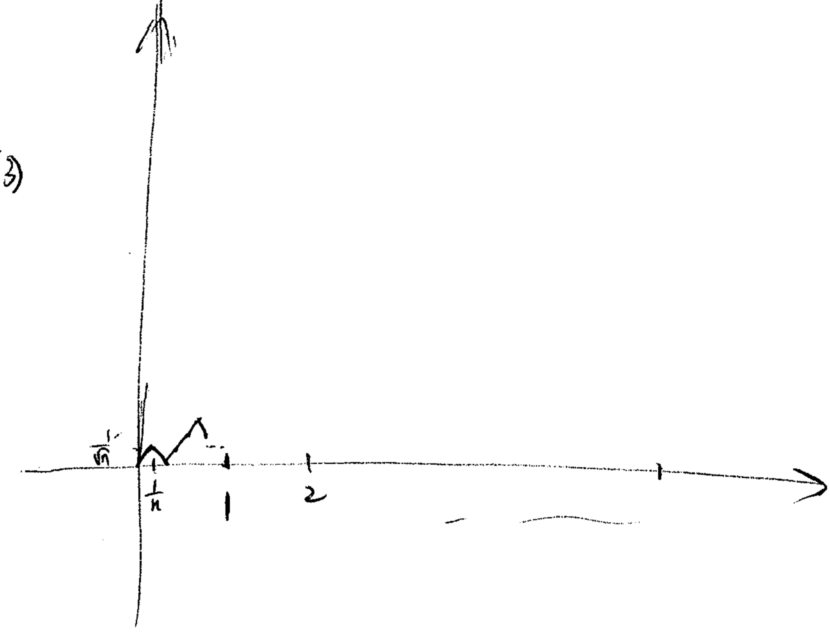
(1), (2)



(3)

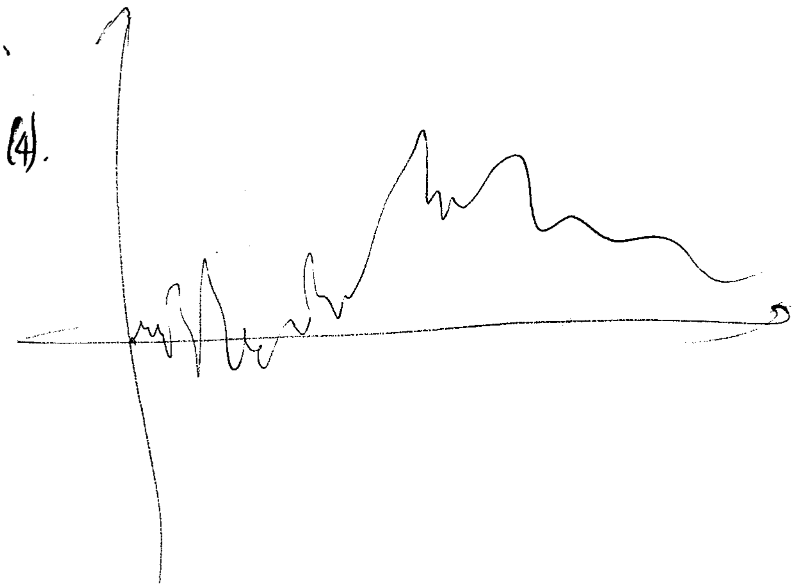
zoom out

$\times n$



12.

(4).



Why scale space by $\sqrt{n}$?

$$B_1^{(n)} = \frac{S_n}{\sqrt{n}} \xrightarrow{d} \underbrace{N(0,1)}_{B_1}, \text{ as } n \to \infty.$$

n

$n^{\frac{1}{4}}$ $\qquad \infty$,

Def. A one-dimensional standard BM is a real-valued process B_t, $t \geq 0$, satisfying

(a) If $t_0 < t_1 < \cdots < t_n$ then $B(t_0), B(t_1) - B(t_0), \cdots, B(t_n) - B(t_{n-1})$ are independent

(b) If $s, t \geq 0$. $B(s+t) - B(s) \sim N(0, t)$

(c) With probability 1, $t \to B_t$ is continuous.

Properties:

- $$P\left(\max_{0 \leq t \leq T} B_t \geq b\right) = 2P(B_t \geq b)$$

- $\sup\{t \leq 1 : B_t = 0\}$, $\inf\{t \leq 1 : B_t = \max_{0 \leq s \leq 1} B_s\}$,

$\int_0^1 1_{\{B_t > 0\}} dt$,

$\sim$ arcsin law.

1.

Ch5: Conditional expectation; Martingales

$E(1_A|B)$
$\parallel$
$P(A|B) = \dfrac{P(AB)}{P(B)}$

$$E(X|Y=y) = \int x \cdot f(x|Y=y)\,dx = \int x \frac{f(x,y)}{f_Y(y)}\,dx.$$

Def: X is a r.v. on $(\Omega, \mathcal{F}, P)$, $E|X| < \infty$.

$\mathcal{A} \subset \mathcal{F}$, $\mathcal{A}$ is a σ-field.

$E(X|\mathcal{A})$, called the conditional expectation of X given $\mathcal{A}$,

is defined as a random variable Y satisfying

(i). $Y \in \mathcal{A}$, i.e., Y is $\mathcal{A}$-measurable.

(ii) $\forall A \in \mathcal{A}$, $E[X 1_A] = E[Y 1_A]$.

- existence and uniqueness to be ~~discussed~~ proved later

- Let $A = \{Y > 0\}$ in (ii):

$$E(Y^+) = E[X 1_A] \le E(|X| 1_A)$$

similarly $E(Y^-) \le E(|X| 1_{A^c})$

$\therefore E|Y| \le E|X| < \infty.$

- $E[Y] = E[X]$

- picture in my mind.

2.

- Uniqueness: Suppose Y_1 and Y_2 both satisfy (i) & (ii)

Let $A = \{Y_1 - Y_2 \geq \varepsilon > 0\}$. $A \in \mathcal{A}$.

$$0 = E(XI_A) - E(XI_A) \overset{(ii)}{=} E(Y_1 I_A) - E(Y_2 I_A) = E[(Y_1 - Y_2) I_A] \geq \varepsilon P(A)$$

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

......

- Existence:

~~Radon~~ Radon–Nikodym theorem: (proof not required) Let μ and ν be σ-finite measures on $(\Omega, \mathcal{A})$

if $\nu \ll \mu$ (i.e., $\mu(A) = 0$ implies $\nu(A) = 0$),

then there is a function $f \in \mathcal{A}$ s.t. $\forall A \in \mathcal{A}$.

$$\int_A f \, d\mu = \nu(A)$$

such f, denoted by $d\nu/d\mu$, is called the Radon–Nikodym derivative.

First suppose $X \geq 0$. Let $\mu = P$. $\nu(A) = E(XI_A)$ for $A \in \mathcal{A}$

[o.w. write $X = X^+ - X^-$] then $\nu \ll \mu$. both σ-finite, ν is a measure

Let $Y = \frac{d\nu}{d\mu}$

$$\int_A f \, d\mu = E(YI_A)$$

$$\nu(A) = E(XI_A)$$

Theorem 5.1.2 $E|X|, E|Y| < \infty$.

(a) $E(aX+Y|\mathcal{A}) = aE(X|\mathcal{A}) + E(Y|\mathcal{A})$ (linearity)

(b) If $X \le Y$, then $E(X|\mathcal{A}) \le E(Y|\mathcal{A})$ (monotonicity)

(c) If $X_n \ge 0$, $X_n \uparrow X$, $EX < \infty$, then

$$E(X_n|\mathcal{A}) \uparrow E(X|\mathcal{A})$$

(monotone convergence theorem)

[proof: (a) $\forall A \in \mathcal{A}$:

$$E\left[aE(X|\mathcal{A}) \cdot 1_A + E(Y|\mathcal{A})1_A\right]$$

$$= aE\left(E(X|\mathcal{A})1_A\right) + E\left(E(Y|\mathcal{A}) \cdot 1_A\right)$$

$$= a \cdot E(X1_A) + E(Y1_A)$$

$$= E((aX+Y)1_A).$$

(b) Let $A = \{E(X|\mathcal{A}) - E(Y|\mathcal{A}) \ge \varepsilon > 0\} \in \mathcal{A}$.

$$E\left(E(X|\mathcal{A})1_A\right) = E(X1_A) \le E(Y1_A) = E\left(E(Y|\mathcal{A}) 1_A\right)$$

$$\Rightarrow E\Big(\underbrace{(E(X|\mathcal{A}) - E(Y|\mathcal{A}))1_A}_{\ge\, \varepsilon \cdot P(A)}\Big) \le 0$$

$$\Rightarrow P(A) = 0$$

(c) $E(X|\mathcal{A}) - E(X_n|\mathcal{A}) \xrightarrow{(a)} E(X - X_n|\mathcal{A}) \overset{(b)}{\downarrow} Z_\infty$.

$Z_\infty \ge 0$, $Z_\infty \in \mathcal{A}$, $\forall A \in \mathcal{A}$: $E(Z_\infty 1_A) \overset{DCT}{=} \lim_{n\to\infty} E\left(E(X-X_n|\mathcal{A}) 1_A\right)$

$$= \lim_{n\to\infty} E((X-X_n)1_A) \overset{DCT}{=} 0 \quad \therefore Z_\infty = 0.$$

3/

- Jensen's inequality: If φ is convex, $E|X|<\infty$, $E|\varphi(X)|<\infty$.

  then $\varphi(E(X|\mathcal{F}))\le E(\varphi(X)|\mathcal{F})$

- Chebyshev's inequality: If $a>0$ then.

  $P(|X|\ge a|\mathcal{F})\le a^{-2}E(X^2|\mathcal{F})$

  - $P(A|\mathcal{F})=E(1_A|\mathcal{F})$

- Hölder's inequality: $p\ge 1$, $\frac{1}{p}+\frac{1}{q}=1$.

  $$|E(XY|\mathcal{A})|\le\left[E(|X|^p|\mathcal{A})\right]^{\frac{1}{p}}\cdot\left[E(|Y|^q|\mathcal{A})\right]^{\frac{1}{q}}.$$

- $p\ge 1$. $\quad |E(X|\mathcal{F})|^p\le E(|X|^p|\mathcal{F})$

~~B.~~ 4. ~~Properties~~

- If $X \in \mathcal{A}$, then $E(X|\mathcal{A}) = X$

- If X and $\mathcal{A}$ are independent, then $E(X|\mathcal{A}) = E(X)$

  [$E(X\mathbb{1}_A) = E(X) \cdot P(A) = E(E(X)\mathbb{1}_A)$, $\forall A \in \mathcal{A}$].

- Thm 5.1.7 If $X \in \mathcal{A}$. $E|XY| < \infty$, $E|Y| < \infty$.

  then $E(XY|\mathcal{A}) = X \cdot E(Y|\mathcal{A})$

  N.T.S. $E(XY\mathbb{1}_A) = E(X \cdot E(Y|\mathcal{A})\mathbb{1}_A)$ $\forall A \in \mathcal{A}$

[proof: 1. $X = \mathbb{1}_B$, $B \in \mathcal{A}$

$$\text{~~RHS~~} = E[\mathbb{1}_B \mathbb{1}_A E(Y|\mathcal{A})]$$

$$= E[E(Y|\mathcal{A})\mathbb{1}_{A\cap B}]$$

$$= E(Y\mathbb{1}_{A\cap B})$$

$$= E(Y\mathbb{1}_A\mathbb{1}_B) = E(XY\mathbb{1}_A)$$

2. ~~simple r.v.~~ $X = \sum_{r=1}^{n} b_r \mathbb{1}_{B_r}$

$$E\left(\sum_{r=1}^{n} b_r \mathbb{1}_{B_r} Y \Big| \mathcal{A}\right) \overset{\text{linearity}}{=} \sum_{r=1}^{n} b_r E(\mathbb{1}_{B_r} Y|\mathcal{A}) \overset{\text{from 1.}}{=} \sum_{r=1}^{n} b_r \mathbb{1}_{B_r} \cdot E(Y|\mathcal{A})$$

$$= X E(Y|\mathcal{A})$$

3. $X \geq 0$, $Y \geq 0$

4. $X = X^+ - X^-$

   $Y = Y^+ - Y^-$.

5.

Thm 5.1.6. If $\mathcal{A}_1 \subset \mathcal{A}_2$, then

(i). $E(E(X|\mathcal{A}_1)|\mathcal{A}_2) = E(X|\mathcal{A}_1)$

(ii) $E(E(X|\mathcal{A}_2)|\mathcal{A}_1) = E(X|\mathcal{A}_1)$

the smaller σ-field always wins.

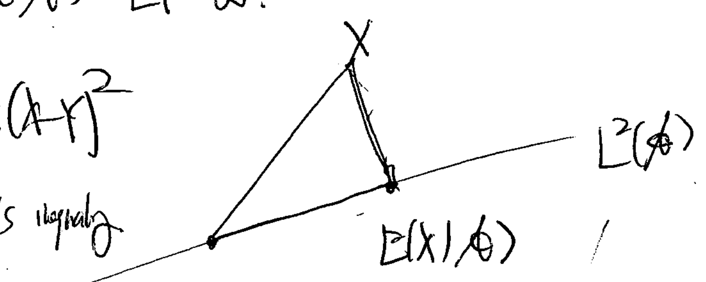
- $E(X) = E(E(X|\mathcal{A}))$, $\forall \mathcal{A}$.
  - [useful when $E(X|\mathcal{A})$ is easier to compute]
- If $E(X|\mathcal{A}_2) \in \mathcal{A}_1$, then $E(X|\mathcal{A}_1) = E(X|\mathcal{A}_2)$

[proof: (i) ✓

(ii). $\forall A \in \mathcal{A}_1$ $E(E(X|\mathcal{A}_1)1_A) = E(X1_A) = E(E(X|\mathcal{A}_2)1_A)$]

Thm 5.1.8: $EX^2<\infty$. $\forall Y \in \mathcal{A}$, $EY^2<\infty$.

$$E(X-\underbrace{E(X|\mathcal{A})}_{\in L^2(\mathcal{A}) \text{ by Jensen's inequality}})^2 \le E(X-Y)^2$$



- In time series models, given historical data $X_1,\dots,X_n$, the "best" prediction for "future X_{n+1}" is $E(X_{n+1}|X_1,\dots,X_n)$.

[proof: $E(X-Y)^2 = E(X-E(X|\mathcal{A})+E(X|\mathcal{A})-Y)^2$

$$= E(X-E(X|\mathcal{A}))^2 + E(Y-E(X|\mathcal{A}))^2 - 2E[(X-E(X|\mathcal{A}))\underbrace{(Y-E(X|\mathcal{A}))}_{\in\mathcal{A}}]$$

$\ge E(X-E(X|\mathcal{A}))^2$]

6. **Def**: $E(X|Y) := E(X|\sigma(Y))$

$$\begin{array}{c} E(X|Y) \\ \parallel \\ h(Y) \end{array}$$

Example 5.1.4 If $X, Y \sim f(x,y)$ joint pdf.

marginal density of Y is $f_Y(y) = \int f(x,y)\,dx$ (assume > 0 $\forall y$)

conditional density of X given $Y=y$ is

$$\frac{f(x,y)}{f_Y(y)}$$

If $E|X| < \infty$ conditional expectation $E(X|Y=y) = \int x \cdot \frac{f(x,y)}{f_Y(y)}\,dx =: h(y)$

Claim: ~~In fact~~ In this case, $E(X|Y) = h(Y)$

[proof: $\forall A \in \sigma(Y)$, write $A = \{Y \in B\}$.

$$E[h(Y)1_A] = \int_B h(y) f_Y(y)\,dy$$

$$= \int_B \int x \frac{f(x,y)}{\cancel{f_Y(y)}}\,dx \cdot \cancel{f_Y(y)}\,dy$$

$$= E[X 1_{\{Y \in B\}}] = E[X 1_A] \quad .\ ]$$

Example 5.1.5 If X and Y are independent

$E|g(X,Y)| < \infty$.

then $E(g(X,Y)|X) = h(X)$.

where $h(x) = Eg(x,Y)$

[proof: $\forall A \in \sigma(X)$, write $A = \{X \in B\}$.

$$E[g(X,Y)1_A] = E[g(X,Y)1_{\{X \in B\}}]$$

$$= \iint g(x,y) 1_B(x)\,\mu(dx)\nu(dy)$$

$$= \int h(x) 1_B(x)\,\mu(dx)$$

$$= E[h(X) 1_{\{X \in B\}}]$$

7. Martingales:

Def: $\mathcal{F}_n \uparrow$ σ-fields (filtration)

S_n, $n \geq 1$ is a sequence of r.v.'s

S_n is called a martingale w.r.t. $\mathcal{F}_n$ if

(i) $E|S_n| < \infty$

(ii) $S_n \in \mathcal{F}_n$.

(iii) $E(S_n | \mathcal{F}_{n-1}) = S_{n-1}$, $n \geq 2$ (a.s.)

- $E(S_{n+m} | \mathcal{F}_n) = S_n$, $m \geq 1$, $n \geq 1$

[$\mathcal{F}_n = \sigma(S_1, \dots, S_n)$]

- (iii) $\Rightarrow$ $E(S_n) = E(S_{n-1}) = \dots = E(S_1)$ ($= 0$, then called mean-zero martingale, $\mathcal{F}_0 = \{\phi, \Omega\}$, $S_0 = 0$

~~Example~~ - note different notation.

Def, If S_n is a martingale w.r.t. $\mathcal{F}_n$, - $E|S_1| \leq E|S_2| \leq \dots \leq E|S_n| \leq \dots$ (Jensen)

$X_1 = S_1$, $X_2 = S_2 - S_1$, ..., $X_n = S_n - S_{n-1}$, ...

are called martingale differences. - (iii). $EX_n = 0$, $n \geq 2$

lecture 10 ends here - If $EX_n^2 < \infty$, then. $E(S_n^2) = E(S_{n-1}^2 + 2S_{n-1}X_n + X_n^2) = E(S_{n-1}^2) + EX_n^2 = \dots = \sum_{i=1}^n EX_i^2$]

Example, $X_1, X_2, \dots$ indep. $EX_i = 0$. $S_n = \sum_{i=1}^n X_i$

then $\{S_n, n \geq 1\}$ is a martingale w.r.t. $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$

$$[\ E(S_n|\mathcal{F}_{n-1}) = E(S_{n-1} + X_n | \mathcal{F}_{n-1}) = E(S_{n-1}|\mathcal{F}_{n-1}) + E(X_n|\mathcal{F}_{n-1}) = S_{n-1} + EX_n = S_{n-1}\ ]$$

Example: $E|X| < \infty$, $\mathcal{F}_n \uparrow$ σ-fields $S_n := E(X|\mathcal{F}_n)$

then $\{S_n : n \geq 1\}$ is a martingale w.r.t. $\mathcal{F}_n$.

$$[\ E(S_n|\mathcal{F}_{n-1}) = E(E(X|\mathcal{F}_n)|\mathcal{F}_{n-1}) = E(X|\mathcal{F}_{n-1}) = S_{n-1}\ ]$$

8.

$Y_1, Y_2, \cdots$

<u>Example</u>. $S_n = g_n(Y_1, \to Y_n)$. $\mathcal{F}_n = \sigma(Y_1 \to Y_n)$ $n=1,2,\cdots$

$E|S_n| < \infty$

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

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

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

Then $T_n := \sum_{r=1}^{n} Y_r$ is a martingale w.r.t. $\mathcal{F}_n$.

<u>Def.</u> S_n is called a supermartingale (submartingale resp.) w.r.t. $\mathcal{F}_n$ if

(i) $E|S_n| < \infty$

(ii). $S_n \in \mathcal{F}_n$

(iii) $E(S_n | \mathcal{F}_{n-1}) \le S_{n-1}$ ($\ge$ resp.), $n \ge 2$

$\{-S_n, \mathcal{F}_n\}$ $ES_1 \le ES_2 \le \cdots$

<u>Thm 5.2.3 & 5.2.4.</u>

①. S_n is a martingale w.r.t. $\mathcal{F}_n$. φ is convex s.t. $E|\varphi(S_n)| < \infty$. $\forall n$

then $\varphi(S_n)$ is a ~~sub mar~~ submartingale

②. S_n is a submartingale w.r.t. $\mathcal{F}_n$ φ is convex s.t. $E|\varphi(S_n)| < \infty$ $\forall n$.
and $\varphi \uparrow$

then $\varphi(S_n)$ is a submartingale

↑ concave

• ~~similar results~~ translates into results for supermartingales. $\boxed{\phi(x)} = -\varphi(-x)$

• $\{S_n\}$ is supermartingale, $\{-S_n\}$ sub. $\varphi\uparrow$ convex $\varphi(-S_n)$ sub $-\varphi(-S_n)$ super

9. $\varphi(x)=|x|^p$, $p\geq 1$.

$\varphi(x)= \max(a,x)$ $\uparrow$

$\varphi(x)=-\log x$, $x>0$

$\varphi(x)= e^{cx}$, $c>0$. $\uparrow$

[proof. ① $E(\varphi(S_n)|\mathcal{F}_{n-1}) \overset{\text{Jensen}}{\geq} \varphi(E(S_n|\mathcal{F}_{n-1}))=\varphi(S_{n-1})$

② - - - - - - - - - - - - - - - - - - $\geq \varphi(S_{n-1})$]

• Recall the proof of SLLN

~~Thm 5.2.8 Martingale convergence theorem.~~

Def. $\mathcal{F}_n \uparrow$ σ-fields H_n, $n\geq 2$ is called a predictable sequence if $H_n \in \mathcal{F}_{n-1}$, $n\geq 2$.

not used below

Thm 5.2.5 $\{S_n\}$ is a supermartingale w.r.t. $\{\mathcal{F}_n\}$, $n\geq 1$. $\{S_n, \mathcal{F}_n, n\geq 1\}$

H_n, $n\geq 2$ predictable, $H_n \geq 0$, H_n bounded.

then. $T_n := \sum_{i=2}^{n} \overset{\text{choice}}{\boxed{H_i}} \underbrace{(S_i - S_{i-1})}$ gain in the ith game (original), $n\geq 2$ $T_1 = 0$.

is a supermartingale w.r.t. $\{\mathcal{F}_n\}$.

$E|T_n|<\infty$ ✓

$T_n \in \mathcal{F}_n$ ✓ For $n\geq 2$.

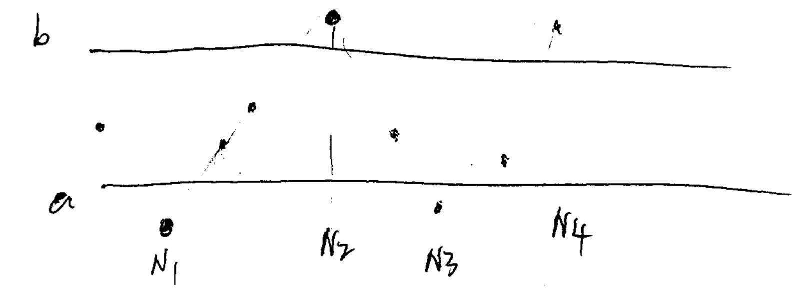
[proof. $E(T_n|\mathcal{F}_{n-1}) = E(T_{n-1} + \underline{H_n}(S_n - S_{n-1})|\mathcal{F}_{n-1})$

$= T_{n-1} + H_n \cdot E(S_n - S_{n-1}|\mathcal{F}_{n-1}) \leq T_{n-1}$.]

∵ 10.

<u>Upcrossings</u>: $\{S_n, n \geq 1\}$ is a submartingale w.r.t. $\mathcal{F}_n$.

H_m	0	0	1	1	1	0	0	1	1	...
m	1	2	3	4	5	6	7	8	9	



Let $N_1 = \inf\{m \geq 1 : S_m \leq a\}$.

$N_2 = \inf\{m > N_1 : S_m \geq b\}$.

$N_3 = \inf\{m > N_2 : S_m \leq a\}$ (stopping times)

$N_4 = \inf\{m > N_3 : S_m \geq b\}$.

$\vdots$

Let $U_n = \sup\{k : N_{2k} \leq n\}$ be the number of upcrossings by time n.

<u>Thm 5.2.7</u> (upcrossing inequality) $\{S_n\}$ submartingale $a < b$

$$\text{Then } E(U_n) \leq \frac{1}{b-a}\left(E(S_n-a)^+ - E(S_1-a)^+\right)$$

Proof: Let $H_m = \begin{cases} 1 & \text{if } \exists k \text{ s.t. } N_{2k-1} < m \leq N_{2k} \\ 0 & \text{otherwise.} \end{cases}$ $m \geq 2$,

Then H_m is predictable. $H_m \in \mathcal{F}_{m-1}$

and $\underline{(b-a)U_n} \leq \sum_{m=2}^{n} H_m(S_m - S_{m-1})$

$$\therefore (b-a)EU_n \leq \sum_{m=2}^{n} E[H_m(S_m - S_{m-1})]$$

$$= \sum_{m=2}^{n} E\left(\underbrace{E[H_m(S_m - S_{m-1}) | \mathcal{F}_{m-1}]}_{H_m \in \{0,1\}}\right) \leq \sum_{m=2}^{n} E(S_m - S_{m-1})$$

$$= E(S_n - S_1)$$

$\because$ 11.

$S_n' = a + (S_n - a)^+$ is also a submartingale.

U_n' upcrossing of $\{S_i'\}$.

then $U_n' = U_n$.

$\therefore (b-a) EU_n = (b-a) EU_n' \le E(S_n' - S_1')$

$= E(S_n - a)^+ - E(S_1 - a)^+$.]

Thm 5.2.8 Martingale convergence theorem.

S_n is a submartingale w.r.t. $\mathcal{F}_n$.

sup

$\liminf ES_n^+ < \infty$.

Then. $S_n \to S$ a.s. $E|S| < \infty$.

$(ES_n \to ES \times)$

[proof: U_n as above $\uparrow$ U.

$EU_n \uparrow EU$.

$EU = \liminf EU_n$

$\le \liminf \frac{E(S_n - a)^+}{b-a} \le \liminf \frac{|a| + ES_n^+}{b-a} < \infty$.

$P(\liminf S_n < \limsup S_n) \le \sum_{\substack{a<b \\ \text{rational numbers}}} P(\liminf S_n < a < b < \limsup S_n) = 0$

$\therefore S_n \to S$ a.s.

12.

$$ES^+ \overset{\text{Fatou}}{\le} \liminf ES_n^+ < \infty.$$

$$ES^- \le \liminf ES_n^- = \liminf \left(ES_n^+ - ES_n \right)$$

$$\le \liminf \left(ES_n^+ - ES_0 \right)$$

$$< \infty.$$

$\therefore$ $E|S| < \infty$.]

- Recall $X_1, X_2, \ldots$ indep $EX_r = 0$. $EX_r^2 < \infty$. $\sum_{r\ge1}^{\infty} E(X_r^2) < \infty$.

then $S_n = \sum_{r=1}^{n} X_r$ converges a.s

[new proof: $ES_n^+ \le E|S_n| \le (ES_n^2)^{1/2} = (\sum_{r\ge1}^{n} EX_r^2)^{1/2} < (\sum_{r\ge1}^{\infty} EX_r^2)^{1/2} < \infty$.]

used to prove LLN

~~$n \ge 1$ $X_1 = S_1$~~

~~S_n martingale. $X_n = S_n - S_{n-1}$, $n \ge 2$ martingale differences.~~

~~Recall $EX_n = 0$, $n \ge 2$.~~

$$\cancel{E(S_n^2) = \sum_{r\ge1}^{n} E(X_r^2)}$$

$$\cancel{\text{If } \sum_{r\ge1}^{n} EX_r^2 < \infty.}$$

13.

$X_1, X_2, \ldots$ not necessarily indep.

· identically distributed ~~identically distributed~~ ~~$E|X_i| < \infty$~~

$E|X_i| < \infty$

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - EX_i) \overset{?}{\longrightarrow} 0$$

$\parallel$ if $X_1 = X_2 = X_3 = \cdots$

$X_1 - EX_1$

$\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$

$S_1 = X_1$

$S_2 = S_1 + X_2 - E(X_2|\mathcal{F}_1)$

$S_n = S_{n-1} + X_n - E(X_n|\mathcal{F}_{n-1})$

is a martingale w.r.t. $\{\mathcal{F}_n\}$

Thm $\frac{S_n}{n} \to 0$ in prob. ~~Suppose $\{X_i\}$ are identically distributed~~

Let $Y_i = X_i I(|X_i| \le i)$, $i \ge 1$, $T_n = T_{n-1} + Y_n - E(Y_n|\mathcal{F}_{n-1})$, $n \ge 2$; $T_1 = Y_1$

① $\sum_{i=1}^{\infty} P(Y_i \ne X_i) = \sum_{i=1}^{\infty} P(|X_i| > i) < \infty \quad (E|X_i| < \infty)$

② $\sum_{i=2}^{\infty} \frac{E[Y_i - E(Y_i|\mathcal{F}_{i-1})]^2}{i^2} \le \sum_{i=2}^{\infty} \frac{E(Y_i^2)}{i^2} \overset{\text{proved before}}{\le} CE|X_1| < \infty$

Martingale convergence theorem $\Rightarrow$ $\sum_{i=2}^{\infty} \frac{[Y_i - E(Y_i|\mathcal{F}_{i-1})]}{i}$ converges a.s.

Kronecker's lemma $\Rightarrow$ $\frac{1}{n}\sum_{i=1}^{n}(Y_i - E(Y_i|\mathcal{F}_{i-1})) \to 0$ a.s.

③. N.T.S.

$$\frac{1}{n}\sum_{i=2}^{n} E(Y_i|\mathcal{F}_{i-1}) - \frac{1}{n}\sum_{i=2}^{n} E(X_i|\mathcal{F}_{i-1}) \to 0 \text{ in prob.}$$

$\Uparrow$

$E\left|E(X_i - Y_i|\mathcal{F}_{i-1})\right|$

$\le E\left|E[X_i I(X_i > i)|\mathcal{F}_{i-1}]\right|$

$\le E|X_i I(X_i > i)|$

$= E|X_1 I(X_1 > i)|$

$\to 0$ as $i \to \infty$.

14. <u>Thm 5.2.9</u> (S_n) is a supermartingale w.r.t. $\{\mathcal{F}_n\}$.

If $S_n \geq 0$, then $S_n \to S$ a.s. and $E(S) \leq ES_1 < \infty$.

[proof: $-S_n \leq 0$ is a submartingale.

$$E((-S_n)^+) = 0. \quad \therefore S_n \to S \text{ a.s.}$$

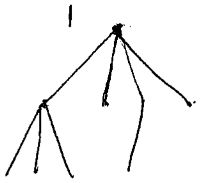
$$E(S) = E(\lim S_n) \overset{\text{Fatou}}{\leq} \liminf ES_n \leq ES_1 \,]$$

<u>Example:</u> Branching processes:

Z_1

Z_2

Z_3



$Z_1 = 1$

~~$Z_2 = \xi$~~

$\xi_i^{(m)} \sim$ i.i.d.

$$Z_{n+1} = \begin{cases} \xi_1^{(n+1)} + \cdots + \xi_{Z_n}^{(n+1)} & \text{if } Z_n > 0 \\ 0 & \text{if } Z_n = 0 \end{cases}$$

$$p_k = P(\xi_i^{(n)} = k), \quad k = 0, 1, 2, \ldots$$

$$\mathcal{F}_n = \sigma(\xi_i^{(m)} : i \geq 1, \ 1 \leq m \leq n)$$

$$\mu = E\xi_i^m \in (0, \infty)$$

Then: $\frac{Z_n}{\mu^n}$ is a martingale w.r.t. $\mathcal{F}_n$.

[proof: $E\left(\frac{Z_{n+1}}{\mu^{n+1}} \middle| \mathcal{F}_n\right) = \frac{1}{\mu^{n+1}} E(\xi_1^{(n+1)} + \cdots + \xi_{Z_n}^{(n+1)} \mid \mathcal{F}_n)$ $\boxed{\text{Lecture 11 ended here}}$

$$= \sum_{k=1}^{\infty} \frac{1}{\mu^{n+1}} I(Z_n = k) E(\xi_1^{(n+1)} + \cdots + \xi_k^{(n+1)} \mid \mathcal{F}_n) = \frac{\mu}{\mu^{n+1}} \sum_{k=1}^{n} k \cdot I(Z_n = k) = \frac{1}{\mu^n} Z_n \,]$$

$\therefore \frac{Z_n}{n} \to \text{limit a.s.}$ if $\mu \leq 1$ $\lim = 0$. If $\mu > 1$: <u>Thm 5.3.10</u>: $\lim \neq 0 \Leftrightarrow \sum p_k \cdot k \log k < \infty$.

0.

$\{S_n, \mathcal{F}_n\}$ $n \geq 1$ submartingale if

(1). $E|S_n| < \infty$.

(2) $S_n \in \mathcal{F}_n$

(3). $E(S_n | \mathcal{F}_{n-1}) \geq S_{n-1}$. $n \geq 2$.

---

predictable sequence $\Leftarrow$

upcrossing inequality.

• Martingale convergence theorem. $\sup_{n \geq 1} E S_n^+ < \infty$.

then . $S_n \to S$ a.s. $E|S| < \infty$.

$\Downarrow$

i.d. $X_1, \ldots$ $EX_1 = 0$.

• LLN. $\dfrac{\sum_{i=1}^{n} \left(X_i - E(X_i | \mathcal{F}_{i-1}) \right)}{n} \to 0$ in prob

Recall Kolmogorov's maximal inequality,

$X_1, \ldots$ indep. $EX_i = 0$. $EX_i^2 < \infty$. $S_k = \sum_{i=1}^{k} X_i$

then $\forall x > 0$. $P(\underbrace{\max_{1 \leq k \leq n} |S_k| \geq x}) \leq \dfrac{E(S_n^2)}{x^2}$

proof by ~~dividing~~ decomposing ~~or~~ Ω according to ~~when the~~ the ~~first~~ ~~or~~ time $\geq x$.

$\Longrightarrow$ Lecture 12.

15. $\mathcal{F}_n \uparrow$ • Recall the proof of Kolmogorov's maximal inequality

Thm 5.2.6 N is a stopping time if $\{N=k\} \in \mathcal{F}_k$

$\{N \geq n\} \in \mathcal{F}_n \quad \forall n \geq 1.$

w.r.t. $\mathcal{F}_n$

Thm $\{S_n\}$ is a submartingale then.

$\{T_n := S_{n \wedge N}\}$ is a submartingale.

[proof:

~~$T_n = \sum_{k=1}^{n} S_k I(N=k) + S_n I(N>k) \in \mathcal{F}_n \qquad E|T_n| < \infty$~~

• $E(T_n | \mathcal{F}_{n-1}) = T_{n-1} + E(X_n I(N \geq n) | \mathcal{F}_{n-1})$

$S_n = \sum_{i=1}^{n} X_i$

• $T_n = S_{n \wedge N} = \sum_{i=1}^{n \wedge N} X_i = \sum_{i=1}^{n} X_i I(N \geq i) \in \mathcal{F}_n. \qquad E|T_n| < \infty$

$= T_{n-1} + \underbrace{I(N \geq n)}_{\geq 0} \cdot \underbrace{E(X_n | \mathcal{F}_{n-1})}_{\geq 0} \geq T_{n-1}.$ ∎]

• Under the above setting:

$$ES_1 = ET_1 \leq ET_n = \sum_{i=1}^{n} E X_i I(N \geq i)$$

$$= \sum_{i=2}^{n} E\, E(X_i I(N \geq i) | \mathcal{F}_{i-1}) + EX_1$$

$$\leq \sum_{i=2}^{n} E\, (E(X_i | \mathcal{F}_{i-1})) + EX_1$$

$$= \sum_{i=1}^{n} EX_i = ES_n.$$

1

16. Thm 5.4.2 Doob's inequality $\{S_n, \mathcal{F}_n\}$

$\{S_n\}$ submartingale w.r.t. $\{\mathcal{F}_n\}$, $n \geq 1$.

$\forall x > 0$.

$$P\left(\max_{1\le k\le n} S_k^{+} \geq x\right) \leq \frac{1}{x} E\left(S_n^{+} \cdot I\left(\max_{1\le k\le n} S_k^{+} \geq x\right)\right)$$

$$\leq \frac{E(S_n^+)}{x}$$

Recall: Kolmogorov maximal inequality. $X_1, X_2, \ldots$ indep. $E(X_i) = 0$.

$$P\left(\max_{1\le k\le n} |S_k|^2 \geq x^2\right) \leq \frac{ES_n^2}{x^2}$$

proof: Let $N = \inf\{k : S_k \geq x\}$.

$$A = \left\{\max_{1\le k\le n} S_k \geq x\right\} = \{S_{n\wedge N} \geq x\} = \{N \leq n\}$$

$$1_A \leq \frac{S_{n\wedge N}}{x} \cdot 1_A$$

$$P(A) = E(1_A) \leq \frac{1}{x} E(S_{n\wedge N} 1_A)$$

$$= \frac{1}{x}\left(E(\underbrace{S_{n\wedge N}}_{T_n}) - E(S_{n\wedge N}) \cdot 1_{A^c}\right)$$

$$\leq \frac{1}{x}\left(E(S_n) - ES_n 1_{A^c}\right) = \frac{1}{x} E(S_n 1_A)$$

$\boxed{\text{change } S_k \text{ by } S_k^+ \text{ as above}}$

17. Thm 5.4.3. L^p maximum inequality.

$\{S_n\}$ submartingale, $p>1$.

then. $$E\left[\left(\max_{1\le k\le n} S_k^+\right)^p\right] \le \left(\frac{p}{p-1}\right)^p \cdot E\left[(S_n^+)^p\right].$$

[proof: $$\boxed{E\left[\left(\max_{1\le k\le n} S_k^+ \wedge M\right)^p\right]}$$

$$= \int_0^\infty p\lambda^{p-1} \cdot \underline{P\left(\max_{1\le k\le n} S_k^+ \wedge M \ge \lambda\right)} d\lambda$$ $M<\lambda$ both $=0$

$$\overset{\text{Doob's ineq.}}{\le} \int_0^\infty p\lambda^{p-1} \cdot \underline{\frac{1}{\lambda} E\left(S_n^+ \cdot I\left(\max_{1\le k\le n} S_k^+ \wedge M \ge \lambda\right)\right)} d\lambda$$

$$= E S_n^+ \int_0^{\max_{1\le k\le n} S_k^+ \wedge M} p\lambda^{p-2} d\lambda$$

$$= \frac{p}{p-1} E S_n^+ \left(\max_{1\le k\le n} S_k^+ \wedge M\right)^{p-1}$$ $\frac{1}{q} = 1 - \frac{1}{p}$

$$\overset{\text{Holder's ineq.}}{\le} \frac{p}{p-1} \cdot \left(E (S_n^+)^p\right)^{1/p} \cdot \boxed{\left(E\left(\max_{1\le k\le n} S_k^+ \wedge M\right)^p\right)^{\frac{1}{q}}}$$ $\frac{p}{q} = (p-1)q$

$$\Rightarrow \quad E\left[\left(\max_{1\le k\le n} S_k^+ \wedge M\right)^p\right] \le \left(\frac{p}{p-1}\right)^p E(S_n^+)^p.\ ]$$

18

- If $\{S_n\}$ is a martingale

then. $$E\left(\max_{1\le k\le n}|S_k|\right)^p \le \left(\frac{p}{p-1}\right)^p E\left(|S_n|^p\right)$$

Thm 5.4.5 L^p convergence theorem.

$p>1$.

$\{S_n\}$ is a martingale, if $\sup E|S_k|^p < \infty$.

then $S_n \to S$ a.s. and $E|S_n - S|^p \to 0$.

[proof.

$$\Rightarrow E\left(\boxed{\sup_{k\ge 1}|S_k|}\right)^p \le \left(\frac{p}{p-1}\right)^p \cdot \sup\left(E|S_n|^p\right) < \infty$$

$$|S_n - S|^p \le \left(2 \sup_{1\le k\le\infty}|S_k|\right)^p$$

DCT: $E|S_n - S|^p \to 0$.]

Burkholder's inequality: If $\{S_n, \mathcal{F}_n, n\ge 1\}$ is a martingale, $p>1$.

then $$E|S_n|^p \le C_p \cdot E\left|\sum_{i=1}^{n} X_i^2\right|^{p/2}$$

- Rosenthal's inequality for independent sums.

19. Martingale limit theory and its application.

P. Hall / C.C. Heyde (1981)

$\{S_n, \mathcal{F}_n\}$ martingale. ~~mean zero~~ zero-mean square integrable $\cancel{ES_1 = 0}$. $\mathcal{F}_0 = \{\emptyset, \Omega\}$ $S_0 = 0$.

$\cancel{X_1 = S_1}$ $X_n = S_n - S_{n-1}$ $n \geq 1$.

$\sigma_n^2 = E(X_n^2 | \mathcal{F}_{n-1})$, $B_n^2 = \sum_{i=1}^{n} E\sigma_i^2$.

$E|X_n|^4 < \infty$. $n \geq 1$.

Then. $\sup_{x \in \mathbb{R}} \left| P\left(\frac{S_n}{B_n} \leq x\right) - \Phi(x) \right|$

$$\leq C \left\{ \frac{1}{B_n^4} \left(\sum_{i=1}^{n} E|X_i|^4 + E\left|\left(\sum_{i=1}^{n} \sigma_i^2\right) - B_n^2\right|^2 \right) \right\}^{\frac{1}{5}}$$

① ②.

Strong approximation to be added.

20.

Skorokhod representation.

Recall Skorokhod's theorem.

Thm. $\{S_n = \sum_{i=1}^{n} X_i,\ \mathcal{F}_n,\ n \geq 1\}$ is a zero-mean, square-integrable martingale.

then, we can define it and a standard Brownian motion $\{B_t, t \geq 0\}$.

and $\tau_1, \tau_2, \cdots \geq 0$ on the same probability space.

such that with $T_n = \tau_1 + \cdots + \tau_n$, and $S_n' = B_{T_n}$, $X_n' = S_n' - S_{n-1}'$

(i) $\{\boxed{S_n}, n \geq 1\} \overset{d}{=} \{\boxed{S_n'}, n \geq 1\}$ $\qquad \mathcal{G}_n = \sigma(S_1', \cdots, S_n', B_{[0,T_n]})$

(ii) $E(\tau_n | \mathcal{G}_{n-1}) = E(X_n'^2 | \mathcal{G}_{n-1})$.

One usage:

distribution of S_n ; determined by $\mathcal{L}(T_n)$

$$T_n = \sum_{i=1}^{n} \tau_i = \underbrace{\sum_{i=1}^{n} (\tau_i - E(\tau_i | \mathcal{G}_{i-1}))}_{①} + \sum_{i=1}^{n} E(\tau_i | \mathcal{G}_{i-1})$$

martingale

$$\| $$

$$\sum_{i=1}^{n} E(X_n'^2 | X_1', \cdots, X_{n-1}')$$

②

If ② $\approx B_n^2$ and ① $\ll$ ②

then $T_n \approx B_n^2$ and.

$$\frac{S_n}{B_n} \overset{d}{\approx} N(0,1).$$