1/22 HTOP 1

Monday, January 22, 2024 11:15 AM

Wei Wu

Office Hour: $\begin{cases} M: 4:30 \sim 5:30 \text{ pm}. \\ W: 3:30 \sim 4:30 \text{ p.m.} \end{cases}$ W910

Difficulties

Prob & Stats → TOP → HTOP → PLT

Some Rrefs:

Probability & Measures: P. Billingsley

Essentials of Stochastic Processes: R. Durret

Introduction to Stoch Process: G. Lawler

Higher Level Refs:

- Probability: Theory and Example, R. Durret
- Probability with martingales, D. Williams
- A course in probability theory, Kai-Lai Chung (outdated a little bit)

Homework: weekly, due Friday at Reci.

Grade: 5% participation, 15% Homework, 40% Midterm, 40% Final ← May 8th

Topics:

- Probability Space, σ-algebras, measures, conditional prob. and independence
- Measurable functions, random variables, and their distributions
- Integration / Expectation:
  Conditional distribution / expectation
  Functions of R.V.s
- Random walks
- Generating functions; characteristic functions
- Branching process
- Convergence of R.V.s, law of large numbers
- Central limit theorems
- Large Deviations (Time permitting)
  Markov Chains (⋯, ⋯)

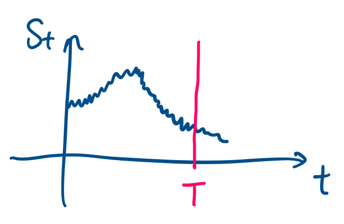
§

Probability Space

eg. Roll a dice, $\Omega = \{1, 2, \cdots, 6\}$

$\mathbb{P}(\{i\}) = \frac{1}{6}$, For any $A \subseteq \Omega$, $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$

eg. Stock Price



(Geometric) Brownian Motion

continuous but nowhere differentiable

Event: $A = \{S_0 = 1, S_T < 1\}$, $\mathbb{P}(A) = ?$ (measure)

Indeed. $\Omega = C[0,T] = \{f: [0,T] \to \mathbb{R},\ f \text{ is continuous}\}$

Probability Space.

$(\Omega, \mathcal{F}, \mathbb{P})$

sample space; σ-algebra (collection of certain sets/events in Ω) ("useful" subset of Ω); measure ($\mathbb{P}: \mathcal{F} \to [0,1]$)

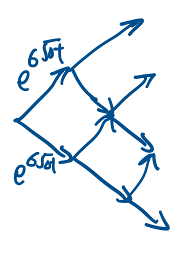
Rmk. In general, one cannot compute prob. for "all" sets

Eg. Discrete Time Stock Model: $t=0$ to $t=T$ ← maturity time

time step $\Delta t \ll T$

price go ↗ by a factor $e^{\sigma \cdot \sqrt{\Delta t}}$
↘ ⋯ $e^{-\sigma \cdot \sqrt{\Delta t}}$, $\sigma > 0$ votality.

$N = \frac{T}{\Delta t}$



if $N = 3$, $\Omega = \{(x_1, x_2, x_3): x_i = 0 \text{ or } 1\}$ (0 = down, 1 = up)

$\Omega_N = \{(x_1, \cdots x_N): x_i = 0 \text{ or } 1\}$

$\forall \omega \in \Omega_N$

$S_N(\omega) = S_0\, e^{\sigma\sqrt{\Delta t} \cdot \sum_{i=1}^{N} x_i} \cdot e^{-\sigma\sqrt{\Delta t} \cdot (N - \sum_{i=1}^{N} x_i)}$

Random Variable

Events: $\{\omega \in \Omega_N,\ S_N \in [0.95 S_0, S_0)\}$

Eg. Gambling.

Start with $\mathbb{N}^+$ amount of money,

each time bet an integer amount

If amt of money $= 0$, stays at 0

Wealth Process:

$\Omega = \mathbb{N} \times \mathbb{N} \times \mathbb{N} \times \cdots$

$= \{(x_1, x_2, \cdots),\ x_i \in \mathbb{N}\}$

Wealth at time N: (R.V.)

$X_N: \Omega \to \mathbb{N}$ projection map

$(x_1, x_2, \cdots) \to x_N$

Event 1: {start at wealth i, reach at wealth j at some points}

$= \{\omega \in \Omega: \text{s.t. } \exists n \in \mathbb{N},\ X_0(\omega) = i,\ X_n(\omega) = j\}$

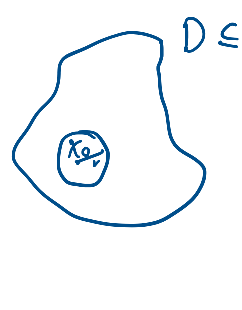
$= \bigcup_{n=0}^{+\infty} \{\omega \in \Omega: X_0(\omega) = i,\ X_n(\omega) = j\}$

Event 2: $\{(X_n) \text{ absorbs at } 0\}$

$= \{\omega \in \Omega: X_n(\omega) = 0\}$

$= \bigcup_{n=0}^{+\infty} \underbrace{\{\omega \in \Omega: X_n(\omega) = 0\}}_{\{X_n = 0\}}$

Eg.



$D \subseteq \mathbb{R}^2$

$\Omega = D$

location of a particle in D

Location: $\forall \omega \in \Omega$, $L(\omega) = \omega$

Events. $\forall x_0 \in D,\ r > 0$,

$\{L \notin B_r(x_0)\} = \{\omega \in \Omega;\ L(\omega) \notin B_r(x_0)\}$

Eg': trajectory of particle in $[0,T]$ marked at time 0

$\Omega = C([0,T]; D)$

$= \{f: [0,T] \to D,\ f \text{ is continuous}\}$

location at time $t \in [0,T]$

$\forall \omega \in \Omega,\ L_t(\omega) = \omega(t)$

1/24 HTOP 2

Wednesday, January 24, 2024 11:16 AM

Probability Space $(\Omega, \mathcal{F}, \mathbb{P})$

$\uparrow$ 6-algebra: collection of "meaningful" subset of Ω

In probability, if A, B are disjoint points, expect, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

$\mathbb{P}(A^c) = \underbrace{\mathbb{P}(\Omega)}_{1} - \mathbb{P}(A)$

would like $\mathcal{F}$ to be closed under set operations: $\cup, \cap, c$

Def. Let $\mathcal{A}$ be a set of subsets of Ω, $\mathcal{A}$ is an algebra iff

(1) $\Omega \in \mathcal{A}$

(2) if $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$

(3) if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$

$\Rightarrow A \cap B = (A^c \cup B^c)^c \in \mathcal{A}$

Rmk.

① $\mathcal{A} = \{\phi, \Omega\}$, trivial algebra (smallest)

② $P(\Omega) = \{A : A \subseteq \Omega\}$ is an algebra (biggest) Power set

③ Let $\mathcal{A}_1, \mathcal{A}_2$ be algebras, $\mathcal{A}_1 \cap \mathcal{A}_2 = \{B \subseteq \Omega : B \in \mathcal{A}_1 \ \& \ B \in \mathcal{A}_2\}$ is an algebra (ex.)

More generally, let $(\mathcal{A}_j)_{j \in J}$ be a family of algebras, then $\bigcap_{j \in J} \mathcal{A}_j$ is an algebra (ex.)

④ Let $\mathcal{E}$ be a set of subsets of Ω. Then $a(\mathcal{E}) = \bigcap_{\substack{\mathcal{A} \text{ algebra} \\ \mathcal{A} \supseteq \mathcal{E} \text{ (all)}}} \mathcal{A}$ is an algebra generated by $\mathcal{E}$.

* The generated algebra is "the smallest algebra that contains $\mathcal{E}$"

eg. Let $A \subsetneq \Omega$

$\mathcal{E} = \{\phi, A\}$, $a(\mathcal{E}) = \{\phi, A, A^c, \Omega\}$

pf. let $\mathcal{F} = \{\phi, A, A^c, \Omega\}$

$a(\mathcal{E}) \subseteq \mathcal{F}$: it is easy to check $\mathcal{F}$ is an algebra.

Then $a(\mathcal{E}) \subseteq \mathcal{F}$ by the minimality of $a(\mathcal{E})$

$\mathcal{F}$ is an algebra

$a(\mathcal{E}) \supseteq \mathcal{F}$: Since $a(\mathcal{E})$ is an algebra,

$\therefore \because \phi, A \in a(\mathcal{E}) \Rightarrow \Omega, A^c \in a(\mathcal{E})$

$\therefore a(\mathcal{E}) \supseteq \mathcal{F}$.

elements in $\mathcal{F}$ are in $\mathcal{A}$

Hence $a(\mathcal{E}) = \mathcal{F} = \{\phi, A, A^c, \Omega\}$ is the generated algebra of $\mathcal{E}$

\#

⑤ Let $\Pi = \{A_1, A_2, \cdots, A_m\}$ be a partition of Ω

i.e. $A_i \cap A_j = \phi$, and $\bigcup_{i=1}^{m} A_i = \Omega$

$a(\Pi) =$ "finite disjoint union of $(A_i)_{i=1}^{m}$"

$= \{\bigcup_{i \in I} A_i,$ for some $I \subseteq \{1, 2, \cdots, m\}\}$ (ex.) This seems just to be the power set of Π

⑥ Let $\mathcal{A}$ be an algebra of $\mathbb{R}$ and $X : \Omega \to \mathbb{R}$

Then $\{\underbrace{X^{-1}(A)}_{w \in \Omega,\ X(w) \in A} : A \in \mathcal{A}\}$ is an algebra of Ω (ex.) Hint: use $X^{-1}(A \cup B) = X^{-1}(A) \cup X^{-1}(B)$

⑦ $\Omega = \mathbb{R}$, $\mathcal{E} =$ "left open, right closed intervals"

$$= \begin{cases} (a, b], & -\infty \le a < b < +\infty \\ (a, +\infty), & a \in \mathbb{R} \end{cases}$$

$a(\mathcal{E}) = \{I_1 \cup I_2 \cup \cdots \cup I_n, I_k \in \mathcal{E}, I_k \cap I_\ell = \phi\}$ (ex.) Hint: $\cdot\ \mathcal{F} \subseteq a(\mathcal{E})$ is simple, since $I_j \in a(\mathcal{E})$, $\bigcup_{j=1}^{n} I_j \in a(\mathcal{E})$

$\cdot\ \mathcal{F}$ is an algebra. $(\cdots)$

6-algebra:

Def. Let $\mathcal{F}$ be a set of subsets of Ω. Then $\mathcal{F}$ is an 6-algebra (6-field) iff

(1) $\Omega \in \mathcal{F}$

(2) if $B_1, B_2, \cdots \in \mathcal{F}$, then $\bigcup_{i=1}^{+\infty} B_i \in \mathcal{F}$

(3) if $B \in \mathcal{F}$, then $B^c \in \mathcal{F}$

$\Rightarrow \bigcap_{i=1}^{+\infty} B_i \in \mathcal{F}$

6-algebra: "collection of information based on observation"

eg. Coin flips each time 0 or 1

Infinite coin flips, $\Omega = \{0,1\}^\infty$

Observing 1st coin flip, $\mathcal{F}_1 = \{\phi, \Omega, A_0, A_1\}$ is a 6-algebra

$A_0 = \{(0, x_2, x_3, \cdots), x_i = 0 \text{ or } 1\}$

$A_1 = \{(1, x_2, x_3, \cdots), x_i = 0 \text{ or } 1\}$

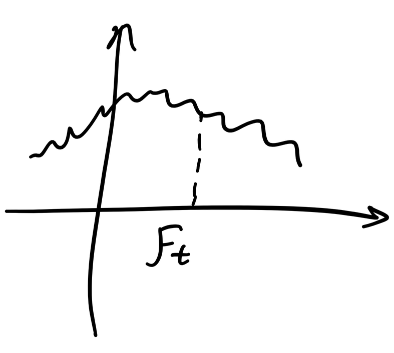
A_{00}	A_{01}
A_{10}	A_{11}

Observing 2nd coin flip, $\mathcal{F}_2 = \{\phi, \Omega, \bigcup_{i \in I} A_i, I = \{00, 01, 10, 11\}\}$ is a 6-algebra

nth coin flip, $\mathcal{F}_n = \{\phi, \Omega, \bigcup_{i \in I} A_i, I = \{0,1\}^n\}$ is a 6-algebra

$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \cdots$

filtration



Brownian motion

stochastic process

Prop. ① Intersections of 6-algebra is 6-algebra

② For any collection of subsets, denoted by $\mathcal{E}$

$6(\mathcal{E}) = \bigcap_{\substack{\mathcal{F} \text{ 6-algebra} \\ \mathcal{F} \supseteq \mathcal{E}}} \mathcal{F}$ is 6-algebra generated by $\mathcal{E}$

③ $a(\mathcal{E}) \subseteq 6(\mathcal{E})$

$6(a(\mathcal{E})) = 6(\mathcal{E})$ (ex.)

1/29 HTOP 3

Monday, January 29, 2024 11:21 AM

Recap:

$\Omega = \mathbb{R}$, $\mathcal{E} = \{$left open, right closed intervals$\} = \begin{cases} (a,b] , & -\infty \le a < b < +\infty \\ (a,+\infty) , & a \in \mathbb{R} \end{cases}$

$a(\mathcal{E}) = \{$ finite, disjoint union of elements of $\mathcal{E}\}$

$\sigma(\mathcal{E})$: " basically contains all nice subset of $\mathbb{R}$ "

Borel sets

- $(a,b) \in \sigma(\mathcal{E})$, $(a,b) = \bigcup_{n\ge 1} \underbrace{(a, b-\tfrac{1}{n}]}_{\in \mathcal{E}} \in \sigma(\mathcal{E})$
- $\{a\} \in \sigma(\mathcal{E})$, $\{a\} = \bigcap_{n\ge1} (a-\tfrac{1}{n}, a]$
- $[a,b) \in \sigma(\mathcal{E})$
- any countable set (say $\mathbb{Q}$) is in $\sigma(\mathcal{E})$
- set of transcedental numbers is in $\sigma(\mathcal{E})$ $[\pi, e]$

  (b/c algebraic numbers is countable)

$(\Omega, \mathcal{F}, \mathbb{P})$

Content & Measure

We want , $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$, if $A \cap B = \emptyset$

Def. Let $\mathcal{A}$ be an algebra (of subsets of Ω). A set function $u : \mathcal{A} \to [0,+\infty)$ is called a $\boxed{\text{content}}$ iff:

① $u(\emptyset) = 0$

② Finite additivity : $u(A\cup B) = u(A) + u(B)$ if $A, B \in \mathcal{A}$, $A \cap B = \emptyset$

Lem. Let $u : \mathcal{A} \to [0,+\infty)$ be a content

Then $\forall A, B \in \mathcal{A}$: ① $u(A\cup B) + u(A\cap B) = u(A) + u(B)$

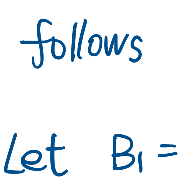
② if $A \subseteq B$, and $u(A) < +\infty$, then $u(B \setminus A) = u(B) - u(A)$

③ if $A \subseteq B$, then $u(A) \le u(B)$

④ if $A_1, \cdots, A_n \in \mathcal{A}$, then $u(\bigcup_{i=1}^n A_i) \le \sum_{i=1}^n u(A_i)$ subadditivity

⑤ if $(A_i)_{i=1}^\infty$ is a sequence of disjoint elements , and $\bigcup_{i=1}^\infty A_i \in \mathcal{A}$

Then $u(\bigcup_{i=1}^\infty A_i) \ge \sum_{i=1}^\infty u(A_i)$

pf. ①  We have $u(A) + u(B\setminus A) = u(A\cup B)$

$u(B) = u(A\cap B) + u(B\setminus A)$ $(*)$

$\Rightarrow u(A) + u(B) = u(A\cup B) + u(A \cap B)$

② follows from $(*)$

③ follows from ②

④ Let $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $\cdots$, $B_n = A_n \setminus \bigcup_{i=1}^{n-1} A_i$, $(B_i)_{i=1}^n$ disjoint

and $\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n A_i$. Therefore, $u(\bigcup_{i=1}^n A_i) = u(\bigcup_{i=1}^n B_i) = \sum_{i=1}^n u(B_i) = \sum_{i=1}^n u(A_i \setminus \bigcup_{j=1}^{i-1} A_i) \le \sum_{i=1}^n u(A_i)$

⑤ We have for $\forall n$

$u(\bigcup_{i=1}^\infty A_i) \ge u(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n u(A_i)$, send $n \to \infty$

we have $u(\bigcup_{i=1}^\infty A_i) \ge \sum_{i=1}^\infty u(A_i)$ since $u(A_i) \ge 0$

#

Rmk. To have $u(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty u(A_i)$, we need $\lim_{n\to\infty} u(\bigcup_{i=1}^n A_i) = u(\bigcup_{i=1}^\infty A_i)$ for disjoint A_i

eg. $\Omega = \mathbb{R}$, $\mathcal{A} = a(\mathcal{E})$ for any $A \in a(\mathcal{E})$

$u(A) := \lim_{L\to+\infty} \frac{|A\cap[0,L)|}{L}$ "density of A in $\mathbb{R}$"

then u is finitely additive but not countably additive

since $1 = u(\bigcup_{i=1}^{+\infty} [i, i+1)) > \sum_{i=1}^{+\infty} u([i,i+1)) = 0$

Def. A $\boxed{\text{measure}}$ on $(\Omega, \mathcal{F})$, u is a function $u : \mathcal{F} \to [0,+\infty]$

$\mathcal{F}$ ↳ σ-algebra

s.t. ① $u(\emptyset) = 0$

② Countable additivity : if $A_1, A_2, \cdots \in \mathcal{F}$ & $A_i \cap A_j = \emptyset$

Then $u(\bigcup_{i=1}^{+\infty} A_i) = \sum_{i=1}^{+\infty} u(A_i)$

If $u(\Omega) = 1$, then u is a $\boxed{\text{probability measure}}$ and $(\Omega, \mathcal{F}, \mathbb{P})$ is called a $\boxed{\text{probability space}}$

eg. $\Omega = \mathbb{R}$. Define a set function $m : a(\mathcal{E}) \to [0,+\infty]$ ($\mathcal{E}$ is in Recap)

Let $m((a,b]) = b-a$, $m((a,+\infty)) = +\infty$

and extend for every $A \in a(\mathcal{E})$, $m(A) = \sum_{j=1}^n m(I_j)$ if $A = \bigcup_{j=1}^n I_j$ and (I_j) disjoint

Then m is a content (ex.)

Lem. $m : a(\mathcal{E}) \to [0,+\infty]$ is countably additive

i.e. if $(A_k)_{k=1}^{+\infty} \subset a(\mathcal{E})$, (A_k) disjoint , and $A = \bigcup_{k=1}^{+\infty} A_k \in a(\mathcal{E})$

then $m(A) = \sum_{i=1}^{+\infty} m(A_k)$

Pf. Let $A = \bigcup_{i=1}^n I_i$ s.t. (I_i) disjoint , $I_i \in \mathcal{E}$

$A_k = \bigcup_{j=1}^{m_k} J_{jk}$, $(J_{jk})_j$ disjoint , $J_{jk} \in \mathcal{E}$

Then $m(A) = \sum_{i=1}^n m(I_i) = \sum_{i=1}^n \sum_{j=1}^{m_k} \sum_k m(I_i \cap J_{jk})$ (check by def.) $\underset{\text{(content)}}{=\!=} \sum_{j=1}^{m_k} \sum_{k=1}^{+\infty} m(J_{jk}) \underset{\text{(content)}}{=\!=} \sum_{k=1}^{\infty} m(A_k)$

#

$$\begin{array}{ll} m: & \mathcal{E} \longrightarrow [0,+\infty] \\ & \big\downarrow \text{extension} \\ & a(\mathcal{E}) \to [0,+\infty] \quad \text{Content, Countably additive} \\ & \big\downarrow \text{extension} \\ & \sigma(\mathcal{E}) \to [0,+\infty] \end{array}$$

Thm. m extends to a measure $m : \underset{\mathcal{B}(\mathbb{R})}{\overset{\shortparallel}{\sigma(\mathcal{E})}} \to [0,+\infty]$

It is the unique measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ s.t. $m([a,b]) = b-a$ $\boxed{\text{Carathéodory's extension theorem}}$

Dynkin π-λ theorem, Lebesgue measure

Def. If $\mathcal{F}$ is a σ-algebra, then $(\Omega, \mathcal{F})$ is called a measurable space

If u is a measure on $(\Omega, \mathcal{F})$. Then $(\Omega, \mathcal{F}, u)$ is a measure space

If $u(\Omega) = 1$, then $(\Omega, \mathcal{F}, u)$ is a probability space

Lem. Let $(\Omega, \mathcal{F}, u)$ be a measure space

(ex.) ① If $A_1, A_2, \cdots \in \mathcal{F}$, then $u(\bigcup_{i=1}^{+\infty} A_i) \le \sum_{i=1}^{+\infty} u(A_i)$ countable subaddivity

② $\boxed{\text{Continuity from below.}}$ If $(A_i)_{i=1}^{+\infty} \subseteq \mathcal{F}$, $A_1 \subseteq A_2 \subseteq \cdots$

then $u(\bigcup_{i=1}^{+\infty} A_i) = u(\lim_{i\to+\infty} A_i) = \lim_{i\to+\infty} u(A_i)$

③ $\boxed{\text{Continuity from above}}$: If $(A_i)_{i=1}^{+\infty} \subseteq \mathcal{F}$, $A_1 \supseteq A_2 \supseteq \cdots$ and $u(A_1) < +\infty$

then $u(\bigcap_{i=1}^{+\infty} A_i) = \lim_{i\to+\infty} u(A_i)$

pf. ② : let $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $\cdots$, $B_n = A_n \setminus A_{n-1}$, $\cdots$, (B_i) disjoint

$u(\bigcup_{i=1}^{+\infty} A_i) = u(\bigcup_{i=1}^{+\infty} B_i) \underset{\substack{\text{countably} \\ \text{additivity}}}{=} \sum_{i=1}^{+\infty} u(B_i) = \lim_{n\to+\infty} \left(\sum_{i=2}^n (u(A_i) - u(A_{i-1})) + u(A_1) \right) = \lim_{n\to+\infty} u(A_n)$

③ : let $B_j = A_1 \setminus A_j$, $B_j \nearrow$ (increasing)

Apply ② : $u(\bigcup_{j=1}^{+\infty} B_j) = \lim_{j\to+\infty} u(B_j) = \lim_{j\to+\infty} u(A_1) - u(A_j)$

$\shortparallel$

$u(A_1 \cap (\bigcap_{j=1}^{+\infty} A_j)^c) = u(A_1) - u(\bigcap_{j=1}^{+\infty} A_j)$

1/31 HTOP 4

Wednesday, January 31, 2024 11:16 AM

Recap: $\Omega = \mathbb{R}$, $\mathcal{E} = \begin{cases} (a,b], & -\infty \leq a < b < +\infty \\ (a,+\infty), & a \in \mathbb{R} \end{cases}$

$a(\mathcal{E}) = \{$ finite disjoint union of elements in $\mathcal{E}\}$

$m: a(\mathcal{E}) \longrightarrow [0,+\infty)$ s.t. $m([a,b)) = b-a$

m is a content on $(\Omega, a(\mathcal{E}))$ + countably additive

$\Rightarrow$ m extends to a measure on $(\Omega, \sigma(\mathcal{E}))$

Carathéodory's Extension Theorem

Let $\mathcal{F}$ be an algebra on Ω, $u(A_i) < +\infty$ and $\bigcup_{i=1}^{\infty} A_i = \Omega$

Then u extends to a measure on $(\Omega, \sigma(\mathcal{F}))$

Lebesgue Measure: "$m(A) = \int_A dx$"

eg. Let $m_F: a(\mathcal{E}) \to \mathbb{R}$ s.t.

$m_F((a,b]) = F(b) - F(a)$ where F is right continuous, increasing $F: \mathbb{R} \to [0,1]$

$F(x) = \lim_{y \to x^+} F(y)$

with convention $F(+\infty) = \sup\{F(x), x \in \mathbb{R}\}$

$F(-\infty) = \inf\{F(x), x \in \mathbb{R}\}$

Extend m_F to $a(\mathcal{E})$ by finite additivity

m_F is countably additive on $a(\mathcal{E})$ (ex.)

By the extension Theorem, m_F extends to a measure on $(\mathbb{R}, \sigma(\mathcal{E}))$

Lebesgue - Stieltjes measure: "$m_F(A) = \int_A dF(x) = \mathbb{P}(X \in A)$"

Random Variable with distribution F

$(\Omega, \mathcal{F}, \mathbb{P})$

Special Case: $\Omega = \{1,2,\cdots,N\}$, $\mathcal{F} = P(\Omega)$ $[2^{\Omega}]$. Assume $\mathbb{P}(\{1\}) = \cdots = \mathbb{P}(\{N\}) = \frac{1}{N}$

for any $E \subseteq \Omega$, we have $\mathbb{P}(E) = \frac{|E|}{|\Omega|}$

prop. **(Inclusion- Exclusion)**

$$\mathbb{P}(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_{i=1}^{n} \mathbb{P}(E_i) - \sum_{i<j} \mathbb{P}(E_i \cap E_j) + \cdots + (-1)^{r+1} \sum_{i_1<i_2<\cdots<i_r} \mathbb{P}(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_r}) + \cdots + (-1)^{n+1} \mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_n)$$

Rmk. ① Prove by induction (ex.)

② Consequences: $\mathbb{P}(E_1 \cup \cdots \cup E_n) \geq \sum_{i=1}^{n} \mathbb{P}(E_i) - \sum_{i<j} \mathbb{P}(E_i \cap E_j)$

$\mathbb{P}(E_1 \cup \cdots \cup E_n) \leq \sum_{i=1}^{n} \mathbb{P}(E_i) - \sum_{i<j} \mathbb{P}(E_i \cap E_j) + \sum_{i<j<k} \mathbb{P}(E_i \cap E_j \cap E_k)$

$\vdots$

eg. (Birthday Problem) n people

$\mathbb{P}[$at least 2 of them have the same birthday$] = ?$

$\parallel$

Sol. $1 - \underbrace{\mathbb{P}[n \text{ different birthdays}]}$

$= \frac{365 \cdot 364 \cdot \cdots \cdot (365-n+1)}{365^n}$

$= 1 \cdot (1-\frac{1}{365})(1-\frac{2}{365}) \cdots (1-\frac{n-1}{365})$

$1-x \approx e^{-x}$ $\quad \approx 1 \cdot e^{-\frac{1}{365}} \cdot e^{-\frac{2}{365}} \cdot \cdots \cdot e^{-\frac{n-1}{365}} = e^{-\frac{n(n-1)}{730}}$

if $n \geq 23$, $\mathbb{P}($at least $2) > 50\%$

eg. (matching problem) N people picking hats at random

$\mathbb{P}[$No one picks his/her own hat$] = ?$

Sol. $E_i = \{ i$th person gets his/her hat $\}$

$1 - \mathbb{P}(E_1 \cup E_2 \cup \cdots \cup E_n)$

$\mathbb{P}(E_i) = \frac{(N-1)!}{N!}$, $\mathbb{P}(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_r}) = \frac{(N-r)!}{N!}$

$(-1)^{r+1} \sum_{i_1<i_2<\cdots<i_r} \mathbb{P}(E_{i_1} \cap \cdots \cap E_{i_r}) = (-1)^{r+1} C_N^r \frac{(N-r)!}{N!} = (-1)^{r+1} \frac{1}{r!}$

$\mathbb{P}(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_{r=1}^{N} (-1)^{r+1} \frac{1}{r!} \xrightarrow[N \to +\infty]{} e^{-1}$

$\mathbb{P}($No one gets the hat$) = 1 - \sum_{r=1}^{N} (-1)^{r+1} \frac{1}{r!} \xrightarrow[N \to +\infty]{} 1 - \frac{1}{e}$

ex. 10 couples sitting at a round table

$\mathbb{P}[$No one sits next to his/her partner$] = ?$

eg. Texas Holder: Hand of 5 cards

① $\mathbb{P}[$Straight$]$, 顺子, (不能同花顺)

② $\mathbb{P}[$full house$]$, 葫芦, ↓

A 2 3 4 5

↓

10 J Q K A

Sol. $|\Omega| = C_{52}^5$. $\mathbb{P}[\text{straight}] = \frac{10 \cdot (4^5 - 4)}{C_{52}^5} = 0.0039$

$\mathbb{P}[\text{full house}] = \frac{13 \times 12 \times C_4^3 \times C_4^2}{C_{52}^5} = 0.0014$

Conditional Probability

If event B occurs, what is the probability that A occurs?

N experiments, $N(B) \triangleq$ # of occurence of B

$N(A \cap B) \triangleq$ # of occurence of both A, B

Def. $\mathbb{P}(A|B) := \frac{N(A \cap B)}{N(B)} = \frac{N(A \cap B)/N}{N(B)/N} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ if $\mathbb{P}(B) > 0$

eg. Two kids Problem

① $\mathbb{P}[$two boys $|$ at least a boy$] = \frac{1}{3}$

because $\Omega = \{GG, GB, BG, BB\}$, $A \cap B = \{BB\}$, $B = \{GB, BG, BB\}$

② $\mathbb{P}[$two boys $|$ second kid is a boy$] = \frac{1}{2}$

③ $\mathbb{P}[$two boys $|$ at least a boy, born on Wednesday$] = \frac{13}{27}$

Sol. $\Omega = \{G_iG_j, G_iB_j, G_jB_i, B_iB_j, i = 1,\cdots,7\}$

$= \frac{\mathbb{P}[\text{two boys} \mid \text{at least one born on Wednesday}]}{\mathbb{P}[\text{at least a boy}]}$

$A \cap B = \{B_3B_j, B_iB_3\}$, $i,j \in \{1,\cdots,7\}$, #13 elements

$B = \{B_3B_j, B_iB_3, G_iB_3, B_3G_j\}$, #27 elements

2/2 HTOP (Reci 2)

Friday, February 2, 2024 11:18 AM

① Let $\mathcal{A}_1, \mathcal{A}_2$ be two algebras

then $\mathcal{A}_1 \cup \mathcal{A}_2 = \{E : E \in \mathcal{A}_1 \text{ or } E \in \mathcal{A}_2\}$ is an algebra iff $\mathcal{A}_1 \subseteq \mathcal{A}_2$ or $\mathcal{A}_2 \subseteq \mathcal{A}_1$

pf. $\Rightarrow$): sps. $\mathcal{A}_1 \subsetneq \mathcal{A}_2$ and $\mathcal{A}_2 \subsetneq \mathcal{A}_1$

take $A \in \mathcal{A}_1 \setminus \mathcal{A}_2$, $B = \mathcal{A}_2 \setminus \mathcal{A}_1$, then $A \cup B \in \mathcal{A}_1 \cup \mathcal{A}_2$

Since $\mathcal{A}_1 \cup \mathcal{A}_2$ is an algebra, $A \setminus B$, $B \setminus A$, $(A \setminus B) \cup (B \setminus A) \in \mathcal{A}_1 \cup \mathcal{A}_2$

$\Rightarrow A \setminus B$, $B \setminus A$, $(A \setminus B) \cup (B \setminus A) \in \mathcal{A}_1$ or $\mathcal{A}_2$, say $\mathcal{A}_1$

but $B = (B \setminus A) \cup (A \cap B) = \underset{\mathcal{A}_1}{(B \setminus A)} \cup (\underset{\mathcal{A}_1}{A} \setminus \underset{\mathcal{A}_1}{(A \setminus B)}) \in \mathcal{A}_2$

Contradiction.

$\Leftarrow$): Obvious

#

② Let Ω be a countably infinite set:

$\mathcal{A} = \{A \subseteq \Omega : A \text{ is finite} / A^c \text{ is finite}\}$

ⓐ Show that $\mathcal{A}$ is an algebra

ⓑ Define a set function: $u : \mathcal{A} \to [0, +\infty]$, $u(A) = \begin{cases} 0 & \text{if } A \text{ is finite} \\ +\infty & \text{if } A^c \text{ is finite} \end{cases}$ ← $|A|$, then u will be a measure

Is u a measure ?

Solution.

ⓐ $\forall A_1, A_2 \in \mathcal{A}$

$A_1 \cup A_2 = \begin{cases} A_1 \cup A_2 & \text{when } A_1^c \text{ and } A_2^c \text{ are infinite} \Leftrightarrow A_1 \text{ and } A_2 \text{ are finte} \\ (A_1^c \cap A_2^c)^c & \text{when } A_1^c \text{ or } A_2^c \text{ is finite} \end{cases} \in \mathcal{A}$

ⓑ No. Let $\Omega = \{w_1, w_2, \cdots\}$, Let $A_i = \{w_i\}$, $u(A_i) = 0$

but $u(\bigcup_i A_i) = u(\Omega) = +\infty$

③ Let $\Omega = \mathbb{R}$

$\mathcal{C}_1 = \{(-\infty, b],\ b \in \mathbb{R}\}$

$\mathcal{C}_2 = \{(a, b] :\ a, b \in \mathbb{R}\}$

$\mathcal{C}_3 = \{(a_1, b_1] \cup (a_2, b_2] \cup \cdots \cup (a_n, b_n],\ n \in \mathbb{N}^+,\ a_1, \cdots, a_n \in \mathbb{R} \cup \{-\infty\},\ b_1, \cdots, b_n \in \mathbb{R}\}$

Show that $\sigma(\mathcal{C}_1) = \sigma(\mathcal{C}_2) = \sigma(\mathcal{C}_3)$

Pf. $\mathcal{C}_1 \subseteq \mathcal{C}_3$, $\mathcal{C}_2 \subseteq \mathcal{C}_3$ $\Rightarrow \sigma(\mathcal{C}_1) \subseteq \sigma(\mathcal{C}_3)$ & $\sigma(\mathcal{C}_2) \subseteq \sigma(\mathcal{C}_3)$

We claim that $\sigma(\mathcal{C}_3) \subseteq \sigma(\mathcal{C}_1)$ ① & $\sigma(\mathcal{C}_3) \subseteq \sigma(\mathcal{C}_2)$ ② (ex.)

Pf. ①: Indeed $\mathcal{C}_3 \subseteq \sigma(\mathcal{C}_1)$, note that for each $(a_i, b_i] = \begin{cases} \overset{\in \mathcal{C}_1}{(-\infty, b_i]} & \text{if } a_i = -\infty \\ \underbrace{(-\infty, b_i] \setminus (-\infty, a_i]}_{\in \sigma(\mathcal{C}_1)} & \text{if } a_i > -\infty \end{cases} \in \sigma(\mathcal{C}_1)$

thus $(a_1, b_1] \cup \cdots \cup (a_n, b_n] \in \sigma(\mathcal{C}_1)$

$\Rightarrow \sigma(\mathcal{C}_3) \subseteq \sigma(\mathcal{C}_1)$ by minimality of $\sigma(\mathcal{C}_3)$

#

② is similar

#

④ Conditional Probability $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

$\boxed{\text{Law of total probability}}$: $\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)$

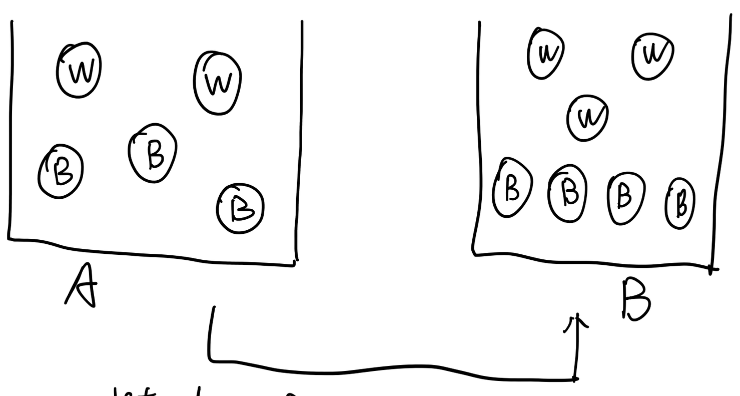
More generally, if $(B_i)_{i=1}^n$ is a partition of Ω, then $\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A|B_i)\mathbb{P}(B_i)$

pf. $RHS = \sum_i \mathbb{P}(A|B_i)\mathbb{P}(B_i) = \sum_i \mathbb{P}(A \cap B_i) = \mathbb{P}(A)$

#

eg.

2nd draw from B



1st draw from A, put it into B

$\mathbb{P}(\text{2nd draw is black})$

$= \mathbb{P}(\text{2nd is B} \mid \text{1st is B})\,\mathbb{P}(\text{1st is B}) + \mathbb{P}(\text{2nd is B} \mid \text{1st is W})\,\mathbb{P}(\text{1st is W})$

$= \frac{5}{8} \times \frac{3}{5} + \frac{1}{2} \times \frac{2}{5} = \frac{3}{8} + \frac{1}{5} = \frac{23}{40}$

#

Conversely, given A which B_i is most likely to occur ?

$\mathbb{P}(B_i|A) = \dfrac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(A)} = \dfrac{\mathbb{P}(A|B_i)\,\mathbb{P}(B_i)}{\sum_{i=1}^n \mathbb{P}(A|B_i)\,\mathbb{P}(B_i)}$ $\boxed{\text{Bayes formula.}}$

law of total probability

Recall: Bayes Formula: $\mathbb{P}(B_i \mid A) = \dfrac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(A)} = \dfrac{\mathbb{P}(A \mid B_i)\,\mathbb{P}(B_i)}{\sum \mathbb{P}(A \mid B_i)\,\mathbb{P}(B_i)}$

eg. Multiple Choice m choice

with prob. p knows the answer

with prob $1-p$ random guess

$\mathbb{P}[\underbrace{\text{knows the answer}}_{K} \mid \underbrace{\text{gets the answer}}_{A}]$

Sol. $\mathbb{P}(K \mid A) = \dfrac{\mathbb{P}(A \cap K)}{\mathbb{P}(A)} = \dfrac{\overset{1}{\mathbb{P}(A \mid K)}\,\overset{p}{\mathbb{P}(K)}}{\underbrace{\mathbb{P}(A \mid K)}_{1}\underbrace{\mathbb{P}(K)}_{p} + \underbrace{\mathbb{P}(A \mid K^c)}_{\frac{1}{m}}\underbrace{\mathbb{P}(K^c)}_{1-p}} = \dfrac{p}{p + \frac{1}{m}(1-p)}$

If $m = 4,\ p = \frac{1}{2} \implies \mathbb{P}(K \mid A) = \dfrac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2} \times \frac{1}{4}} = \dfrac{4}{5}$

#

eg. Covid tests

"False negative": $\mathbb{P}[\text{negative} \mid \text{contracted the virus}] = 5\%$

"False positive": $\mathbb{P}[\text{positive} \mid \text{healthy}] = 1\%$

Sps. 5% of the population got the virus

$\mathbb{P}[\underbrace{\text{contract the virus}}_{V} \mid \underbrace{\text{positive}}_{P}] = ?$

Sol. $\mathbb{P}[V \mid P] = \dfrac{\mathbb{P}[V \cap P]}{\mathbb{P}[P]} = \dfrac{\overset{0.05}{\mathbb{P}[V]}\,\overset{0.95}{\mathbb{P}[P \mid V]}}{\underbrace{\mathbb{P}[V]}_{0.05} \cdot \underbrace{\mathbb{P}[P \mid V]}_{0.95} + \underbrace{\mathbb{P}[V^c]}_{0.95} \cdot \underbrace{\mathbb{P}[P \mid V^c]}_{0.01}} \approx 83.3\%$

When only 0.1% of the population gets the virus $\approx 8.68\%$

eg. Prisoner's Paradox

A B C

2 executed, 1 pardoned

A asks: please tell me the name of someone else who will be killed

Guard: B will be killed

$\mathbb{P}[\underbrace{A \text{ survive}}_{A\,[\text{survive}]} \mid \underbrace{\text{Guard says } B \text{ will be killed}}_{B\,[\text{kill}]}]$

$= \dfrac{\mathbb{P}(B \mid A) \times \mathbb{P}(A)}{\mathbb{P}(B \mid A) \times \mathbb{P}(A) + \underbrace{\mathbb{P}(B \mid B)}_{0} \times \mathbb{P}(B) + \mathbb{P}(B \mid C) \times \mathbb{P}(C)}$

$= \dfrac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{6} + 1 \times \frac{1}{3}} = \dfrac{1}{3}$

$\mathbb{P}[C \text{ survive} \mid B] = \dfrac{\mathbb{P}[C \text{ survive}] \times \mathbb{P}[B \mid C \text{ survive}]}{\mathbb{P}[B]} = \dfrac{\frac{1}{3} \times 1}{\frac{1}{2}} = \dfrac{2}{3}$

eg. Two envelope

x 2x

switch ?

Wrong: # amount of money y: $\frac{1}{2} \times \underbrace{2y}_{\text{condition that } y \text{ is the smaller one}} + \frac{1}{2} \cdot \underbrace{\frac{y}{2}}_{\text{Condition that } y \text{ is the larger one}} = \frac{5}{4}y$

Independence:

If $\mathbb{P}[A \mid B] = \mathbb{P}[A]$, then we say A, B are independent

Def. (Elementary)

Events A, B are $\boxed{\text{independent}}$ iff $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$

Rmk. A, B are independent $\rightarrow$ A, B^c are independent

Multiple Events:

Def. The events $A_1, A_2 \cdots A_n$ are independent iff $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n)$

Def. The events $A_1, \cdots, A_n$ are pairwise independent iff $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\,\mathbb{P}(A_j)$

eg. Coin Flips X_1, X_2, X_3 with probability $\frac{1}{2} \rightarrow$ "0", $\frac{1}{2} \rightarrow$ "1"

$A_1 = \{X_1 = X_2\}$, $A_2 = \{X_2 = X_3\}$, $A_3 = \{X_1 = X_3\}$

- Pairwise Independence: $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(X_1 = X_2 = X_3) = \frac{1}{4}$

  $\mathbb{P}(A_i) = \frac{1}{2}$, $\therefore \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\,\mathbb{P}(A_j)$

- Not independent: $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(X_1 = X_2 = X_3) = \frac{1}{4} \neq \frac{1}{8}$

Def. A, B are $\boxed{\text{conditionally independent}}$ given C iff:

$\mathbb{P}(A \cap B \mid C) = \mathbb{P}(A \mid C)\,\mathbb{P}(B \mid C)$

eg. Independent trial, with success prob. p, failure prob. $1-p$

$\mathbb{P}[n \text{ success occurs before the } m^{th} \text{ failure}] = ?$

Sol. [Pascal]

Law of total probability:

$\mathbb{P}[\underbrace{n \text{ success before } m^{th} \text{ failure}}_{P_{n,m}}] = \mathbb{P}[\overset{n-1 \text{ success}}{n \text{ success before } m^{th} \text{ failure}} \mid 1\text{st success}]\ \mathbb{P}[1\text{st success}]$

$\quad + \mathbb{P}[\overset{\cdots \text{ before } m-1 \text{ failure}}{\cdots} \mid 1\text{st failure}] \cdot \mathbb{P}[1\text{st failure}]$

$\implies P_{n,m} = P_{n-1,m} \cdot p + P_{n,m-1} \cdot (1-p)$

Recursion:
$P_{n,m} = p \cdot P_{n-1,m} + (1-p) \cdot P_{n,m-1}$
Boundary Condition:
$P_{0,m} = (1-p)^m$
$P_{n,0} = 0$

Way: generating function

Sol. [Fermat] $\{n \text{ success before the } m^{th} \text{ failure}\}$

$= \{\text{at least } n \text{ success in the first } m+n-1 \text{ trials}\}$ (ex.)

$\mathbb{P}[k \text{ success in the first } m+n-1 \text{ trial}] = C_{m+n-1}^{k}\, p^k (1-p)^{m+n-1-k}$ $\boxed{\text{Binomial Distribution}}$

$\mathbb{P}[n \text{ success before } m^{th} \text{ failure}] = \sum_{k \geq n} C_{m+n-1}^{k}\, p^k (1-p)^{m+n-1-k}$

#

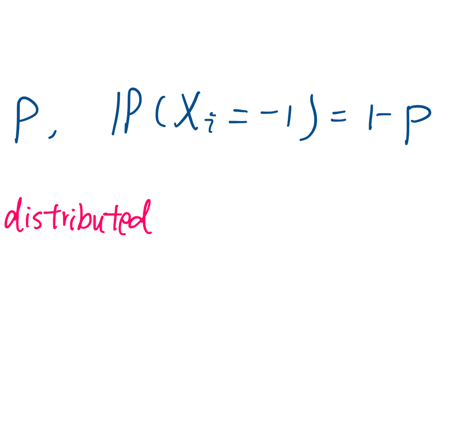
eg. Gambler's Ruin (1-d Random Walk)

Every time bet 1, with prob. $(1-p)$ lose

sps. initial amount of money is i

sps. stop either money reaches N or 0

$\mathbb{P}_i[\text{reach } N \text{ before } 0] = ?$



Sol. Let $S_n = X_1 + X_2 + \cdots + X_n + i$, X_i i.i.d, $\mathbb{P}(X_i = 1) = p$, $\mathbb{P}(X_i = -1) = 1-p$

i.i.d: independent & identically distributed

Let $P_i = \mathbb{P}[S \text{ hit } N \text{ before } 0]$

Condition on the first step:

$P_i = p \cdot P_{i+1} + (1-p) \cdot P_{i-1}$

Boundary Condition:

$P_0 = 0$, $P_N = 1$

* See the solution on the following note

2/5 HTOP 6

Monday, February 5, 2024 11:19 AM

1-d Random Walk

$S_n = X_1 + \cdots + X_n$ (X_i) i.i.d. $\mathbb{P}(X_i = 1) = p$, $\mathbb{P}(X_i = -1) = 1-p$

$P_i \triangleq \mathbb{P}(S \text{ reaches } N \text{ before } 0 \mid S_0 = i)$

$$\begin{cases} P_i = p \cdot P_{i+1} + (1-p) P_{i-1} \\ P_0 = 0 \\ P_N = 1 \end{cases}$$

Sol. $p \cdot (P_{i+1} - P_i) = \underbrace{(1-p)}_{q} (P_i - P_{i-1})$ $p+q=1$

$\rightarrow P_{i+1} - P_i = \frac{q}{p}(P_i - P_{i-1}) = \left(\frac{q}{p}\right)^i \cdot P_1$

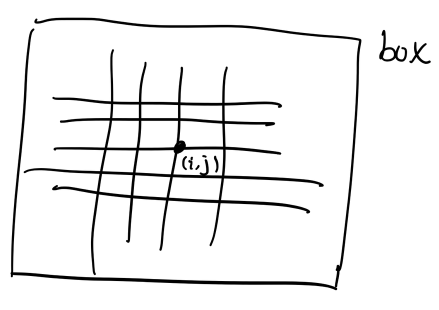
$\therefore P_i = \sum_{j=0}^{i-1} (P_{j+1} - P_j) = \left(1 + \cdots + \left(\frac{q}{p}\right)^{i-1}\right) P_1$

use $P_N = 1 \Rightarrow P_1 = \dfrac{1}{1 + \cdots + \left(\frac{q}{p}\right)^{N-1}}$

$$\therefore P_i = \frac{1 + \cdots + \left(\frac{q}{p}\right)^{i-1}}{1 + \cdots + \left(\frac{q}{p}\right)^{N-1}} = \begin{cases} \frac{i}{N}, & p = q \\ \frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^N} & \text{if } p \neq q \end{cases}$$

d-dimensional Random Walk

$\mathbb{Z}^2$



$S_n = X_1 + \cdots + X_n$, i.i.d.

$\mathbb{P}(X_1 = \pm e_1) = \mathbb{P}(X_1 = \pm e_2) = \frac{1}{4}$

Stop the R.W. when exiting the box $\square$ (as soon as touching the boundary)

What is $\mathbb{P}(S \text{ exit the box through } A) = ?$

Condition on the 1st step:

$$\begin{cases} P_{(i,j)} = \frac{1}{4} P_{(i+1,j)} + \frac{1}{4} P_{(i-1,j)} + \frac{1}{4} P_{(i,j+1)} + \frac{1}{4} P_{(i,j-1)} \\ P_{(i,j)} = 1 \quad \text{if } (i,j) \in A \\ P_{(i,j)} = 0 \quad \text{if } (i,j) \in \partial\square \setminus A \end{cases} \Longleftarrow \frac{1}{4} \sum_{u \sim (i,j)} (P_u - P_{(i,j)}) = 0$$

discrete laplacian

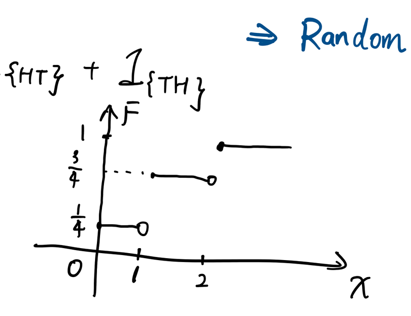
Random Variables & Measurable Function

eg. 2 coin flips: $\Omega = \{HH, HT, TH, TT\}$, $\mathcal{F} = \mathcal{P}(\Omega) = 2^\Omega$

$X = \#$ of heads

$= 2 \cdot \mathbb{1}_{\{HH\}} + \mathbb{1}_{\{HT\}} + \mathbb{1}_{\{TH\}}$ $\Rightarrow$ Random Variable

$F_X(x) = \mathbb{P}(X \leq x)$



Right Continuous

increasing

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

ⓐ A map $X: \Omega_1 \to \Omega_2$ is $\boxed{\text{m'ble}}$ iff $\forall A \in \mathcal{F}_2$, $X^{-1}(A) \in \mathcal{F}_1$

ⓑ If $(\Omega, \mathcal{F}_1, \mathbb{P})$ is a probability space, then m'ble function $X: \Omega_1 \to \Omega_2$ is called a Random Variable and $\mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A))$, $\forall A \in \mathcal{F}_2$ defines a $\boxed{\text{probability measure}}$

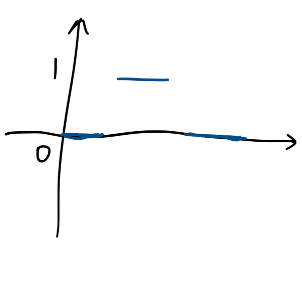
ⓒ If $(\Omega_2, \mathcal{F}_2) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then $X: \Omega_1 \to \mathcal{B}(\mathbb{R})$ is called a $\boxed{\text{Borel function}}$

ⓓ If $X: (\Omega, \mathcal{F}_1, \mathbb{P}_1) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then X is called an $\mathbb{R}$-valued random variable and $F_X(x) \triangleq P_X((-\infty, x]) = \mathbb{P}(X \leq x)$ is the $\boxed{\text{distribution function}}$ of X

$X^{-1}((-\infty, x]) = \{X \leq x\}$

eg. Consider $(\Omega, \mathcal{F})$ and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $A \in \mathcal{F}$

Then $X = \mathbb{1}_A$ is a m'ble function



To see this, take any $B \in \mathcal{B}(\mathbb{R})$, then $X^{-1}(B) = \begin{cases} \Omega, & B \ni \{0,1\} \\ A^c, & B \ni 0, B \not\ni 1 \\ A, & B \ni 1, B \not\ni 0 \\ \emptyset, & B \not\ni 1, B \not\ni 0 \end{cases}$

#

Indeed, we only need to check the preimage on a smaller set.

lemma: Let $(\Omega_1, \mathcal{F}_1)$, $(\Omega_2, \mathcal{F}_2)$ be m'ble spaces. $\mathcal{E} \subseteq \mathcal{F}_2$ and $\sigma(\mathcal{E}) = \mathcal{F}_2$

Then $X: \Omega_1 \to \Omega_2$ is m'ble iff $\forall A \in \mathcal{E}$, $X^{-1}(A) \in \mathcal{F}_1$

pf. Let $\mathcal{G} := \{B \in \Omega_2, X^{-1}(B) \in \mathcal{F}_1\}$

Then $\mathcal{G}$ is a σ-algebra (ex.)

By condition $\mathcal{E} \subseteq \mathcal{G}$, Therefore $\underset{\mathcal{F}_2}{\underset{\parallel}{\sigma(\mathcal{E})}} \subseteq \mathcal{G} \Rightarrow X$ is m'ble

#

Cor. Let $(\Omega, \mathcal{F})$ be a m'ble space, the following are equivalent:

ⓐ $X: \Omega \to \mathbb{R}$ is a Borel function

ⓑ $\{X < a\} = \{\omega \in \Omega, x(\omega) < a\} \in \mathcal{F}$, $\forall a \in \mathbb{R}$

ⓑ' $\{X \geq a\} \in \mathcal{F}$, $\forall a \in \mathbb{R}$ $\nwarrow$ $\forall a \in \mathbb{R}$, $X^{-1}((-\infty, a)) \in \mathcal{F}$

ⓒ $\{X \leq a\} \in \mathcal{F}$, $\forall a \in \mathbb{R}$

ⓒ' $\{X > a\} \in \mathcal{F}$, $\forall a \in \mathbb{R}$

Sketch of the pf.

By lemma, it suffices to show:

$\mathcal{E}_2 = \{(-\infty, a) : a \in \mathbb{R}\}$

satisfies $\sigma(\mathcal{E}_2) = \mathcal{B}(\mathbb{R})$

We know $\mathcal{B}(\mathbb{R}) \triangleq \sigma(\mathcal{E}_0)$

where $\mathcal{E}_0 = \{(-a, b], -\infty \leq a < b < +\infty\}$

Note: $\mathcal{E}_0 \subseteq \sigma(\mathcal{E}_2)$ because $(-\infty, b] = \bigcap_{n \geq 1} (-\infty, b + \frac{1}{n}) \in \sigma(\mathcal{E}_2)$

$\Leftrightarrow \sigma(\mathcal{E}_0) \subseteq \sigma(\mathcal{E}_2)$

$(a, b] = (-\infty, b] \setminus (-\infty, a] \in \sigma(\mathcal{E}_2)$

$\Leftrightarrow \sigma(\mathcal{E}_2) \subseteq \sigma(\mathcal{E}_0)$

$\mathcal{E}_2 \subseteq \sigma(\mathcal{E}_0)$ because $(-\infty, a) = \bigcup_{n \geq 1} (-\infty, a - \frac{1}{n}] \in \sigma(\mathcal{E}_0)$

$\Rightarrow \sigma(\mathcal{E}_2) = \sigma(\mathcal{E}_0) = \mathcal{B}(\mathbb{R})$

#

Lemma: A distribution function F satisfies

① $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$

② F is increasing: $F(x) \leq F(y)$ if $x < y$

③ F is right continuous: $F(x+h) \downarrow F(x)$ as $h \downarrow 0$

pf. ② $\{X \leq x\} \subseteq \{X \leq y\}$,

$F(x) \leq F(y)$ follows from the monotonicity of $\mathbb{P}$

① First show: $\lim_{n \to -\infty} F(n) = 0$ for $n \in \mathbb{Z}$

$F(n) = \mathbb{P}(\{X \leq n\})$, by continuity from above

$\lim_{n \to -\infty} F(n) = \mathbb{P}\left(\bigcap_{n \in \mathbb{Z}} \{X \leq n\}\right) = \mathbb{P}(\emptyset) = 0$.

by monotonicity of F: $\lim_{x \to -\infty} F(x) = 0$

by continuity from below,

$\lim_{n \to +\infty} F(n) = \mathbb{P}\left(\bigcup_{n \in \mathbb{Z}} \{X \leq n\}\right) = \mathbb{P}(\Omega) = 1$.

by monotonicity of F: $\lim_{x \to +\infty} F(x) = 1$

③ First show $\lim_{n \to +\infty} F(x + \frac{1}{n}) = F(x)$ for $n \in \mathbb{Z}$

$\text{LHS} = \lim_{n \to +\infty} \mathbb{P}(\{X \leq x + \frac{1}{n}\}) \underset{\uparrow}{=} \mathbb{P}\left(\bigcap_{n \geq 1} \{X \leq x + \frac{1}{n}\}\right) = \mathbb{P}(X \leq x) = \text{RHS}$

Cont. from above

Then by monotonicity of $\mathbb{P}$:

$F(x+h) \searrow F(x)$ as $h \downarrow 0$

#

2/7 HTOP (Reci 3)

Wednesday, February 7, 2024 11:17 AM

Multiple Conditioning

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1 | E_2)\,\mathbb{P}(E_2)$$

$$\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_n) = \mathbb{P}(E_1)\,\mathbb{P}(E_2|E_1)\,\mathbb{P}(E_3|E_1 \cap E_2) \cdots \mathbb{P}(E_n | E_1 \cap \cdots \cap E_{n-1})$$

eg. Independent coin flips with p Head
with $1-p$ Tail

Previously,
$\mathbb{P}[n^{th}$ success occur before m^{th} failure$] = ?$

$$\underbrace{\mathbb{P}[\,n \text{ consecutive } \underset{H}{\text{success}} \text{ before } m \text{ consecutive } \underset{T}{\text{failure}}\,]}_{E} = ?$$

Sol. Conditioning on 1^{st} step : $\mathbb{P}(E) = p \cdot \mathbb{P}(E|H) + (1-p)\,\mathbb{P}(E|T)$

Let $F = \{\overset{2nd}{H} H \cdots \overset{n^{th}}{H}\}$, $\mathbb{P}(E|H) = \underbrace{\mathbb{P}(F)}_{p^{n-1}} \cdot \underbrace{\mathbb{P}(E|H\cap F)}_{1} + \underbrace{\mathbb{P}(F^c)}_{1-p^{n-1}} \cdot \underbrace{\mathbb{P}(E|H\cap F^c)}_{\overset{\text{(ex.)}}{=}\mathbb{P}(E|T)}$

$$\therefore \mathbb{P}(E|H) = p^{n-1} + (1-p^{n-1})\,\mathbb{P}(E|T)$$

let $G = \{\overset{2nd}{T} \cdots \overset{n^{th}}{T}\}$, $\mathbb{P}(E|T) = \mathbb{P}(G)\cdot \mathbb{P}(E|T\cap G) + \mathbb{P}(G^c)\cdot \underbrace{\mathbb{P}(E|T\cap G^c)}_{\mathbb{P}(E|H)}$

Answer: $\mathbb{P}(E) = \dfrac{p^{n-1}\cdot(1-q^m)}{p^{n-1}+q^{m-1}-p^{n-1}q^{m-1}}$, $q = 1-p$.

Recall: $X: \Omega \to \mathbb{R}$ $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$

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

$(\Omega, \mathcal{F}, \mathbb{P})$

$X: \Omega \to \mathbb{R}$ is a Borel function if $\forall B \in \mathcal{B}(\mathbb{R})$, $X^{-1}(B) \in \mathcal{F}$
(Real-valued R.V.)

Equivalent Condition: X is a Borel function iff $\{X \le a\} \in \mathcal{F}$, $\forall a \in \mathbb{R}$

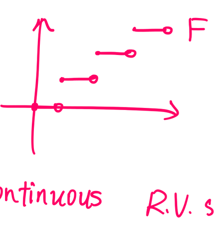
or $\{X < a\} \in \mathcal{F}$, $\forall a \in \mathbb{R}$

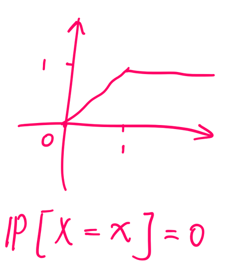
Distribution Function $F_X(x) = \mathbb{P}(X \le x)$ is increasing, right continuous,

$F(-\infty) = 0$, $F(+\infty) = 1$

lemma: ① $\mathbb{P}(x < X \le y) = F(y) - F(x)$

② $\mathbb{P}(X = x) = F(x) - \lim_{y \uparrow x} F(y)$ → Discrete / (absolutely) continuous R.V.s





$\mathbb{P}[X = x] = 0$

"density funct. at x"

pf. of ②: Let $B_n = \{x - \frac{1}{n} < X \le x\}$

by ①: $\mathbb{P}(B_n) = F(x) - F(x - \frac{1}{n})$

send $n \nearrow +\infty$, $\lim_{n \to +\infty} B_n = \bigcap_{n \ge 1} B_n = \{x\}$. by continuity of measure

$$\mathbb{P}(X = x) = \lim_{n \to +\infty} \mathbb{P}(B_n) = F(x) - \lim_{n \to +\infty} F(x - \tfrac{1}{n}) \underset{\substack{\uparrow \\ F \text{ is increasing}}}{=\!=} F(x) - \lim_{y \uparrow x} F(y)$$

#

Properties of Borel functions.

Let $\{X < Y\} = \{w \in \Omega, X(w) < Y(w)\}$

$\{X > Y\} = \{w \in \Omega, X(w) > Y(w)\}$

lemma. Let $(\Omega, \mathcal{F})$ be a measurable space, X, Y are Borel functions

ⓐ $\{X < Y\}, \{X > Y\}, \{X = Y\}, \{X \ne Y\} \in \mathcal{F}$

$\{X \le Y\}, \{X \ge Y\} \in \mathcal{F}$

ⓑ $X + Y$, $X - Y$, XY are Borel functions

pf. ⓐ Use $\mathbb{Q}$ is dense in $\mathbb{R}$

$$\{X < Y\} = \bigcup_{q \in \mathbb{Q}} (\underset{\substack{\cap \\ \mathcal{F}}}{\{X < q\}} \cap \underset{\substack{\cap \\ \mathcal{F}}}{\{q < Y\}}) \in \mathcal{F}$$

$\{X \ge Y\} = \{X < Y\}^c \in \mathcal{F}$

$\{X = Y\} = \{X \le Y\} \cap \{X \ge Y\} \in \mathcal{F}$

ⓑ We show $\{X + Y \le a\} \in \mathcal{F}$, $\forall a \in \mathbb{R}$

Easy to show: if Y is Borel, then $a \pm Y$ is Borel (ex.)

$\{X + Y \le a\} = \{X \le \underbrace{a - Y}_{\text{Borel}}\} \in \mathcal{F}$ by part ⓐ

Note: if X is Borel, then X^2 is Borel: $\{X^2 > a\} = \underset{\substack{\cap \\ \mathcal{F}}}{\{X < -\sqrt{a}\}} \cup \underset{\substack{\cap \\ \mathcal{F}}}{\{X > \sqrt{a}\}} \in \mathcal{F}$, $\forall a \in \mathbb{R}$

Then $XY = \frac{1}{4}\underbrace{(X+Y)^2}_{\text{Borel}} - \frac{1}{4}\underbrace{(X-Y)^2}_{\text{Borel}}$ is Borel

#

Lemma: Let $(X_n)_{n \ge 1}$ is a seq. of Borel function

Then $\sup_{n \ge 1} X_n$, $\inf_{n \ge 1} X_n$, $\limsup_{n \to +\infty} X_n$, $\liminf_{n \to +\infty} X_n$ are Borel.

pf. We claim $\{\sup_{n \ge 1} X_n \le a\} \in \mathcal{F}$, $\forall a \in \mathbb{R}$

$\{\sup_{n \ge 1} X_n\} = \bigcap_{n \ge 1} \underset{\substack{\cap \\ \mathcal{F}}}{\{X_n < a\}} \in \mathcal{F}$

$\inf_{n \ge 1} X_n = -\sup_{n \ge 1} (-X_n)$ is Borel

$\limsup_{n \to +\infty} = \inf_{n \ge 1} \sup_{m \ge n} X_m \in \mathcal{F}$

$\liminf_{n \to +\infty} = \sup_{n \ge 1} \inf_{m \ge n} X_m \in \mathcal{F}$

#

Cor. For seq. of Borel funct. X_n,

if $\lim_{n \to +\infty} X_n$ exists, then $\lim_{n \to +\infty} X_n$ is Borel.

lem. Let $(\Omega_1, \mathcal{F}_1)$, $(\Omega_2, \mathcal{F}_2)$, $(\Omega_3, \mathcal{F}_3)$ be m'ble spaces. $X: \Omega_1 \to \Omega_2$, $Y: \Omega_2 \to \Omega_3$ are m'ble functions,

then $Y \circ X: \Omega_1 \to \Omega_3$ is measurable.

pf. $\forall B \in \mathcal{F}_3$, $(Y \circ X)^{-1}(B) = X^{-1} \underbrace{Y^{-1}(B)}_{\substack{\cap \\ \mathcal{F}_2}} \in \mathcal{F}_1$ #

Def. If $X: \Omega \to \mathbb{R}$ is a R.V.

Then $\sigma(X) \triangleq \{X^{-1}(B), B \in \mathcal{B}(\mathbb{R})\}$ is called the $\boxed{\sigma\text{-algebra generated by } X}$

If $(X_i)_{i \in I}$ is a family of R.V. s, then

$\sigma(X_i, i \in I) \triangleq \sigma(\bigcup_{i \in I} \sigma(X_i))$, $\boxed{\sigma\text{-algebra generated by } (X_i)_{i \in I}}$

Rmk. $\sigma(X)$ is the smallest σ-algebra s.t. X is m'ble, "information of X"

eg. $(\Omega, \mathcal{F}, \mathbb{P})$, $A_1, A_2, \cdots, A_n \in \mathcal{F}$, $A_i \cap A_j = \emptyset$

$X = b_1 \mathbb{1}_{A_1} + b_2 \mathbb{1}_{A_2} + \cdots + b_n \mathbb{1}_{A_n}$, all (b_i) are distinct

$\sigma(X) = \sigma(\{A_1, \cdots, A_n\})$

pf. $\supseteq$: $\forall i$, $A_i = X^{-1}(\{b_i\}) \in \sigma(X) \Rightarrow \sigma(\{A_1, \cdots, A_n\}) \subseteq \sigma(X)$

$\subseteq$: lemma: $\sigma(X) = \sigma(\{X \le a\}, a \in \mathbb{R})$ (ex.)

$\forall a \in \mathbb{R}$, $\{X \le a\} =$ "disjoint union of $(A_i)_{i=1}^n$, $(\bigcup_{i=1}^n A_i)^c$" $\in \sigma(\{A_1, \cdots, A_n\})$

$\Rightarrow \sigma(X) = \sigma(\{X \le a\}, a \in \mathbb{R}) \subseteq \sigma(\{A_1, \cdots, A_n\})$

#

Product Measure:

Two experiments, probability spaces $(\Omega, \mathcal{F}_1, \mathbb{P}_1)$, $(\Omega, \mathcal{F}_2, \mathbb{P}_2)$

(X_1, X_2) takes values in $\Omega \times \Omega = \{(w_1, w_2), w_1, w_2 \in \Omega\}$

We want all sets of the form $\{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$ to be in the product σ-algebra

Define $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\})$

Let $\mathbb{P}_1 \otimes \mathbb{P}_2 : \mathcal{F}_1 \otimes \mathcal{F}_2 \to [0,1]$

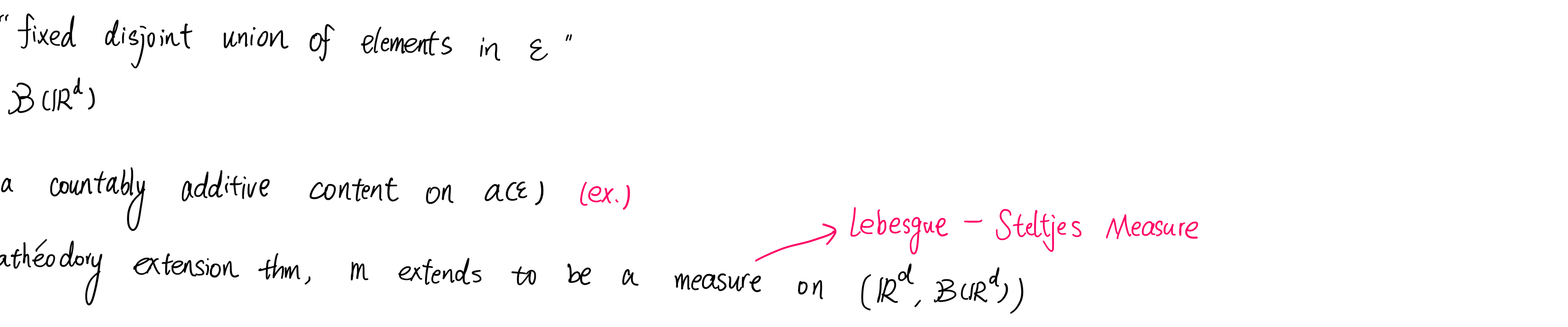
given by $\mathbb{P}_1 \otimes \mathbb{P}_2 (A_1 \times A_2) = \mathbb{P}_1(A_1) \mathbb{P}_2(A_2)$ for $A_1 \in \mathcal{F}_1$, $A_2 \in \mathcal{F}_2$

By Carathéodory extension thm., this extends to a measure on $\sigma(\{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}) = \mathcal{F}_1 \otimes \mathcal{F}_2$

Lebesgue measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$

$\mathcal{E} =$ "left open, right closed cubes" $= \{\text{cube}(a, b), a, b \in \mathbb{R}^d, -\infty \le a_i \le b_i \le +\infty, \forall i\}$

e.g. $d = 2$: elements in $\mathcal{E}$:



Define: $m(\text{cube}(a, b)) = \prod_{i=1}^{d} (b_i - a_i)$

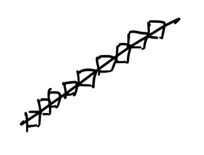
$a(\mathcal{E}) =$ "fixed disjoint union of elements in $\mathcal{E}$"

$\sigma(\mathcal{E}) = \mathcal{B}(\mathbb{R}^d)$

m is a countably additive content on $a(\mathcal{E})$ (ex.)

By Carathéodory extension thm, m extends to be a measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ → Lebesgue – Steltjes Measure

e.g. Line segment in $\mathbb{R}^d$, $d \ge 2$.



$A = \{x : x = \underset{\substack{\cap \\ \mathbb{R}^d}}{x_0} + r \cdot \underset{\substack{\cap \\ \mathbb{R}^d}}{a}, r \in [0, 1]\}$

$m(A) = \lim_{n \to \infty} m(A_n) = 0$

Rmk. If A has dimension $< d$,

then $m_d(A) = 0$.

2/21 HTOP 8

Wednesday, February 21, 2024 11:16 AM

Recall: $F_X(x) = \mathbb{P}(X \leq x)$

F_X is $\nearrow$, right continuous, $\mathbb{P}(X=x) = F_X(x) - \lim_{y \nearrow x} F_X(y)$

Two special classes of R.V.: Discrete & (absolutely) Continuous

Def. $\boxed{\text{A R.V. } X \text{ is discrete}}$ if it takes value in a countable set $\{x_1, x_2, \cdots\}$

Probability mass function : $f(x) \triangleq \mathbb{P}(X=x)$

We say that $\{x_1, x_2, \cdots\}$ are atoms of F_X

Def. $\boxed{\text{A R.V. } X \text{ is (absolutely) continuous}}$ if

$F_X(x) = \int_{-\infty}^{x} f(u)\,du$ for some integrable $f: \mathbb{R} \to [0, +\infty)$. f is the probability density function

$f: I \to \mathbb{R}$ is absolutely continuous if $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t. for every $(a_k, b_k) \subseteq I$,

$\sum |b_k - a_k| < \delta$, we have $\sum |f(b_k) - f(a_k)| < \varepsilon$

Rmk. F_X is absolutely continuous

Rmk. There exists R.V. s.t. F_X is cont. but not absolutely cont.

eg. singular R.V. : F_X = Cantor Function

(F_X has all mass concentrated on an uncountable set of measure 0)

Discrete R.V.

$F_X(x) = \sum_{x_i: x_i < x} f(x_i)$, $f(x) = F_X(x) - \lim_{y \nearrow x} F_X(x)$

Expectation. [discrete]

Def. $\boxed{\text{The mean / expectation / expected value}}$ of a R.V. X with a probability mass funct. f

is $\mathbb{E}X = \sum_{x: f(x)>0} x f(x) = \sum_{x: f(x)>0} x\, \mathbb{P}(X=x)$ whenever the sum is absolutely convergent

eg. 2 Coin Flips : X = # heads

$\mathbb{E}X = \sum x \cdot \mathbb{P}(X=x) = 0 \cdot \mathbb{P}(X=0) + 1 \cdot \mathbb{P}(X=1) + 2 \cdot \mathbb{P}(X=2) = 0 + \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$

lem 1: (Change of Variable) If X is a R.V. with prob. mass funct. f,

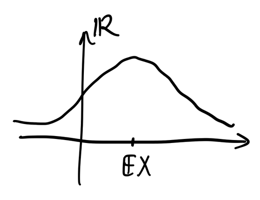
$g: \mathbb{R} \to \mathbb{R}$, Then $\boxed{\mathbb{E}g(X) = \sum_{x: f(x)>0} g(x) \cdot f(x)}$ (ex.)

lem 2: Let X be a R.V. taking values in $\mathbb{N}$

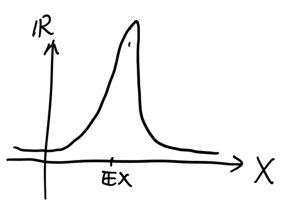
Then $\boxed{\mathbb{E}X = \sum_{n \in \mathbb{N}} \mathbb{P}(X \geq n)}$ (ex.)

Def. $k \in \mathbb{N}$. The $\boxed{\text{k-th moment}}$ of X, $m_k = \mathbb{E}(X^k)$

$\boxed{\text{k-th central moment}}$ of X, $\sigma_k = \mathbb{E}\left[(X - \mathbb{E}X)^k\right]$ "deviate from the mean"



σ_k large



σ_k small : there are more concentrated at the center

Def. When $k=2$, $\operatorname{Var} X = \mathbb{E}\left[(X - \mathbb{E}X)^2\right]$, $\boxed{\text{the variance of } X}$

$\sigma = \sqrt{\operatorname{Var} X}$, $\boxed{\text{the standard deviation}}$

$\operatorname{Var} X = \mathbb{E}\left[X^2 - 2X \cdot \mathbb{E}X + (\mathbb{E}X)^2\right] = \mathbb{E}X^2 - 2\mathbb{E}\left[X \cdot \mathbb{E}X\right] + (\mathbb{E}X)^2$

$\Leftrightarrow \boxed{\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2}$

eg. $\boxed{\text{Bernoulli}(p)}$, $\mathbb{P}(X=1) = p$, $\mathbb{P}(X=0) = q = 1-p$

$\mathbb{E}X = 1 \cdot \mathbb{P}(X=1) = p$

$\operatorname{Var} X = \underbrace{\mathbb{E}X^2}_{\mathbb{E}X} - (\mathbb{E}X)^2 = p - p^2 = pq$.

eg. $\boxed{\text{Binomial}(n,p)}$ = # success in n independent trials $\hat{=}$ X

Sol 1:

$f(k) = \mathbb{P}(X=k) = C_n^k \cdot p^k \cdot q^{n-k}$, $q = 1-p$

$\mathbb{E}X = \sum_{k=1}^{n} k \cdot C_n^k \cdot p^k \cdot q^{1-k}$

Binomial Expansion: $(1+x)^n = \sum_{k=0}^{n} C_n^k x^k$

$\frac{d}{dx}$: $n(1+x)^{n-1} x = \sum_{k=1}^{n} C_n^k\, k \cdot x^k$

take $x = \frac{p}{q}$ $\Rightarrow$ $n\left(1+\frac{p}{q}\right)^{n-1} \cdot \frac{p}{q} = \sum_{k=1}^{n} k \cdot C_n^k\, p^k\, q^{-k}$

$\overset{\times q^n}{\Rightarrow}$ $np = \sum_{k=1}^{n} k C_n^k p^k q^{n-k} = \mathbb{E}X$

Similarly, $\operatorname{Var} X = npq$ (ex.)

Sol 2:

$X = Y_1 + Y_2 + \cdots + Y_n$

Y_i, i.i.d., $Y_i \sim \text{Bernoulli}(p)$

Then $\mathbb{E}X = \mathbb{E}(Y_1 + Y_2 + \cdots + Y_n) = \sum_{i=1}^{n} \mathbb{E}Y_i = np$

$\operatorname{Var} X = \mathbb{E}\left[(X - \mathbb{E}X)^2\right]$

$= \mathbb{E}\left(\sum_{i=1}^{n} (Y_i - \mathbb{E}Y_i)\right)^2$

$= \sum_{i=1}^{n} \mathbb{E}(Y_i - \mathbb{E}Y_i)^2 + \sum_{i \neq j} \underbrace{2\mathbb{E}\left[(Y_i - \mathbb{E}Y_i)(Y_j - \mathbb{E}Y_j)\right]}_{0 \text{ (next lemma)}}$

$= \sum_{i=1}^{n} \operatorname{Var} Y_i = npq$

$= \mathbb{E}\left[Y_i Y_j + \mathbb{E}Y_i \cdot \mathbb{E}Y_j - \mathbb{E}Y_i\, Y_j - \mathbb{E}Y_j\, \mathbb{E}Y_i\right]$

$= 2\mathbb{E}(Y_i)\mathbb{E}(Y_j) - 2\mathbb{E}(Y_i)\mathbb{E}(Y_j) = 0$

lem. ① $\forall a, b \in \mathbb{R}$, $\mathbb{E}\left[aX + bY\right] = a\mathbb{E}X + b\mathbb{E}Y$

② If X, Y are independent, then $\mathbb{E}[XY] = \mathbb{E}X \cdot \mathbb{E}Y$

③ If X, Y are independent, then $\operatorname{Var}(X+Y) = \operatorname{Var} X + \operatorname{Var} Y$

pf. ② $\mathbb{E}XY = \sum x_i y_j\, \mathbb{P}(X = x_i, Y = y_j) \overset{\text{indep.}}{=} \sum x_i\, \mathbb{P}(X = x_i) \cdot \sum y_j\, \mathbb{P}(Y = y_j)$

$= \mathbb{E}X \cdot \mathbb{E}Y$

Rmk 1: Since X, Y are independent

iff $\mathbb{P}(X=i, Y=j) = \mathbb{P}(X=i)\mathbb{P}(Y=j)$

③ $\operatorname{Var}(X+Y) = \mathbb{E}(X - \mathbb{E}X)^2 + \mathbb{E}(Y - \mathbb{E}Y)^2 + 2\underbrace{\mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y)}_{0}$

$= \operatorname{Var}(X) + \operatorname{Var}(Y)$

#

Rmk 2: X, Y independent

iff $\sigma(X)$, $\sigma(Y)$ independent

iff $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

for $\forall A \in \sigma(X)$ & $B \in \sigma(Y)$

Def. $\boxed{\text{Covariance, } \operatorname{Cov}(X,Y) = \mathbb{E}\left[(X - \mathbb{E}X)(Y - \mathbb{E}Y)\right]}$

Def. We say X, Y are $\boxed{\text{uncorrelated}}$ iff $\operatorname{Cov}(X,Y) = 0$

Rmk. independence $\underset{\not\Leftarrow}{\Rightarrow}$ uncorrelated

Counter-example: X is a symmetric R.V. $\mathbb{P}(X=x) = \mathbb{P}(X=-x)$

$Y = |X|$

then $\operatorname{Cov}(X,Y) = \mathbb{E}\left[XY + \mathbb{E}X \cdot \mathbb{E}Y - \mathbb{E}X \cdot Y - \mathbb{E}Y \cdot X\right]$

$= \mathbb{E}[XY] - \mathbb{E}X \cdot \mathbb{E}Y$

$= \sum x_i y_j\, \mathbb{P}(X = x_i, Y = y_j) - \underbrace{\sum x_i\, \mathbb{P}(X = x_i)}_{0} \sum y_j\, \mathbb{P}(Y = y_j)$

$= \sum_i x_i \sum_j y_j\, \mathbb{P}(X = x_i, Y = y_j)$

$= 0$

But Y depends on X.

Def.

eg. A R.V. X taking value in $\mathbb{N}$ is a $\boxed{\text{Poisson R.V.}}$ with parameter $\lambda > 0$

if $f(i) = \mathbb{P}(X = i) = \frac{\lambda^i}{i!} e^{-\lambda}$, $i \in \mathbb{N}$

$\mathbb{E}X = \lambda$ (ex.)

$\operatorname{Var} X = \lambda$ (ex.)

• Poisson arises from Binomial (n,p) in the limit $n \to +\infty$ but $np = \lambda$

Binomial $(n, \frac{\lambda}{n})$: $\mathbb{P}(X = i) = C_n^i p^i (1-p)^{n-i}$

$= C_n^i \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}$

$= \frac{n(n-1)\cdots(n-i+1)}{i!} \cdot \frac{\lambda^i}{n^i} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-i}$

$\xrightarrow[i, \lambda \text{ fixed}]{n \to \infty} \frac{\lambda^i}{i!} e^{-\lambda}$

Hence in $n \to \infty$, the Binomial R.V. converges to a Poisson R.V.

2/23 HTOP (Reci 4)

Friday, February 23, 2024 9:42 AM

Rmk. for the HW2 PDF problem 1:

$0 \le a \le b \le c \le d \le 1$

$a + d \ge b + c$

$$\begin{cases} m(\liminf_n A_n) = a \\ \liminf_n m(A_n) = b \\ \limsup_n m(A_n) = c \\ m(\limsup_n A_n) = d \end{cases}$$

$\liminf_n A_n = \bigcup_{n\ge1}\bigcap_{k\ge n} A_k = \{A_n, \text{e.v.}\}$ → A_n happens eventually

$\limsup_n A_n = \bigcap_{n\ge1}\bigcup_{k\ge n} A_k = \{A_n, \text{i.o.}\}$ → A_n happens infinitely open

eg. ① Show that if X is a R.V. taking value in $\mathbb{N}$

Then $\boxed{\mathbb{E}X = \sum_{n=1}^{+\infty} \mathbb{P}(X \ge n)}$

② Assume offer (X_i), i.i.d.

T = first time see an offer $> X_1$, Compute $\mathbb{E}T$

Sol. ① $\mathbb{E}X = \sum_{n\in\mathbb{N}} n\,\mathbb{P}(X=n) = \sum_{n=1}^{+\infty} n\,\mathbb{P}(X=n)$

$$= \underbrace{\sum_{n=1}^{+\infty}\sum_{k=n}^{\infty} \mathbb{P}(X=k)}_{\sum_{k=1}^{\infty}\mathbb{P}(X=k) + \sum_{k=2}^{\infty}\mathbb{P}(X=k) + \sum_{k=3}^{\infty}\mathbb{P}(X=k) + \cdots} = \sum_{n=1}^{+\infty} \mathbb{P}(X \ge n)$$

② $X_1, X_2, X_3, \dots$

T takes value in $\mathbb{N}$

$\mathbb{E}T = \sum_{n\ge1} \mathbb{P}(T \ge n)$ (use ①)

where $\mathbb{P}(T > n) = \mathbb{P}(X_1 = \max\{X_1, \cdots, X_n\})$

$= \frac{1}{n}$

$\therefore \mathbb{E}T = \sum_{n\ge1} \frac{1}{n} \to +\infty$

Rmk. As time goes by, each time you see an offer $> X_i$'s probability is $\frac{1}{n}$

but it does not decay fast enough, thus $\mathbb{E}T \to \infty$.

Hence it is not suggested to turn down the first offer

#

eg. ① $(\Omega_2, \mathcal{F}_2)$ is a m'ble space. $f: \Omega_1 \to \Omega_2$

- Show that $\mathcal{F} = \{f^{-1}(A) : A \in \mathcal{F}_2\}$ is a σ-algebra on Ω_1
- Show that if $\mathcal{G}$ is a σ-algebra on Ω_1 s.t. $\underbrace{f \text{ is m'ble}}_{f^{-1}(\mathcal{F}_2)\in\mathcal{G}}$, then $\mathcal{G} \supseteq \mathcal{F}$

$f: \mathcal{G} \to \mathcal{F}_2$

② $(\Omega_1, \mathcal{F}_1)$ is a m'ble space. $f: \Omega_1 \to \Omega_2$

- Show that $\tilde{\mathcal{F}} = \{A \subseteq \Omega_2 : f^{-1}(A) \in \mathcal{F}_1\}$ is a σ-algebra
- If $\mathcal{G}$ is a σ-algebra on Ω_2 s.t. $\underbrace{f \text{ is m'ble}}_{f^{-1}(\mathcal{G})\in\mathcal{F}_1}$, then $\mathcal{G} \subseteq \tilde{\mathcal{F}}$

$f: \mathcal{F}_1 \to \mathcal{G}$

Pf. ① Let $B \in \mathcal{F}$

then $\exists A \in \mathcal{F}_2$ s.t. $f^{-1}(A) = B$

Since f is $\mathcal{G}/\mathcal{F}_2$ m'ble,

then $\underset{\underset{B}{\shortparallel}}{f^{-1}(A)} \in \mathcal{G}$

$\therefore \mathcal{F} \subseteq \mathcal{G}$

② Let $B \in \mathcal{G}$

Since f is $\mathcal{F}_1/\mathcal{G}$ m'ble,

then $f^{-1}(B) \in \mathcal{F}_1$

then $B \in \tilde{\mathcal{F}}$

$\therefore \mathcal{G} \subseteq \tilde{\mathcal{F}}$

#

Recall: $X \sim \text{Poisson}(\lambda)$ iff $\mathbb{P}(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k \in \mathbb{N}$

eg. Show that if $X \sim \text{Poisson}(\lambda_1)$, $Y \sim \text{Poisson}(\lambda_2)$ ①, X, Y are independent, ②

then $\boxed{X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)}$ ③

① + ② ⇒ ③

① + ③ ⇒ ②

Sol. $\mathbb{P}(X=k) = \frac{\lambda_1^k}{k!}e^{-\lambda_1}$

$\mathbb{P}(Y=k) = \frac{\lambda_2^k}{k!}e^{-\lambda_2}$

$$\mathbb{P}(X+Y=k) \overset{\text{indep.}}{=} \sum_{k_1+k_2=k} \frac{\lambda_1^{k_1}}{k_1!}e^{-\lambda_1} \cdot \frac{\lambda_2^{k_2}}{k_2!}e^{-\lambda_2}$$

$$= \sum_{k_1=0}^{k} \frac{\lambda_1^{k_1}\cdot\lambda_2^{k-k_1}}{k_1!\cdot(k-k_1)!} e^{-(\lambda_1+\lambda_2)} \qquad = \frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)}$$

$$= \frac{\sum_{k_1=0}^{k} \lambda_1^{k_1}\lambda_2^{k-k_1}\cdot C_k^{k_1}}{k!} e^{-(\lambda_1+\lambda_2)}$$

$$= \sum_{k_1=0}^{k} \frac{\lambda_1^{k_1}\lambda_2^{k-k_1}\cdot\frac{k!}{k_1!\cdot(k-k_1)!}}{k!} e^{-(\lambda_1+\lambda_2)}$$

#

eg. ① $\boxed{\text{Geometric}(p)}$ = first success in indep. trial with success prob. p

$p \in (0,1)$

$\mathbb{P}(X=k) = (1-p)^{k-1}\cdot p$, $k \in \mathbb{N}^+$

$\mathbb{E}X = \frac{1}{p}$, $\text{Var}\,X = \frac{1-p}{p^2}$

② Coupon collector. N types X = first time to get a complete set $\{1, \cdots, N\}$

*blind box

$\mathbb{E}X = ?$

②.

Sol. Y_k = first time to get k distinct coupons

$X = Y_N = \sum_{k=1}^{N}(Y_k - Y_{k-1}) + Y_1$

$Y_k - Y_{k-1} \sim \text{Geometric}\left(\frac{N-(k-1)}{N}\right)$

$\therefore \mathbb{E}X = \sum_{k=1}^{N} \mathbb{E}(Y_k - Y_{k-1}) + 1$

$= \sum_{k=1}^{N} \frac{N}{N-(k-1)} + 1$

$\sim N \log N$

#

Discrete R.V.

- Prob. mass function
- Expectation
- Change of Variable
- K-th moment

K-th central moment

(absolutely) continuous R.V.

- Prob. density function

$\boxed{\mathbb{E}X = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-\infty}^{\infty} x \, dF_X(x)}$

$\boxed{\mathbb{E}g(X) = \int g(x) f(x)\,dx}$ (ex.)

$\mathbb{E}X^k = \int x^k f(x)\,dx$

Recall: If X takes value in $\mathbb{N}$, then $\mathbb{E}X = \sum_{n\geq 1} \mathbb{P}(X \geq n)$

Lemma: If X is a non-negative abs. cont. R.V. with density function f, Then $\boxed{\mathbb{E}X = \int_0^{+\infty} \mathbb{P}(X > x)\,dx = \int_0^{+\infty} (1 - F_X(x))\,dx}$ [Continuous]

pf. $\int_0^{+\infty} \mathbb{P}(X \geq x)\,dx = \int_0^{+\infty}\int_x^{+\infty} f(y)\,dy\,dx \overset{\text{Fubini}}{=} \int_0^{+\infty} \left(\int_0^y dx\right) f(y)\,dy$

$= \int_0^{+\infty} y f(y)\,dy$

$= \mathbb{E}X.$

#

Lebesgue Integral & Expectation

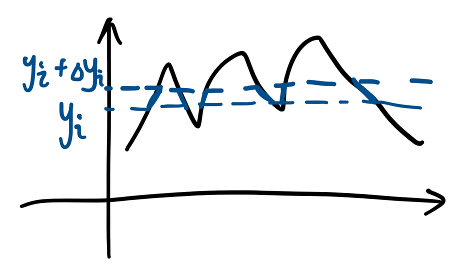
Recall: Riemann Integral

$\Downarrow$

works for f having finitely many discontinuity

but for $f(x) = 1_{\mathbb{Q}}$ in probability, Riemann Integral doesn't work

Lebesgue:



$\{x : y_i \leq f(x) \leq y_i + \Delta y_i\}$

$= f^{-1}[y_i, y_i + \Delta y_i]$

Goal: $\sum \Delta y_i \, m(f^{-1}[y_i, y_i + \Delta y_i]) \longrightarrow \int f(x)\,dx$

Today's Goal: Given any m'ble space $(\Omega, \mathcal{F}, \mu)$, Borel function f.

Define $\underline{\int_\Omega f\,d\mu}$

Construction: $\int_\Omega 1_A\,d\mu \Rightarrow$ simple function $\Rightarrow$ non-negative function $\Rightarrow$ general Borel function

Step 1: Let $A \in \mathcal{F}$, $1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$ Define: $\int_\Omega 1_A\,d\mu = \mu(A)$

Step 2: Let f be a normal representation of simple function:

$f = \sum_{i=1}^{n} a_i 1_{A_i}$, $A_i \in \mathcal{F}$, $A_i \cap A_j = \emptyset$, $a_i \geq 0$

Define: $\int_\Omega f\,d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$

Fact: If f, g are simple functions, then $f+g$, $f \cdot g$, $\min(f,g)$, $\max(f,g)$ are simple functions. (ex.)

Prop. If f, g are non-negative simple functions, then

1. $\int af + bg\,d\mu = a\int f\,d\mu + b\int g\,d\mu$
2. If $f \leq g$, then $\int f\,d\mu \leq \int g\,d\mu$
3. $w(A) = \int_A f\,d\mu = \int_\Omega 1_A \cdot f\,d\mu$, then

   $w: \mathcal{F} \to [0, +\infty)$ is a measure (ex.)

Step 3. Approximate non-negative Borel function by simple functions

Let $f \geq 0$ be Borel

Then $f = \sup f_i$, $f_i = \sum_{k=1}^{i \cdot 2^i} \frac{k-1}{2^i} 1_{\{\frac{k-1}{2^i} \leq f \leq \frac{k}{2^i}\}} + i \cdot 1_{\{f > i\}}$

$\uparrow$

simple, increasing w.r.t. i

Def. Let $f = \sup_i f_i$, where (f_i) is an increasing sequence of simple functions s.t. $\lim_{i\to\infty} f_i = f$, define:

$\int_\Omega f\,d\mu = \lim_{i\to\infty} \int_\Omega f_i\,d\mu = \sup_i \int_\Omega f_i\,d\mu$

Question: Since f can have multiple simple function representations.

say $f = \sup_i g_i$, $\sup_i \int g_i\,d\mu \overset{?}{=} \sup_i \int f_i\,d\mu$

Theorem: (Monotone Convergence) For every increasing seq. f_n of m'ble functions, $\lim_{n\to\infty} \int_\Omega f_n\,d\mu = \int_\Omega \lim_{n\to\infty} f_n\,d\mu$.

If we assume MCT, then $\sup_i \int g_i\,d\mu = \int \lim_i g_i\,d\mu$

$= \int \lim_i f_i\,d\mu = \sup_i \int f_i\,d\mu$

Step 4: (final step) General Borel function f:

$f^+ = \max\{0, f\}$, $f^- = \max\{0, -f\}$

Def. A Borel function f is said to be (Lebesgue) integrable if $\int_\Omega f^+\,d\mu < +\infty$ and $\int_\Omega f^-\,d\mu < +\infty$

define: $\int_\Omega f\,d\mu = \int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu$

Rmk.

prop. Let f, g be Borel functions. Then:

- $\int af+bg \, d\mu = a\int f \, d\mu + b \int g \, d\mu$
- If $f \le g$, then $\int f \, d\mu \le \int g \, d\mu$

pf. (ex.) Simple functions $\Rightarrow$ nonnegative Borel funct. $\Rightarrow$ general Borel funct.

prop. f is integrable iff $\int |f| \, d\mu < +\infty$

In particular, there exists a Borel function Y s.t. $|f| \le Y$, and Y is integrable, then f is integrable.

pf. • Note that $f = f^+ - f^-$ where $f^+ = \max\{f, 0\}$, $f^- = \max\{0, -f\}$

$|f| = f^+ + f^-$

By definition, f is integrable $\Leftrightarrow \int f^+ \, d\mu < +\infty$, $\int f^- d\mu < +\infty \Leftrightarrow \int |f| \, d\mu < +\infty$

• If $|f| \le Y$, then $\int |f| \, d\mu \le \int Y \, d\mu < +\infty \Rightarrow |f|$ is integrable $\Leftrightarrow f$ is integrable

#

eg. $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$, $f(x) = \sin x$, is f integrable over $\mathbb{R}$?

$\int_{\mathbb{R}} f(x) \, dx$ No, because $\int_{\mathbb{R}} |f(x)| dx = +\infty$

Def. We say $f = g$ almost everywhere (a.e.) if $\{f \ne g\}$ has measure 0.

$X = Y$ almost surely (a.s.) if $\{X \ne Y\}$ has prob. measure 0.

ex. Prove that $\mu(A) = 0$, f is Borel, then $\int_A f \, d\mu = 0$

Simple $\Rightarrow$ nonnegative Borel $\Rightarrow$ general Borel

Cor. If $f = g$ a.e., then $\int f \, d\mu = \int g \, d\mu$

Expectation: $\boxed{\mathbb{E}X \triangleq \int_\Omega X \, d\mathbb{P}}$ (general)

propability density function.

For absolutely cont. R.V.: $\mathbb{E}X \triangleq \int_{-\infty}^{+\infty} x f(x) \, dx$

Rmk. They shall coincide $\Rightarrow$ Goal

§ Absolute Continuity of Measures & Radon - Nikodym Derivative

prop. Let $(\Omega, \mathcal{F}, \mu)$ be a measurable space. $f: \Omega \to [0, +\infty)$ be a Borel function.

Then $\nu(A) = \int_A f \, d\mu$ defines a measure.

Def. If $\forall A \in \mathcal{F}$, $\nu(A) = \int_A f \, d\mu$, then we say f is the $\boxed{\text{Radon - Nikodym derivative}}$ (density) of ν w.r.t. μ. Write $f = \frac{d\nu}{d\mu}$

pf of prop.: • If $A = \emptyset$, $\nu(A) = 0$ (follows from ex.)

• Countable additivity: Let (A_j) be disjoint.

$$\nu\Big(\bigcup_{j=1}^{+\infty} A_j\Big) = \int_{\bigcup_{j=1}^{+\infty} A_j} f \, d\mu = \int_\Omega f \cdot 1_{\bigcup_{j=1}^{+\infty} A_j} \, d\mu$$

$$= \int_\Omega f \cdot \lim_{n\to+\infty} \big(1_{\bigcup_{j=1}^{n} A_j}\big) \, d\mu$$

$$\overset{\text{(MCT)}}{=} \lim_{n\to+\infty} \int_\Omega f \cdot 1_{\bigcup_{j=1}^{n} A_j} \, d\mu$$

$$= \lim_{n\to+\infty} \int_\Omega f \cdot \Big(\sum_{j=1}^{n} 1_{A_j}\Big) \, d\mu$$

$$= \lim_{n\to+\infty} \sum_{j=1}^{n} \nu(A_j) = \sum_{j=1}^{+\infty} \nu(A_j)$$

#

Def. We say ν is absolutely cont. w.r.t. μ iff $\forall A \in \mathcal{F}$, $\mu(A) = 0$, we have $\nu(A) = 0$

Write $\nu \ll \mu$

eg: If $\nu(A) = \int_A f \, d\mu$, f is Borel, then $\nu \ll \mu$

eg: Lebesgue measure - m, $m \ll 2m$, $2m \ll m$.

Def. If $\mu \ll \nu$ and $\nu \ll \mu$, then we say μ and ν are equivalent.

Write $\mu \sim \nu$

eg. $\mu = \mu_{\text{counting}}$, $\mu(A) = \begin{cases} |A| & \text{if } |A| \text{ is finite} \\ +\infty & \text{if } |A| \text{ is infinite} \end{cases}$ cardinality

Radom - Nikodym Thm.

• If μ, ν are σ-finite measures on $(\Omega, \mathcal{F})$ and $\nu \ll \mu$, then there is a m'ble function f s.t. $\nu(A) = \int_A f \, d\mu$, $\forall A \in \mathcal{F}$

μ is σ-finite, if $\Omega = \bigcup_{j=1}^{+\infty} A_j$ where $A_j \in \mathcal{F}$, $\mu(A_j) < +\infty$

pf. (Idea): $G \triangleq \{f : \int_A f \, d\mu \le \nu(A), \forall A \in \mathcal{F}\}$

Note that if $f_1, f_2 \in G$, then $\max\{f_1, f_2\} \in G$ since $\int_A \max\{f_1, f_2\} d\mu = \int_{A \cap \{f_1 \ge f_2\}} f_1 \, d\mu + \int_{A \cap \{f_1 < f_2\}} f_2 \, d\mu$

$\le \nu(A \cap \{f_1 \ge f_2\}) + \nu(A \cap \{f_1 < f_2\})$

$= \nu(A)$

Then take a seq. $\{f_n\} \subseteq G$ s.t. $\lim_{n\to+\infty} \int_\Omega f_n \, d\mu = \sup_{f \in G} \int_\Omega f \, d\mu$

WLOG, $\{f_n\}$ is increasing. $g = \lim_{n\to+\infty} f_n$ exists.

Then prove that $g \in G$ & $\int_A g \, d\mu = \nu(A)$, $\forall A \in \mathcal{F}$

#

Prop. (Equiv. Charact. of absolute continuity)

$\nu \ll \mu \Leftrightarrow \forall \varepsilon > 0, \exists \delta > 0$ s.t. $\forall A$ with $\mu(A) < \delta$, we have $\nu(A) < \varepsilon$ (Rmk. more general)

pf. $\Rightarrow$ Assume on the contrary that $\exists \varepsilon > 0$ s.t. $\forall \delta > 0$ we can find A with

$\mu(A) < \delta$ but $\nu(A) \ge \varepsilon$

$\forall n \in \mathbb{N}$, take A_n s.t. $\mu(A_n) \le \frac{1}{2^n}$ and $\nu(A_n) \ge \varepsilon$

Take $B = \limsup_{n\to+\infty} A_n = \bigcap_{n\ge1} \underbrace{\bigcup_{k\ge n} A_k}_{B_n} = \bigcap_{n\ge1} B_n$

We have $\mu(B_n) \le \sum_{k\ge n} \mu(A_k) \le \sum_{k\ge n} \frac{1}{2^k} = \frac{1}{2^{n-1}}$

$\nu(B_n) \ge \nu(A_n) \ge \varepsilon$

By continuity of measure: $\mu(B) = \lim_{n\to+\infty} \mu(B_n) = 0$ but $\nu(B) = \lim_{n\to+\infty} \nu(B_n) \ge \varepsilon$

which contradicts the definition of $\nu \ll \mu$.

$\Leftarrow$ (ex.)

#

Absolutely cont. R.V. $\Leftrightarrow$ distribution $\mathbb{P}_X$ is absolutely cont. w.r.t. Lebesgue measure

$\mathbb{P}_X((-\infty, a]) = \mathbb{P}(X \le a)$

$\mathbb{P}_X(A) = \mathbb{P}(X \in A)$

• Examples of abs. cont. R.V.

eg. $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$, $f(x) = \lambda e^{-\lambda x} \cdot 1_{[0,+\infty)}(x)$

Then $\mathbb{P}_\lambda(B) = \int_B f(x) \, dx$ defines a prob. measure

$\mathbb{P}_\lambda(\mathbb{R}) = \int_0^{+\infty} \lambda e^{-\lambda x} dx = 1$

ex. Show that if X with density function $\boxed{Exp(\lambda) = \lambda e^{-\lambda x} \ \text{[Exponential]}}$ $(x \ge 0)$

$\mathbb{E}X = \frac{1}{\lambda}$, $Var\, X = \frac{1}{\lambda^2}$.

3/1 HTOP (Reci 5)
Friday, March 1, 2024 11:19 AM

Cauchy - Schwarz Inequality :

Let X, Y be R.V. s , $(\mathbb{E}[XY])^2 \leq \mathbb{E}X^2\, \mathbb{E}Y^2$

pf. Use that $\mathbb{E}(aX-bY)^2 \geq 0$, $\forall a,b \in \mathbb{R}$

$a^2\, \mathbb{E}(X^2) - 2ab \cdot \mathbb{E}(XY) + b^2\, \mathbb{E}(Y^2) \geq 0$

Deterministic : $\frac{1}{4}\Delta = (\mathbb{E}XY)^2 - \mathbb{E}(X^2)\, \mathbb{E}(Y^2) \leq 0$

#

Rmk. " = " iff $X = cY$ for some $c \in \mathbb{R}$

Cor. $|Cov(X,Y)| \leq \sqrt{Var X \cdot Var Y}$

pf. $|Cov(X,Y)| = \left|\mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)]\right|$

$\leq \sqrt{\mathbb{E}[(X-\mathbb{E}X)^2] \cdot \mathbb{E}[(Y-\mathbb{E}Y)^2]}$

$= \sqrt{Var X \cdot Var Y}$

#

ex. Let $X \geq 0$ be a R.V. with $\mathbb{E}X = 1$, $t \in (0,1)$

Show that $\mathbb{P}(X \geq t) \geq \frac{(1-t)^2}{\mathbb{E}X^2}$

Sol. Take $Y = 1_{X \geq t}$.

$\mathbb{P}(X \geq t) \cdot \mathbb{E}(X^2)$

$= \mathbb{E}(Y^2) \cdot \mathbb{E}(X^2)$

$\overset{C-S}{\geq} [\mathbb{E}(XY)]^2 = [\mathbb{E}(X \cdot 1_{X \geq t})]^2$

$= [\underbrace{\mathbb{E}X}_{1} - \underbrace{\mathbb{E}(X \cdot 1_{X<t})}_{\leq \mathbb{E}(t \cdot 1_{X<t}) \leq t}]^2 \quad (*)$

$\geq (1-t)^2$

#

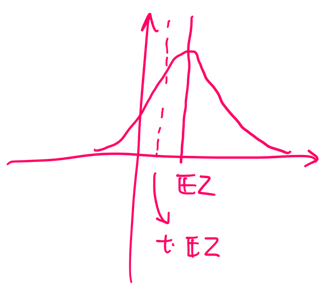
At $(*)$: If we don't have $\mathbb{E}X = 1$ & let $t = t \cdot \mathbb{E}X$,

we can prove the following theorem :

Thm. [Payley - Zygmund Inequality]

If $Z \geq 0$ be a R.V. , $t \in (0,1)$

Then $\mathbb{P}(Z \geq t \cdot \mathbb{E}Z) \geq (1-t)^2 \cdot \frac{(\mathbb{E}Z)^2}{\mathbb{E}(Z^2)}$



Concentrated : $(\mathbb{E}Z)^2 \approx \mathbb{E}(Z^2)$, i.e., $Var Z \approx 0$

Recall : Exp (λ) : $f_\lambda(x) = \lambda e^{-\lambda x} \cdot 1_{\{x \geq 0\}} \iff \mathbb{P}(X > t) = e^{-\lambda t}$

ex. ① If X is the R.V. with $f_\lambda(x) = \lambda e^{-\lambda x} \cdot 1_{\{x \geq 0\}}$,

Show that $\forall s,t > 0$, $\mathbb{P}[X > t+s \mid X > s] = \mathbb{P}[X > t]$

" lack of memory prop. "

② Find all cont. R.V. s s.t. $\mathbb{P}[X > t+s \mid X > s] = \mathbb{P}[X > t]$

Sol. ① $\mathbb{P}(X > t) = \int_t^{+\infty} \lambda e^{-\lambda x} \cdot 1_{\{x \geq 0\}}\, dx = e^{-\lambda t}$

$\mathbb{P}[X > t+s \mid X > s] = \frac{\mathbb{P}[X > t+s,\, X > s]}{\mathbb{P}[X > s]} = \frac{\mathbb{P}[X > t+s]}{\mathbb{P}[X > s]} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(X > t)$

#

② Let $f(t) = \mathbb{P}[X > t]$. Then $\frac{f(t+s)}{f(s)} = f(t)$ for $\forall s,t \in \mathbb{R}$

Let $g = \log f$

Claim : If $g(s+t) = g(s) + g(t)$ for $\forall t,s \in \mathbb{R}$, g is continuous . (further works for Borel) by Lusin's Thm.

Then $g(x) = \lambda x$ for some $\lambda \in \mathbb{R}$

- $g(0) = 0$, $g(n) = n \cdot g(1)$
- $g(q) = q \cdot g(1)$, $\forall q = \frac{m}{n} \in \mathbb{Q}$
- By continuity $\Rightarrow g(x) = \lambda x$ for $\lambda \in \mathbb{R}$

Lusin's thm :

$(\mathbb{R}, \mathcal{F}, \mu)$. Let $f : [a,b] \to \mathbb{R}$ be a m'ble function .

Then $\forall \varepsilon > 0$, there is a compact set $K_\varepsilon \subseteq [a,b]$ s.t. $\mu(K_\varepsilon) > b-a-\varepsilon$

and $f|_{K_\varepsilon}$ is (uniformly) continuous.

then $f(t+s) = f(t) \cdot f(s)$

$\iff g(t+s) = g(t) + g(s)$

$\Rightarrow g(x) = \lambda x$

$\therefore f(x) = e^{\lambda x}$. (f can only be in the form of exponential)

#

ex. $X_1 \sim Exp(\lambda_1)$, $X_2 \sim Exp(\lambda_2)$, indep.

Compute ① : density funct. for $\min\{X_1, X_2\}$ $\quad Exp(\lambda_1 + \lambda_2)$

② : $\mathbb{P}(X_1 < X_2)$ $\quad \frac{\lambda_1}{\lambda_1 + \lambda_2}$

Sol. ① $\mathbb{P}(\min(X_1, X_2) > t) \overset{indep.}{=} \mathbb{P}(X_1 > t)\, \mathbb{P}(X_2 > t) = e^{-\lambda_1 t}\, e^{-\lambda_2 t}$

② $\mathbb{P}(X_1 < X_2) = \int_0^{+\infty} f_{X_1}(t)\, \underbrace{\mathbb{P}[X_2 > t \mid X_1 = t]}_{\text{Cond. prob. density}}\, dt = \int_0^{+\infty} f_{X_1}(t) \cdot \mathbb{P}(X_2 > t)\, dt = \int_0^{+\infty} \lambda_1 e^{-\lambda_1 t} \cdot e^{-\lambda_2 t}\, dt = \frac{\lambda_1}{\lambda_1 + \lambda_2}$

formally, we should write : $1 - F_{X_2|X_1}(t,t)$

#

eg. $\mathbb{P}_{\lambda,t}(A) \triangleq \int_A \frac{1}{\Gamma(t)} \cdot \lambda^t x^{t-1} e^{-\lambda x}\, dx$, $\lambda, t > 0$

$\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x}\, dx$ Gamma Function

$\boxed{\text{Gamma } (\lambda, t)}$

- $t=1$, $\Gamma(1)=1$, $f_{\lambda,1}(x) = \lambda e^{-\lambda x} \triangleq \text{Exp}(\lambda)$
- $t=n$, $\Gamma(n) = (n-1)!$, $f_{\lambda,n}(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}$

Fact: If $X_1, \cdots, X_n$ are indep. with density funct. $\text{Exp}(\lambda) = \lambda e^{-\lambda x}$

Then $X_1 + \cdots + X_n \sim \text{Gamma}(\lambda, n)$

Sketch: discrete r.v. $\mathbb{P}(X+Y=x) = \sum_a \mathbb{P}(X=a, Y=x-a)$

$\overset{\text{ind.}}{=} \sum_a \mathbb{P}(X=a)\,\mathbb{P}(Y=x-a)$

$= \sum_a f_X(a) \cdot f_Y(x-a)$ Convolution

continuous r.v. $f_{X+Y}(a) = \int f_X(a) \cdot f_Y(x-a)\, da$

[proved by induction]

$\mathbb{P}(X+Y \le a) = \int_{\{X+Y\le a\}} f_{X,Y}(x,y)\, dx dy = \int_{\{X+Y\le a\}} f_X(x) f_Y(y)\, dx dy$

$= \int f_X(x) \cdot \int_{-\infty}^{a-x} f_Y(y)\, dy \cdot dx$

$\boxed{\text{Poisson Process:}}$

Def. A poisson process $(N_s)_{s\ge 0}$ with rate λ satisfies

① $N_0 = 0$

② $\forall\, 0 \le s \le t$, $N_t - N_s \sim \text{Poisson}(\lambda(t-s))$, where $\text{Poisson}(\lambda) = \frac{\lambda^n}{n!} e^{-\lambda}$

making N_t, N_s indep.

③ $\forall\, 0 \le t_1 \le t_2 \le \cdots \le t_n$, $N_{t_1}, N_{t_2} - N_{t_1}, \cdots, N_{t_n} - N_{t_{n-1}}$ are mutually indep.

Alternative Construction:

Prop. Let $\tau_1, \tau_2, \cdots, \tau_n$ be indep. r.v.s with $\text{Exp}(\lambda)$ distribution $= \lambda e^{-\lambda x}$

Let $T_n = \sum_{i=1}^n \tau_i$ and $N_s = \max\{n : T_n \le s\}$

Then $(N_s)_{s\ge 0}$ is Poisson process with rate λ

pf. 1° $N_0 = 0$ trivial

2° Assume $S=0$. $\mathbb{P}(N_t = n) = \mathbb{P}(T_n \le t, T_{n+1} > t)$

$= \int_0^t f_{T_n}(s) \cdot \mathbb{P}[T_{n+1} > t \mid T_n = s]\, ds$

$= \int_0^t \underbrace{f_{T_n}(s)}_{\text{Gamma}(\lambda, n)}\ \underbrace{\mathbb{P}[\tau_{n+1} > t-s]}\, ds$

If $\tau \sim \text{Exp}(\lambda)$,
Then $\mathbb{P}(\tau > t) = e^{-\lambda t}$

$= \int_0^t \frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s} e^{-\lambda(t-s)}\, ds$

$= \frac{\lambda^n}{(n-1)!} e^{-\lambda t} \cdot \int_0^t s^{n-1}\, ds = \frac{\lambda^n t^n}{n!} e^{-\lambda t} = \text{Poisson}(\lambda t)$

To show $\forall t, s \ge 0$, $N_{t+s} - N_s \sim \text{Poisson}(\lambda t)$

Assume $N_s = n$, by lack of memory,

$\mathbb{P}[\tau_{n+1} > s+t-T_n \mid \tau_{n+1} > s - T_n]$

$= \mathbb{P}[\tau_{n+1} > t] = \mathbb{P}[\tau_1 > t] = e^{-\lambda t}$

Thus $N_{t+s} - N_s \sim N_t \sim \text{Poisson}(\lambda t)$

3° If $s \ge r$, we have $N_{t+s} - N_s$ indep. of N_r for $\forall r \le s \le t$

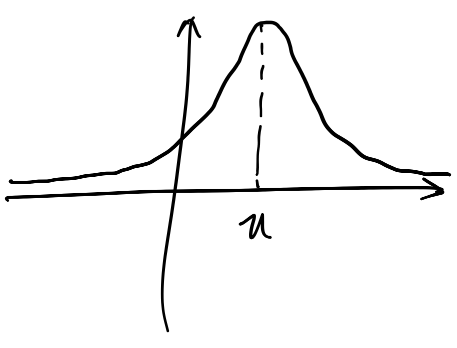
$r \quad s \quad t \quad s{+}t$

because $N_{t+s} - N_t \sim \text{Poisson}(\lambda t)$

$\Rightarrow N_{t_2} - N_{t_1}, N_{t_3} - N_{t_2}, \cdots, N_{t_n} - N_{t_{n-1}}$ are indep.

#

$\boxed{\text{Normal Distribution:}}$



Eg. $N(0,1)$: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$N(\mu, \sigma^2)$: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, Note that $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx = \sqrt{2\pi}$

In fact, if $X \sim N(\mu, \sigma^2)$, then $\mathbb{E}X = \mu$

$\text{Var}\, X = \sigma^2$

$\boxed{(X: \text{ Gaussian})}$

Hint:

$\mathbb{E}X = \underbrace{\int (x-\mu) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx}_{0} + \mu$

For variance, use $\int x^2 e^{-\lambda x^2}\, dx = -\frac{d}{d\lambda} \overbrace{\int e^{-\lambda x^2}\, dx}^{\sqrt{\frac{\pi}{\lambda}}}$

Fact: If $X \sim N(\mu, \sigma^2)$

Then $Y = \frac{X-\mu}{\sigma}$ is $N(0,1)$

pf. $\mathbb{P}(Y \le a) = \mathbb{P}(\frac{X-\mu}{\sigma} \le a)$

$= \mathbb{P}(X \le a\sigma + \mu)$

$= \int_{-\infty}^{\sigma a + \mu} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$

$\overset{y = \frac{x-\mu}{\sigma}}{=\!=} \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy$

#

Fact: If $X \sim N(0,1)$, $\mathbb{E} e^{tX} = e^{\frac{1}{2}t^2}$

$\boxed{\text{moment generating function}}$

pf. $\mathbb{E}e^{tX} = \frac{1}{\sqrt{2\pi}} \int e^{tx} e^{-\frac{1}{2}x^2}\, dx = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}(x-t)^2} e^{\frac{t^2}{2}}\, dx = e^{\frac{t^2}{2}} \cdot \int \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}}\, dx = e^{\frac{t^2}{2}}$

#

Ex. If $Y \sim N(\mu, \sigma^2)$, compute $\mathbb{E}e^{tY}$.

pf. $\mathbb{E}e^{tY} = \frac{1}{\sqrt{2\pi\sigma^2}} \int e^{ty} \cdot e^{-\frac{(y-\mu)^2}{2\sigma^2}}\, dy = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \int e^{\frac{1}{2\sigma^2}(2\sigma^2 ty - y^2 - \mu^2 + 2y\mu)}\, dy = e^{\frac{t(\sigma^2 t + 2\mu)}{2}} \cdot \int \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y-(\sigma^2 t+\mu))^2}{2\sigma^2}}\, dy = e^{\frac{\sigma^2 t^2 + 2\mu t}{2}}$

3/6 HTOP 12

Wednesday, March 6, 2024   11:17 AM

Recap: $\upsilon \ll \mu$ if $\mu(A)=0$, we have $\upsilon(A)=0$

- Radon–Nikodym thm: $\upsilon(A) = \int_A f \, d\mu$ for some f ($f = \frac{d\upsilon}{d\mu}$)
- $\upsilon \sim \mu$ iff $\mu \ll \upsilon$, $\upsilon \ll \mu$

If $\upsilon \sim \mu$, "change of frame"

prop. (Chain Rule) sps. $\upsilon \ll \mu$ and f is (Lebesgue) integrable w.r.t. υ

$$\Leftrightarrow f \cdot \frac{d\upsilon}{d\mu} \text{ is integrable w.r.t. } \mu \text{ and } \int f \, d\upsilon = \int f \cdot \frac{d\upsilon}{d\mu} d\mu$$

[Probabilistic: if $Q \ll \mathbb{P}$, then $\mathbb{E}^Q[X] = \mathbb{E}^{\mathbb{P}}[X \cdot \frac{dQ}{d\mathbb{P}}]$)

pf.
- If $f = 1_A$, $A \in \mathcal{F}$, then $\int 1_A \, d\upsilon = \upsilon(A) = \int_A \frac{d\upsilon}{d\mu} \cdot d\mu = \int 1_A \cdot \frac{d\upsilon}{d\mu} \cdot d\mu$
- By linearity, if $\varphi = \sum_{i=1}^{n} a_i 1_{A_i}$, $a_i \geq 0$, $A_i \cap A_j = \phi$

  Then $\int \varphi \, d\upsilon = \int \varphi \cdot \frac{d\upsilon}{d\mu} \cdot d\mu$
- Let $g \geq 0$ be a Borel function, take $(\varphi_n)_{n\geq 1}$ be a seq. of increasing simple function $\varphi_n \uparrow g$

By MCT: $\int g \, d\upsilon = \lim_{n\to+\infty} \int \varphi_n \, d\upsilon = \lim_{n\to+\infty} \int \varphi_n \cdot \frac{d\upsilon}{d\mu} \cdot d\mu = \int g \cdot \frac{d\upsilon}{d\mu} \cdot d\mu$

- Finally, for g be Borel, write $g = g^+ - g^-$

#

eg. $(\Omega, \mathcal{F}, \mathbb{P})$, $X \sim \mathcal{N}(0,1)$, $X+\theta \sim \mathcal{N}(\theta, 1)$ where $\theta \in \mathbb{R}$

There is a new propability measure Q s.t. under Q, $X + \theta \sim \mathcal{N}(0,1)$

If $Y \sim \mathcal{N}(\mu, \sigma^2)$

Then $X = \frac{Y-\mu}{\sigma} \sim \mathcal{N}(0,1)$

Then $\mathbb{E} e^{tX} = e^{\frac{1}{2}t^2}$

Sol. Take $\frac{dQ}{d\mathbb{P}} = e^{-\theta X - \frac{1}{2}\theta^2}$

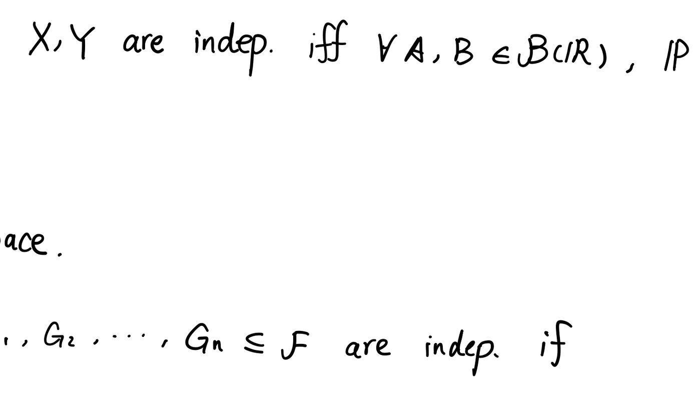
Then
$$\mathbb{E}^Q e^{t(X+\theta)} = \mathbb{E}^{\mathbb{P}}\left[e^{t(X+\theta)} \cdot \frac{dQ}{d\mathbb{P}}\right]$$
$$= \mathbb{E}^{\mathbb{P}}\left[e^{tX} \cdot e^{t\theta} \cdot e^{-\theta X} \cdot e^{-\frac{1}{2}\theta^2}\right]$$
$$= \mathbb{E}^{\mathbb{P}}\left[e^{(t-\theta)X}\right] e^{\theta t - \frac{1}{2}\theta^2}$$

$X \overset{\mathbb{P}}{\sim} \mathcal{N}(0,1)$
$$= e^{\frac{1}{2}(t-\theta)^2} \cdot e^{\theta t - \frac{1}{2}\theta^2} = e^{\frac{1}{2}t^2}$$

i.e. $X + \theta \overset{Q}{\sim} \mathcal{N}(0,1)$

#

Risk-Neutral Measure



Girsanov theorem

§. Independence

Intuitively, two r.v. X, Y are indep. iff $\forall A, B \in \mathcal{B}(\mathbb{R})$, $\mathbb{P}(X\in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y\in B)$

General definition:

$(\Omega, \mathcal{F}, \mathbb{P})$ prob. space.

Def. ⓐ a seq. of $G_1, G_2, \cdots, G_n \subseteq \mathcal{F}$ are indep. if

$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n)$ for all $A_i \in G_j$, $i = 1, 2, \cdots, n$

ⓑ a seq. of r.v. $X_1, \cdots, X_n$ are indep. if $\sigma(X_1), \cdots, \sigma(X_n)$ are indep.

where $\sigma(X_i) = \{X_i^{-1}(B), B \in \mathcal{B}(\mathbb{R})\}$

Rmk#1: ⓐ, ⓑ $\Longleftrightarrow$ $\forall B_1, \cdots, B_n \in \mathcal{B}(\mathbb{R})$

$\mathbb{P}(X_1 \in B_1, X_2 \in B_2, \cdots, X_n \in B_n)$

$= \underbrace{\mathbb{P}(X_1 \in B_1)}_{X_1^{-1}(B_1)} \underbrace{\mathbb{P}(X_2 \in B_2)}_{X_2^{-1}(B_2)} \cdots \underbrace{\mathbb{P}(X_n \in B_n)}_{X_n^{-1}(B_n)}$

Rmk #2: $A_1, \cdots, A_n$ are indep. $\Leftrightarrow 1_{A_1}, 1_{A_2}, \cdots, 1_{A_n}$ are indep. r.v. s (ex.)

§. Sufficient condition for independence

Main Thm. sps. $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_n$ are indep., each $\mathcal{A}_i$ is a π-system

Then $\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), \cdots, \sigma(\mathcal{A}_n)$ are indep.

Def (π-system): $\mathcal{G} \subseteq \mathcal{F}$ is a π-system if
- $\phi \in \mathcal{G}$
- $\forall A_1, A_2 \in \mathcal{G}$, $A_1 \cap A_2 \in \mathcal{G}$

Def. (λ-system): $\mathcal{L}$ is a λ-system if
- $\Omega \in \mathcal{L}$
- if $A, B \in \mathcal{L}$, $A \subseteq B$, then $B \setminus A \in \mathcal{L}$
- if $A_n \uparrow A$, and each $A_n \in \mathcal{L}$, then $A \in \mathcal{L}$

thm. If $\mathcal{A}$ is a π-system and a λ-system,

then it is a σ-algebra

- Examples of π-systems (generates $\mathcal{B}(\mathbb{R})$) pf.

ⓐ $\mathcal{E}_1 = \{(a, b], -\infty \leq a \leq b \leq +\infty\}$ (ex.)

ⓑ $\mathcal{E}_2 = \{(a, b), -\infty \leq a \leq b \leq +\infty\}$ (ex.)

ⓒ $\mathcal{E}_{open} = \{A \subseteq \mathbb{R}, A \text{ is open}\}$

ⓓ $\mathcal{E}_{closed} = \{A \subseteq \mathbb{R}, A \text{ is closed}\}$ (ex.)

pf. of ⓒ: Let $A \subseteq \mathbb{R}$ be open

Then $A = \bigcup_{i=1}^{+\infty} (x_i - \varepsilon_i, x_i + \varepsilon_i)$

$\Rightarrow \mathcal{E}_2 \subseteq \mathcal{E}_{open}$

and $\mathcal{E}_{open} \subseteq \sigma(\mathcal{E}_2)$

$\Rightarrow \sigma(\mathcal{E}_2) = \sigma(\mathcal{E}_{open})$

#

Dynkin's π-λ Theorem: If $\mathcal{G}$ is a π-system, $\mathcal{L}$ is a λ-system, $\mathcal{L} \supseteq \mathcal{G}$

Then $\mathcal{L} \supseteq \sigma(\mathcal{G})$

Cor. for Main Thm.

$X_1, X_2, \cdots, X_n$ are indep. iff $\forall x_1, \cdots, x_n \in \mathbb{R}$

$$\underbrace{\mathbb{P}[X_1 \leq x_1, \cdots, X_n \leq x_n]}_{\text{joint prob. distribution function}} = \mathbb{P}[X_1 \leq x_1] \cdots \mathbb{P}[X_n \leq x_n] \quad (*)$$

$\Leftarrow$):

pf. Let $\mathcal{A}_i = \{\{X_i \leq x\}, x \in \mathbb{R}\}$ then $\mathcal{A}_i$ is a π-system

since $\{X_i \leq x\} \cap \{X_i \leq y\} = \{X_i \leq \min(x, y)\}$

and $\sigma(\mathcal{A}_i) = \sigma(X_i)$ (ex.)

$(*) \Rightarrow \mathcal{A}_1, \cdots, \mathcal{A}_n$ are indep. $\xRightarrow{\text{Main Thm.}}$ $\sigma(X_1), \cdots, \sigma(X_n)$ are indep.

$\Rightarrow$): Each $\mathcal{A}_i$ is a specific $B \in \mathcal{B}(\mathbb{R})$, trivial.

#

pf. of the Main Thm:

We prove if $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_n$ are indep.

then $\sigma(\mathcal{A}_1), \mathcal{A}_2, \cdots, \mathcal{A}_n$ are indep. then iterate.

Take $A_2 \in \mathcal{A}_2, \cdots, A_n \in \mathcal{A}_n$. Set $B = A_2 \cap A_3 \cap \cdots \cap A_n$

Define: $\mathcal{L} = \{A \in \Omega : \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)\}$

if $A \in \mathcal{A}_1$, then $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B) \Rightarrow \mathcal{L} \supseteq \mathcal{A}_1$

We claim that $\mathcal{L}$ is a λ-system:

① $\Omega \in \mathcal{L}$, since $\mathbb{P}(\Omega \cap B) = \mathbb{P}(\Omega)\mathbb{P}(B)$

② If $A_1', A_2' \in \mathcal{L}$, $A_1' \subset A_2'$, $\mathbb{P}((A_2' - A_1') \cap B) = \mathbb{P}(A_2' \cap B) - \mathbb{P}(A_1' \cap B)$

$$= \mathbb{P}(A_2')\mathbb{P}(B) - \mathbb{P}(A_1')\mathbb{P}(B)$$
$$= \mathbb{P}(A_2' - A_1')\mathbb{P}(B)$$

③ If $A_n \in \mathcal{L}$, $A_n \uparrow A$.

Then
$$\mathbb{P}(A \cap B) \overset{\text{cont.}}{=} \lim_{n\to+\infty} \mathbb{P}(A_n \cap B)$$
$$= \lim_{n\to+\infty} \mathbb{P}(A_n)\mathbb{P}(B)$$
$$\overset{\text{conti.}}{=} \mathbb{P}(A)\mathbb{P}(B)$$

Then by Dynkin's π-λ Thm, then $\mathcal{L} \supseteq \sigma(\mathcal{A}_1)$

This implies $\forall A \in \sigma(\mathcal{A}_1)$, $\mathbb{P}(A \cap (A_2 \cap A_3 \cap \cdots \cap A_n)) \overset{A\in\mathcal{L}}{=} \mathbb{P}(A)\mathbb{P}(A_2 \cap A_3 \cap \cdots \cap A_n)$

$$= \mathbb{P}(A)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n)$$

at this stage, we assume $\mathcal{A}_2, \cdots, \mathcal{A}_n$ are indep.

$\Rightarrow \sigma(\mathcal{A}_1), \mathcal{A}_2, \cdots, \mathcal{A}_n$ are indep.

Then by iteration, we have $\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), \cdots, \sigma(\mathcal{A}_n)$ are indep.

#

3/8 HTOP (Reci 6)

Friday, March 8, 2024 11:31 AM

ex. $(\Omega, \mathcal{F}, \mathbb{P})$ $X: \Omega \to [0, +\infty)$, R.V.

Assume $\exists M < +\infty$ s.t. $\mathbb{E}X^n \le M$ for $\forall n = 1, 2, \cdots$

① Prove that $\mathbb{P}(X > 1) = 0$ and $\mathbb{P}(X = 1) \le M$

② Compute: $\lim_{n \to +\infty} \mathbb{E}X^n$

③ If $\mathbb{E}X^n = M$ for $\forall n = 1, 2, \cdots$, Show that $\mathbb{P}(X \in \{0, 1\}) = 1$

Sol.

① $\forall \varepsilon > 0$, $M \ge \mathbb{E}X^n = \int X^n d\mathbb{P} \ge \int_{\{X \ge 1+\varepsilon\}} X^n d\mathbb{P} \ge (1+\varepsilon)^n \int_{\{X \ge 1+\varepsilon\}} d\mathbb{P} = (1+\varepsilon)^n \mathbb{P}(X \ge 1+\varepsilon)$

sending $n \to +\infty$: $\mathbb{P}(X > 1+\varepsilon) = 0$

By continuity of measure: $\mathbb{P}(X > 1) = \lim_{k \to +\infty} \mathbb{P}(X \ge 1 + \frac{1}{k}) = 0$

$M \ge \mathbb{E}X^n \ge \int_{\{X=1\}} X^n d\mathbb{P} = \mathbb{P}(X = 1)$ #

② Claim: $\lim_{n \to +\infty} \mathbb{E}X^n = \mathbb{P}(X = 1)$

$\mathbb{E}X^n = \int_{\{0<X<1\}} X^n d\mathbb{P} + \int_{\{X=1\}} X^n d\mathbb{P}$, note that for $0 < X < 1$, $X^n \downarrow 0$

By MCT: $\lim_{n \to +\infty} \int_{\{0<X<1\}} X^n d\mathbb{P} = \int_{\{0<X<1\}} \lim_{n \to +\infty} X^n d\mathbb{P} = 0$

Hence $\lim_{n \to +\infty} \mathbb{E}X^n = \lim_{n \to +\infty} \int_{\{X=1\}} X^n d\mathbb{P} = \mathbb{P}(X = 1)$

#

$\boxed{\text{Bounded Convergence Thm.}}$

$(\Omega, \mathcal{F}, \mu)$, $\mu(\Omega) < +\infty$

Then if $f_n \to f$ a.e. and

$\sup_{n \ge 1} |f_n| < +\infty$, then $\lim_{n \to +\infty} \int f_n d\mu = \int f d\mu$

or ($\sup_{n \ge 1} |f_n| \le g$ and $\int g d\mu < +\infty$)

Two approaches to prove it:

① MCT

② Fatou's Lemma

③ It suffices to show that $\mathbb{P}(0 < X < 1) = 0$

We claim: $\forall \varepsilon > 0$, $\mathbb{P}(\varepsilon < X < 1 - \varepsilon) = 0$

$\mathbb{E}X^n \ge \int_{\{X=1\}} X^n d\mathbb{P} + \int_{\{\varepsilon < X < 1-\varepsilon\}} X^n d\mathbb{P} \ge \mathbb{P}(X=1) + \varepsilon^n \mathbb{P}(\varepsilon < X < 1 - \varepsilon)$

Since $\mathbb{E}X^n$ is constant, but by ②: $\lim_{n \to +\infty} \mathbb{E}X^n = \mathbb{P}(X=1)$ $\Rightarrow$ $\mathbb{E}X^n = \mathbb{P}(X=1) = M$

$\Rightarrow \mathbb{P}(\varepsilon < X < 1 - \varepsilon) = 0$

Then by conti. of measure,

$\mathbb{P}[0 < X < 1] = \lim_{k \to +\infty} \mathbb{P}[\frac{1}{k} < X < 1 - \frac{1}{k}] = 0$

#

Concentration Inequality:

Last time: [Payley - Zygmund inequality]. (Reci 5)

- If $X \ge 0$, $\theta \in (0,1)$, $\mathbb{P}(X > \theta \mathbb{E}X) \ge (1-\theta)^2 \cdot \frac{(\mathbb{E}X)^2}{\mathbb{E}(X^2)}$

$\boxed{\text{[Chebyshev Inequality]}}$:

- $\mathbb{P}(|X| \ge t) \le \frac{1}{t^k} \mathbb{E}|X|^k$ $(k \in \mathbb{R}^+)$

Pf. $\mathbb{E}|X|^k = \int |X|^k d\mathbb{P}$

$\ge \int_{\{|X| \ge t\}} |X|^k d\mathbb{P}$

$\ge t^k \int_{\{|X| \ge t\}} d\mathbb{P}$

$= t^k \cdot \mathbb{P}(|X| \ge t)$ #

Cor. $\boxed{\text{[Markov Inequality]}}$

- $\mathbb{P}(|X - \mathbb{E}X| \ge t) \le \frac{1}{t^2} \mathbb{E}|X - \mathbb{E}X|^2 = \frac{1}{t^2} \mathrm{Var}\, X$
- $\mathbb{P}(|X - \mathbb{E}X| \ge t) \le \frac{1}{t^k} \mathbb{E}|X - \mathbb{E}X|^k$

$\boxed{\text{Improved Chebyshev}}$:

- If $\mathbb{E}e^{tX} < +\infty$, $\forall t \in \mathbb{R}$

Then $\mathbb{P}(X > a) = \mathbb{P}[e^{tX} > e^{aX}] \le \underbrace{e^{-at} \mathbb{E}e^{tX}}_{\text{hence we can optimize over } t}$

$\boxed{\text{Chernoff Bound}}$.

ex. Let $X_1, \cdots, X_n$ be indep. $X_i \sim \mathrm{Bernoulli}(p)$ $\Rightarrow$ we can change $p \to p_i$

and $\mu = \sum_i p_i$,

then $\mathbb{P}[|X - \mu| \ge \delta\mu] \le e^{-\frac{\delta^2}{3}\mu}$ $\Rightarrow$ $\boxed{\text{Chernoff Bound}}$

Let $X = \sum_{i=1}^n X_i$, $\mu = \mathbb{E}X = n \cdot p$

Show that $\mathbb{P}(X \ge (1+\delta)\mu) \le \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$ ①

$\mathbb{P}(X \le (1-\delta)\mu) \le \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^\mu$ ②

for $\forall \delta > 0$

In fact, $\frac{e^\delta}{(1+\delta)^{1+\delta}} \le e^{-\frac{\delta^2}{3}}$ (ex.)

for $\delta \in [0, 1]$.

where $\mu = np$

Sol. $\mathbb{P}(X \ge (1+\delta)\mu) = \mathbb{P}(e^{Xt} \ge e^{(1+\delta)\mu t}) \le e^{-(1+\delta)\mu t} \mathbb{E}e^{tX}$, by Chebyshev. picking $k=1$

$\mathbb{E}e^{tX} = \mathbb{E}(e^{tX_1} e^{tX_2} \cdots e^{tX_n})$

$\overset{\text{indep.}}{=} [\mathbb{E}e^{tX_1}]^n = (1 - p + e^t p)^n$

$= (1 + p(e^t - 1))^n$

$\le e^{p(e^t-1) \cdot n} = e^{\mu(e^t - 1)}$

Rmk 1. We use the idea of moment generating function here. (see Lec 20)

Rmk 2. See the essential ideas at Reci 11.

$\Rightarrow \mathbb{P}(X \ge (1+\delta)\mu) \le \frac{e^{\mu(e^t-1)}}{e^{\mu(1+\delta)t}} = e^{\mu(e^t - 1 - t(1+\delta))}$

to minimize $\mu(e^t - 1 - t(1+\delta)) = f(t)$,

$f'(t) = \mu(e^t - (1+\delta)) = 0$ $\Rightarrow$ $t = \ln(1+\delta)$

Thus $\mathbb{P}(X \ge (1+\delta)\mu) \le e^{\mu(\delta - t - t\delta)} = \left(\frac{e^\delta}{e^{t(1+\delta)}}\right)^\mu = \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$

② is proved similarly.

#

Recall: $X_1, X_2, \cdots, X_n$ are indep.

$\Leftrightarrow \mathbb{P}(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n)$ ← joint distribution function

$= \mathbb{P}(X_1 \le x_1)\mathbb{P}(X_2 \le x_2)\cdots\mathbb{P}(X_n \le x_n)$

(ex.) $\Omega = [0,1]\times[0,1]$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^2)\cap[0,1]^2$, m = Lebesgue measure in $\mathbb{R}^2$

$\mathcal{A}_1 = \{[0,1]\times A : A \in \mathcal{B}(\mathbb{R})\cap[0,1]\}$

$\mathcal{A}_2 = \{A\times[0,1] : A \in \mathcal{B}(\mathbb{R})\cap[0,1]\}$

Show that $\mathcal{A}_1, \mathcal{A}_2$ are σ-algebras & they are indep. w.r.t. m

Thm. (Uniqueness of Extension)

Let $\mathcal{C}$ be a π-system. Let μ_1, μ_2 be two finite measures (σ-finite) on $(\Omega, \sigma(\mathcal{C}))$ s.t. $\mu_1(\Omega) = \mu_2(\Omega)$ and $\mu_1(\Omega) = \mu_2(\Omega)$ on $\mathcal{C}$.

Then $\mu_1 = \mu_2$ on $\sigma(\mathcal{C})$.

finite: $\mu(\Omega) < +\infty$

σ-finite: $\Omega = \bigcup_{j=1}^{+\infty} A_j$, where $A_j \in \mathcal{F}$, $\mu(A_j) < +\infty$

π-system:

λ-system:

π-λ thm:

Pf. Let $D = \{A \in \sigma(\mathcal{C}) . \mu_1(A) = \mu_2(A)\}$

Then we have $D \supseteq \mathcal{C}$.

We claim: D is a λ-system.

Indeed, • $\Omega \in D$, since $\mu_1(\Omega) = \mu_2(\Omega)$

• If $A, B \in D$, $A \subseteq B$

Then $\mu_2(B\setminus A) = \mu_2(B) - \mu_2(A) \overset{A,B\in D}{=} \mu_1(B) - \mu_1(A) = \mu_1(B\setminus A)$

• If $A_n \in D$, $A_n \uparrow A$, $\mu_2(A) \underset{\text{continuity from below}}{=} \lim_{n\to+\infty} \mu_2(A_n) \underset{A_n\in D}{=} \lim_{n\to+\infty} \mu_1(A_n) = \mu_1(A)$

Applying the π-λ thm: $D \supseteq \sigma(\mathcal{C})$

Hence $D = \sigma(\mathcal{C})$

#

Construct indep. R.V.s

• Let $F_1, F_2, \cdots, F_n$ be distribution functions. How to construct independent R.V.s $X_1, \cdots, X_n$ s.t. $\mathbb{P}(X_i \le x) = F_i(x)$

Sol. Let $\Omega = \mathbb{R}^n$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^n)$, Let $X_i : \mathbb{R}^n \to \mathbb{R}$

$\vec{\omega} \to X_i(\vec{\omega}) = w_i$

Let $\mathbb{P}_n$ be the measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ s.t.

$\mathbb{P}_n([a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_n,b_n]) = (F_1(b_1) - F_1(a_1))\cdots(F_n(b_n) - F_n(a_n))$

Pf. Since $\{[a_1,b_1]\times\cdots\times[a_n,b_n]\}$ is a π-system that generates $\mathcal{B}(\mathbb{R}^n)$.

By extension Thm., $\mathbb{P}_n$ extends uniquely to a measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

Indeed, $\mathbb{P}_n = m_{F_1}\otimes m_{F_2}\otimes\cdots\otimes m_{F_n}$, m_{F_i} is 1-d Lebesgue-Steltjes measure 和 Lebesgue measure 区别?

#

$m_F((a,b]) = F(b) - F(a)$ for right-continuous function F

• How to construct infinite seq. of indep. R.V.s $X_1, X_2, \cdots$?

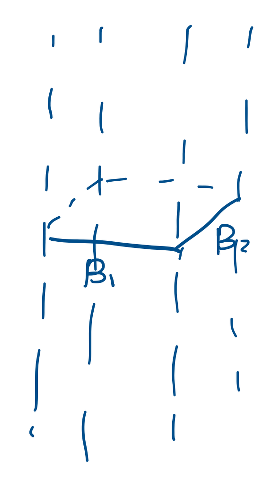
Sol. Try a similar construction, $\Omega = \mathbb{R}^{\mathbb{N}} = \{(w_1, w_2, \cdots), w_i \in \mathbb{R}\}$

Goal: define σ-algebra, prob. measure based on approximation in finite dimension

$\mathcal{A} = \bigcup_{n=1}^{+\infty}\{B_1\times B_2\times\cdots\times B_n\times\mathbb{R}\times\mathbb{R}\times\cdots, B_i \in \mathcal{B}(\mathbb{R})\}$

Name: Cylinder set

Rmk. $\mathcal{A}$ is an algebra (ex.), but it's NOT a σ-algebra



Counter-example: let $A_n = \mathbb{R}\times\cdots\times\mathbb{R}\times[0,1]\times\mathbb{R}\times\cdots$, ← n-th coordinate

Then $\bigcap_{n=1}^{+\infty} A_n = [0,1]^{\mathbb{N}} \notin \mathcal{A}$

Hence, we take $\mathcal{F} = \sigma(\mathcal{A})$

Let $X_i : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$

$\vec{\omega} \to X_i(\vec{\omega}) = w_i$

We write content here because $\mathcal{A}$ is just an algebra

Natural to define a content (countably additive) on $\mathcal{A}$: $\forall A \in \mathcal{A}$, $\exists n \in \mathbb{N}$ s.t. $A = B_1\times\cdots\times B_n\times\mathbb{R}\times\mathbb{R}\times\cdots$

Set $\mathbb{P}(B_1\times\cdots\times B_n\times\mathbb{R}\times\cdots) = \mathbb{P}_n(B_1\times B_2\times\cdots\times B_n) \triangleq m_{F_1}\otimes\cdots\otimes m_{F_n}$

• $\boxed{\text{Kolmogorov Extension Thm:}}$

sps. the prob. measures $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P}_n)_{n\ge1}$ are consistent:

$\mathbb{P}_{n+1}((a_1,b_1]\times\cdots\times(a_n,b_n]\times\mathbb{R}) = \mathbb{P}_n((a_1,b_1]\times\cdots\times(a_n,b_n])$, $\forall n \ge 1$

then there is a unique prob. measure $\mathbb{P}$ on $(\mathbb{R}^{\mathbb{N}}, \sigma(\mathcal{A}))$ s.t.

$\mathbb{P}((a_1,b_1]\times\cdots\times(a_n,b_n]\times\mathbb{R}\times\cdots) = \mathbb{P}_n((a_1,b_1]\times\cdots\times(a_n,b_n])$

• Applications.

$(X_i)_{i\in\mathbb{N}}$ ($\to (X_t)_{t\in[a,b]}$) is a random walk / Markov Chain. (an example of)

If we have disjoint distribution of $(X_{t_1}, \cdots, X_{t_n})$, $\forall n \ge 1$.

Then we can uniquely determine the law of $(X_i)_{i\in\mathbb{N}}$

Markov Chain:

$\mathbb{P}[X_n \in A \mid X_1, \cdots, X_{n-1}]$

$= \mathbb{P}[X_n \in A \mid X_{n-1}]$

§ Joint Distribution Function

Rmk. There also exists joint density function for multi-variable case

• The $\boxed{\text{joint distribution function}}$ of X and Y is $F: \mathbb{R}^2 \to [0,1]$

where $F(x,y) \triangleq \mathbb{P}(X \le x, Y \le y)$

Def. If X and Y are absolutely continuous R.V.s, then the $\boxed{\text{joint density function}}$ $f: \mathbb{R}^2 \to [0,+\infty)$ is defined by

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv, \quad \forall x,y \in \mathbb{R}$$

$$\Rightarrow f(x,y) \triangleq \frac{\partial^2}{\partial x\partial y}F(x,y)$$

Moreover, $\mathbb{P}(X \in (a,b], Y \in (c,d]) = \int_a^b\int_c^d f(u,v)\,du\,dv = F(a,c) + F(b,d) - F(a,d) - F(b,c)$

By Extension Thm: $\mathbb{P}(X \in A, Y \in B) = \int_A\int_B f(u,v)\,du\,dv$, $A, B \in \mathcal{B}(\mathbb{R})$

• May recover the distribution for individual R.V.s

① $\mathbb{P}(X \le x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{x} f(u,v)\,du\,dv$; $f_X(x) = \int_{-\infty}^{+\infty} f(x,v)\,dv$

② $\mathbb{P}(Y \le y) = \int_{-\infty}^{y}\int_{-\infty}^{+\infty} f(u,v)\,du\,dv$; $f_Y(y) = \int_{-\infty}^{+\infty} f(u,y)\,du$

Rmk. X, Y are indep. $\Leftrightarrow F(x,y) = F_X(x)F_Y(y) \Leftrightarrow f(x,y) = f_X(x)f_Y(y)$

eg. If X, Y have joint distribution function $f(x,y) = \frac{\alpha^x\beta^y}{x!\,y!}e^{-\alpha-\beta}$, $x, y \in \mathbb{N}$

Then X, Y are indep: $f_X(x) = \sum_{y\in\mathbb{N}} f(x,y) = \frac{\alpha^x}{x!}e^{-\alpha}\cdot\underbrace{\sum_{y\in\mathbb{N}}\frac{\beta^y}{y!}e^{-\beta}} = \frac{\alpha^x}{x!}e^{-\alpha}$. Thus, $X \sim \text{Poisson}(\alpha)$, $Y \sim \text{Poisson}(\beta)$

3/13 HTOP 14

Wednesday, March 13, 2024 11:18 AM

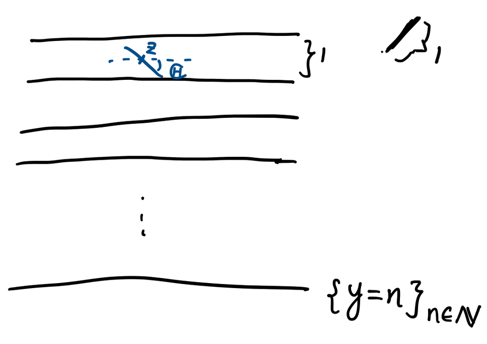
Recall:

Joint probability distribution function $F(x,y)$

density function $f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y)$, marginal density: $f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy$

$f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx$

eg. $\boxed{\text{(Buffon's needle):}}$



$\mathbb{P}[\text{Intersect some lines}] = ?$

$\{y=n\}_{n\in\mathbb{N}}$

Z – center, Θ – angle

$X \sim \text{unif}[0,1]$

$f_X(x) = \begin{cases} 1, & \text{if } x\in[0,1] \\ 0, & \text{else} \end{cases}$

$Z \sim \text{unif}[0,1]$: $f_Z(z) = 1$ if $z\in[0,1)$

$\Theta \sim \text{unif}[0,\pi]$: $f_\Theta(\theta) = \frac{1}{\pi}$ if $\theta\in[0,\pi)$

$\mathbb{P}(X\in[a,b]) = b-a$

$\{\text{intersection}\} = \{(z,\theta) : \overbrace{\tfrac{1}{2}\sin\theta \geq z \text{ or } \tfrac{1}{2}\sin\theta \geq 1-z}^{B}\}$

$\mathbb{P}(X\in[x,x+\Delta x]) = f_X(x)\cdot\Delta x$

$$\mathbb{P}(\{\text{intersection}\}) = \int_B f(z,\theta)\,dz\,d\theta = \frac{1}{\pi}\int_0^\pi d\theta \cdot \int_0^{\frac{1}{2}\sin\theta} dz + \frac{1}{\pi}\int_0^\pi d\theta \int_{1-\frac{1}{2}\sin\theta}^1 dz = \frac{2}{\pi}\int_0^\pi \tfrac{1}{2}\sin\theta\,d\theta = \frac{2}{\pi}$$

\# Assuming center & angle are indep.

Sum of Random Variables:

Lemma: Let X, Y be conti. R.V. s $f_{X+Y}(z) = \int_{-\infty}^{+\infty} f(x, z-x)\,dx$

pf. $F_{X+Y}(z) = \mathbb{P}(X+Y\leq z) = \int_{\{X+Y\leq z\}} f(x,y)\,dx\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{z-x} f(x,y)\,dy\,dx$

$\Rightarrow f_{X+Y}(z) = \int_{-\infty}^{+\infty} f(x,z-x)\,dx$

\#

Rmk. If X, Y are indep., $f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(x)\cdot f_Y(z-x)\,dx = f_X * f_Y(z)$

eg. $X\sim\mathcal{N}(0,1)$, $Y\sim\mathcal{N}(0,1)$, indep. Then $X+Y\sim\mathcal{N}(0,2)$

Sol. $f_{X+Y}(z) = \int f_X(x) f_Y(z-x)\,dx$

$$= \int \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}}\,dx$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{4}z^2} \cdot \int \frac{1}{\sqrt{2\pi}} e^{-(x-\frac{z}{2})^2}\,dx$$

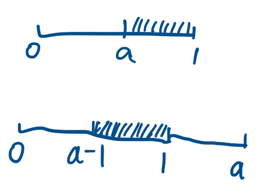
$$= \frac{1}{\sqrt{2\pi}\cdot\sqrt{2}} e^{-\frac{1}{4}z^2} \Rightarrow X+Y\sim\mathcal{N}(0,2)$$

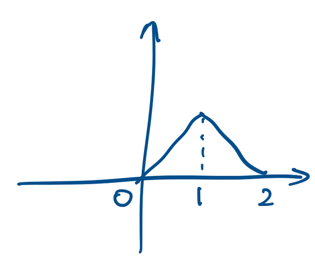
\#

ex. If $X\sim\mathcal{N}(\mu_1,\sigma_1^2)$, $Y\sim\mathcal{N}(\mu_2,\sigma_2^2)$ + indep.: $X+Y\sim\mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

eg. Let X, Y be indep., unif $[0,1]$, what is $f_{X+Y}(a)$

Sol. $f_{X+Y}(a) = \int f_X(x) f_Y(a-x)\,dx = \begin{cases} \int_0^a dx = a & \text{if } a\leq 1 \\ \int_{a-1}^1 dx = 2-a & \text{if } a>1 \end{cases}$





\#

eg. Let $X_1, X_2, \cdots, X_n$ indep. $X_i\sim\text{unif}[0,1]$

① Show that $F_n(x) = \mathbb{P}(X_1+\cdots+X_n\leq x) = \frac{x^n}{n!}$ if $0\leq x\leq 1$

② Let $N = \min\{n : X_1+\cdots+X_n > 1\}$, Compute $\mathbb{E}N$

pf. ① By induction, assume $F_{n-1}(x) = \frac{x^{n-1}}{(n-1)!}$ for $x\in[0,1]$

$$F_n(z) = \int_0^1 f_{X_n}(x)\, F_{n-1}(z-x)\,dx = \int_0^1 1\cdot\frac{(z-x)^{n-1}}{(n-1)!}\cdot 1_{\{x\leq z\}}\,dx = \int_0^z \frac{(z-x)^{n-1}}{(n-1)!}\,dx = \frac{z^n}{n!}$$

② $\mathbb{P}[N>n] = \mathbb{P}[X_1+\cdots+X_n\leq 1] = F_n(1) = \frac{1}{n!}$

$\mathbb{E}N = \sum_{n=0}^{+\infty}\mathbb{P}(N>n) = \sum_{n=0}^{+\infty}\frac{1}{n!} = e$

\#

eg. $\boxed{\text{Bivariate normal}}$:

- Std bivariate normal:

  sps. $X, Y\sim\mathcal{N}(0,1)$, $f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}$, $-1<\rho<1$ $\leftarrow (x-\rho y)^2+(1-\rho^2)y^2$

- Fact: ① $\int f(x,y)\,dx\,dy = 1$ (ex.)

  ② If $\rho=0$, $f(x,y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$, X, Y are indep. $\mathcal{N}(0,1)$

  ③ $\text{Cov}(X,Y) = \rho$

③: $\text{Cov}(X,Y) \triangleq \mathbb{E}XY - \underbrace{\mathbb{E}X\cdot\mathbb{E}Y}_{0} = \int xy\cdot f(x,y)\,dx\,dy$

$$= \int xy\cdot\underbrace{\frac{1}{\sqrt{2\pi}\cdot\sqrt{1-\rho^2}} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}}}_{\mathcal{N}(\rho y,1-\rho^2)}\cdot\underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}}_{\mathcal{N}(0,1)}\,dx\,dy = \int y\cdot\rho y\cdot\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\,dy = \rho$$

\#

- $\boxed{\text{General bivariate normal:}}$ sps. $X\sim\mathcal{N}(\mu_1,\sigma_1^2)$, $Y\sim\mathcal{N}(\mu_2,\sigma_2^2)$

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(x,y)}, \text{ where } Q(x,y) = \frac{1}{1-\rho^2}\cdot\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\cdot\frac{x-\mu_1}{\sigma_1}\cdot\frac{y-\mu_2}{\sigma_1} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right], \ \rho\in(-1,1)$$

ex. • marginal density: $X\sim\mathcal{N}(\mu_1,\sigma_1^2)$, $Y\sim\mathcal{N}(\mu_2,\sigma_2^2)$

• $\boxed{\text{Correlation:}}$ $\rho(x,y) \triangleq \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}X\cdot\text{Var}Y}} = \rho$ ($\rho=0 \Leftrightarrow X, Y$ are independent)

Rmk. ① If X, Y are Gaussians, X, Y are independent $\Leftrightarrow$ X, Y are uncorrelated

② If $X\sim\mathcal{N}(0,1)$, we know $\mathbb{E}[e^{tX}] = e^{\frac{1}{2}t^2} = \sum\frac{1}{k!}\frac{t^{2k}}{2^k}$

$$\mathbb{E}[e^{tX}] = \sum\frac{1}{k!}\cdot t^k\cdot\mathbb{E}X^k$$

Identify the coefficient: $\mathbb{E}X^{2k+1} = 0$

$\mathbb{E}X^{2k} = (2k-1)!! = (2k-1)(2k-3)\cdots 1$

$= \#\{\text{pairings of } \{1,2,\cdots,2k\}\}$

1: $2k-1$ to choose, say 2.

3: $2k-3$ to choose

⋮

Catalan number

③ $\boxed{\text{Wick's Theorem:}}$ (generalized ②)

If $X_1, X_2, \cdots, X_{2n}$ normal, $\mathbb{E}X_i = 0$

Then $\mathbb{E}[X_1X_2\cdots X_{2n}] = \sum_{\substack{\text{pairings of } \{1,2,\cdots,2n\} \\ \triangleq \pi}} \prod_{(i,j)\in\pi} \mathbb{E}[X_iX_j]$

eg. $\mathbb{E}[X_1X_2X_3X_4] = \mathbb{E}[X_1X_2]\mathbb{E}[X_3X_4] + \mathbb{E}[X_1X_3]\mathbb{E}[X_2X_4] + \mathbb{E}[X_1X_4]\mathbb{E}[X_2X_3]$

3/18 HTOP 15

Monday, March 18, 2024 11:14 AM

§. Conditional Distribution & Conditional Expectations

Recall: joint distribution function: $F(x,y) = \mathbb{P}(X \le x ; Y \le y)$

If X, Y are indep., $F(x,y) = F_X(x) \cdot F_Y(y)$

Question: What if X, Y are not independent? $\quad \mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B)$

Def. (discrete R.V.)

The $\boxed{\text{conditional distribution function}}$ of Y given $X=x$ is $F_{Y|X}(y|x) \triangleq \mathbb{P}(Y \le y \mid X=x)$, for every $x \in \mathbb{R}$ s.t. $\mathbb{P}(X=x) > 0$

The $\boxed{\text{cond. mass funct.}}$ $f_{Y|X}(y|x) \triangleq \mathbb{P}(Y=y \mid X=x)$

Rmk.

① $f(x,y) = f_X(x) \cdot f_{Y|X}(y|x)$

② If X, Y are indep. $f_{Y|X}(y|x) = f_Y(y)$, $\forall x \in \mathbb{R}$

Def. For $\forall x$, the $\boxed{\text{conditional expectation}}$ of Y given $X=x$

is $\varphi(x) \triangleq \mathbb{E}[Y|X=x] \triangleq \sum_y y\, f_{Y|X}(y|x)$

The conditional expectation of Y given X

$\mathbb{E}[Y|X] \triangleq \varphi(X)$

Rmk.

$\mathbb{E}[Y|X]$ is a R.V., "best guess" of Y given the info. of X

Prop. [Towering]: $\mathbb{E}\big[\mathbb{E}[Y|X]\big] = \mathbb{E}[Y]$

Pf. $\mathbb{E}\big[\mathbb{E}[Y|X]\big] = \mathbb{E}[\varphi(X)]$

$$= \sum_x \mathbb{E}[Y|X=x]\, f_X(x)$$

$$= \sum_x \sum_y y\, f_{Y|X}(y|x)\, f_X(x)$$

$$= \sum_x \sum_y y\, f(x,y) = \sum_y y\, f_Y(y) = \mathbb{E}Y$$

#

Rmk.

① useful to compute $\mathbb{E}Y$

② Let $\{A_i\}$ be a partition of Ω, then

$\mathbb{E}Y = \sum_i \mathbb{P}(A_i)\, \mathbb{E}[Y|A_i]$

eg. Customer: $N \sim \text{Poisson}(\lambda)$

Each customer $\begin{cases} p, & \text{carry a dog} \\ 1-p, & \text{no dog} \end{cases}$

Dog: K

Compute $\mathbb{E}[K|N]$, $\mathbb{E}K$, $\mathbb{E}[N|K]$

Sol. $\mathbb{E}[K|N=n] = np$, since given $N=n$, $K \sim \text{Binomial}(n,p)$

$\therefore \mathbb{E}[K|N] = p \cdot N$

$\therefore \mathbb{E}[K] = \mathbb{E}\big[\mathbb{E}[K|N]\big] = p \cdot \mathbb{E}N = p\lambda$

$$f_{N|K}(n|k) \underset{\text{Bayes}}{=} \frac{\overbrace{f_{K|N}(k|n)}^{\text{Bin}(n,p)} \cdot \overbrace{f_N(n)}^{\text{poi}(\lambda)}}{\underbrace{f_K(k)}_{\sum_{n \ge k} f(k,n)}} = \frac{\binom{n}{k} p^k (1-p)^{n-k} \cdot \frac{\lambda^n}{n!} e^{-\lambda}}{\sum_{n \ge k} \underbrace{\binom{n}{k}}_{\frac{n!}{k!(n-k)!}} p^k (1-p)^{n-k} \cdot \frac{\lambda^n}{n!} e^{-\lambda}}$$

Bayes: $f_{N|K}(n|k) f_K(k) = f(n,k) = f(k,n) = f_{K|N}(k|n) \cdot f_N(n)$

$$\Rightarrow \frac{p^k \lambda^k}{k!} e^{-p\lambda} \cdot \underbrace{\sum_{n \ge k} \frac{(1-p)^{n-k}}{(n-k)!} \lambda^{n-k} \cdot e^{-(1-p)\lambda}}_{1}$$

$$= \frac{(1-p)^{n-k}}{(n-k)!} \lambda^{n-k} \cdot e^{-(1-p)\lambda}$$

$\therefore$ Given $K=k$, $N-K$ is Poisson$((1-p)\lambda)$

$\mathbb{E}[N|K=k] = \mathbb{E}[N-K|K=k] + k$

$= (1-p)\lambda + k$.

$\therefore \mathbb{E}[N|K] = (1-p)\lambda + K$

#

Examinable Content:

§. Measure theoretic part:

- algebra; $a(\mathcal{E})$; examples from $\mathbb{R}$ (intervals) and discrete sets
- Measure; $\sigma(\mathcal{E})$; Countable additivity; continuity from above / below
  - examples from $\mathbb{R}$ (intervals)
- $\mathcal{B}(\mathbb{R})$, Lebesgue measure
- π-system, examples of π-system that generates $\mathcal{B}(\mathbb{R})$
- Measurable function / Random variables, Distribution $\mathbb{P}_X(\cdot)$, $\sigma(X)$
- equivalent cond. for a funct. being Borel (random variable)
- Construction of Lebesgue integral / Expectation, Monotone Convergence Thm.
- Indep. for σ-algebras, random variables, events
- Sufficient (Equivalent) cond. for independence

§ Non-measure part.

- Chebyshev Ineq.
- Moment generating funct. for Gaussians
- Inclusion / Exclusion
- Cond. Prob. / Bayes
  - Gambler's ruins; Cond. on the 1^{st} step $\to$ recursion
- Discrete / Conti. R.V.s
- Expectation: Change of variables. $\mathbb{E}X = \begin{cases} \sum \mathbb{P}(X \ge n) \\ \int \mathbb{P}(X > x)\, dx \end{cases}$
- Variance & Moments.
  - Covariance, uncorrelated, correlation
- Joint distri. funct. / density funct. / marginal distri. Computations
  - Sum of R.V.s, bivariate Gaussian

Discrete R.V.

Bernoulli(p) (indep. trials): $f(1) = p$, $f(0) = q = 1-p$ $\quad \begin{cases} \mathbb{E}: p \\ Var: pq \end{cases}$

Binomial(n,p) (k success in n indep. trials): $f(k) = C_n^k p^k q^{n-k}$ $\quad \begin{cases} \mathbb{E}: np \\ Var: npq \end{cases}$

Poisson(λ) ($= \lim_{n \to +\infty} \text{Bin}(n, \frac{\lambda}{n})$): $f(k) = \frac{\lambda^k}{k!} e^{-\lambda}$ $\quad \begin{cases} \mathbb{E}: \lambda \\ Var: \lambda \end{cases}$

Geometric(p) (first success in indep. trials): $f(k) = p q^{k-1}$ $\quad \begin{cases} \mathbb{E}: \frac{1}{p} \\ Var: \frac{q}{p^2} \end{cases}$

Cont. R.V.

Exponential(λ): $f(x) = \lambda e^{-\lambda x}$ $(x > 0)$ $\quad \begin{cases} \mathbb{E}: \frac{1}{\lambda} \\ Var: \frac{1}{\lambda^2} \end{cases}$

Normal(u, σ^2): $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-u)^2}{2\sigma^2}}$ $\quad \begin{cases} \mathbb{E}: u \\ Var: \sigma^2 \end{cases}$

Uniform$[a,b]$: $f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & \text{elsewhere} \end{cases}$ $\quad \begin{cases} \mathbb{E}: \frac{a+b}{2} \\ Var: \frac{(b-a)^2}{12} \end{cases}$

3/25 HTOP 16

Monday, March 25, 2024 11:16 AM

Recall: Discrete R.V. Conditional mass function

$f_{Y|X}(y|x) \triangleq \mathbb{P}(Y=y \mid X=x)$

$\varphi(X) \triangleq \mathbb{E}[Y|X]$ (Lec. 15)

$\varphi(x) \triangleq \mathbb{E}[Y|X=x] = \sum_y y \cdot f_{Y|X}(y|x)$

eg. X, Y indep. $X \sim Poi(\lambda_1)$, $Y \sim Poi(\lambda_2)$

What is the conditional distribution of X given $X+Y=n$?

Sol. $\mathbb{P}[X=k \mid X+Y=n] = \dfrac{\mathbb{P}[X=k, Y=n-k]}{\mathbb{P}[X+Y=n]}$

X, Y indep.
$X+Y \sim Poi(\lambda_1+\lambda_2)$

$$= \frac{\lambda_1^k}{k!} e^{-\lambda_1} \cdot \frac{\lambda_2^{n-k}}{(n-k)!} e^{-\lambda_2} \cdot \left(\frac{(\lambda_1+\lambda_2)^n}{n!} e^{-(\lambda_1+\lambda_2)} \right)^{-1}$$

$$= C_n^k \cdot \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \cdot \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}$$

Binomail $(n, \frac{\lambda_1}{\lambda_1+\lambda_2})$

#

Conditional distri. for continuous R.V. s

Heuristic: $\mathbb{P}\left[Y \le y \mid X \in [x, x+dx]\right] = \dfrac{\mathbb{P}[Y \le y, X \in [x, x+dx]]}{\mathbb{P}[X \in [x, x+dx]]}$

$$\approx \frac{\int_{-\infty}^{y} f(x,v)dv \cdot dx}{f_X(x) \cdot dx}$$

Def. $\boxed{\text{The Conditional Distribution of } Y \text{ given } X > x}$ [Conti.]

$$F_{Y|X}(y|x) \triangleq \frac{\int_{-\infty}^{y} f(x,v)dv}{f_X(x)}, \quad \text{for } \forall x \text{ s.t. } f_X(x) > 0$$

[Conti.]

Def. $\boxed{\text{The conditional density funct.}}$: $f_{Y|X}(y|x) \triangleq \dfrac{f(x,y)}{f_X(x)} = \dfrac{f(x,y)}{\int_{-\infty}^{+\infty} f(x,y)dy}$

[Conti.]

Def. $\boxed{\text{Conditional Expectation}}$: $\mathbb{E}[Y|X] = \varphi(X)$

where $\varphi(x) \triangleq \mathbb{E}[Y|X=x] = \int y f_{Y|X}(y|x) dy$

Thm. (Towering Prop.)

$$\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}Y = \int f_X(x) \cdot \mathbb{E}[Y|X=x]\, dx$$

In particular, if $Y = 1_A$, then $\mathbb{P}(A) = \int f_X(x) \cdot \mathbb{P}[A|X=x]dx$

pf. (ex.)

eg. Sps. the joint density function of X, Y

$$f(x,y) = \begin{cases} \frac{1}{y} e^{-\frac{x}{y}} e^{-y}, & x, y \in \mathbb{R}^+ \\ 0, & \text{else} \end{cases}$$

Find $\mathbb{P}[X>1 \mid Y=y]$

Sol. $f_{X|Y}(x|y) = \dfrac{f(x,y)}{f_Y(y)} = \dfrac{f(x,y)}{\int_{-\infty}^{+\infty} f(x,y)dx} = \dfrac{\frac{1}{y} e^{-\frac{x}{y}} \cdot e^{-y}}{\frac{e^{-y}}{y} \int_0^{+\infty} e^{-\frac{x}{y}} dx} = \frac{1}{y} e^{-\frac{x}{y}}$

$$\Rightarrow \mathbb{P}[X>1 \mid Y=y] = \int_1^{+\infty} \frac{1}{y} e^{-\frac{x}{y}} dx = -e^{-\frac{x}{y}} \Big|_{x=1}^{x=+\infty} = e^{-\frac{1}{y}}$$

#

eg. $\boxed{\text{Bivariate Normal}}$ (Lec. 14)

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \cdot \underbrace{(x^2 - 2\rho xy + y^2)}_{(y-\rho x)^2 + (1-\rho^2)x^2}}$$

Compute $\mathbb{E}[Y|X]$

Sol. $f(x,y) = \underbrace{\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}}_{\mathcal{N}(0,1) \,=\, f_X(x)} \cdot \underbrace{\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}}}_{\mathcal{N}(\rho x, 1-\rho^2) \,=\, f_{Y|X}(y|x)}$

$\therefore \mathbb{E}[Y|X=x] = \rho x$, $\mathbb{E}[Y|X] = \rho X$

$\mathbb{E}[Y|X=x] = \int y \cdot f_{Y|X}(y|x)$

#

• $\boxed{\text{Multiple Conditioning:}}$

$\mathbb{E}[X \mid Y, Z] \triangleq \varphi(Y, Z)$

s.t. $\varphi(y,z) = \mathbb{E}[X \mid Y=y, Z=z] = \begin{cases} \sum_x x \cdot \mathbb{P}[X=x \mid Y=y, Z=z], & \text{discrete} \\ \dfrac{\int x \cdot f(x,y,z)dx}{\underbrace{f_{Y,Z}(y,z)}_{\int f(x,y,z)dx}}, & \text{continuous} \end{cases}$

m'ble funct.

Thm. $\mathbb{E}[\mathbb{E}[Y|X] \cdot g(X)] = \mathbb{E}[Y \cdot g(X)]$ since $\mathbb{E}\Big[\mathbb{E}[Y \cdot g(X) | X]\Big] \overset{\text{pull-out property}}{=} \mathbb{E}(g(X) \cdot \mathbb{E}[Y|X]) \overset{\text{Towering property}}{=} \mathbb{E}[Y \cdot g(X)]$

$\boxed{\text{for multiple conditioning ?}}$ ①

Measure Theoretical Definition for Conditional Expectation:

Def. $(\Omega, \mathcal{F}, \mathbb{P})$ prob. space, let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra, the cond. exp. $\mathbb{E}[X|\mathcal{G}]$ is a real R.V. s.t.

① $Y_{\mathcal{G}} \triangleq \mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-m'ble
$(\forall B \in \mathcal{B}(\mathbb{R}),\ Y_{\mathcal{G}}^{-1}(B) \in \mathcal{G})$

② for $\forall A \in \mathcal{G}$, we have: $\int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_A X\, d\mathbb{P}$

Answer for ①:

General Towering: if $\mathcal{F} \subseteq \mathcal{G}$, then $\mathbb{E}\big[\mathbb{E}[Y|\mathcal{G}] \mid \mathcal{F}\big] = \mathbb{E}[Y|\mathcal{F}]$

for instance, $\mathcal{F} = \sigma(X_1, \cdots, X_n)$

We define $\mathbb{E}[X|Y] \triangleq \mathbb{E}[X|\sigma(Y)]$

"best estimate for X given info. of $\mathcal{G}$"

• If $\mathcal{G} = \mathcal{P}(\Omega)$, Then $\mathbb{E}[X|\mathcal{G}] = X$ almost surely.

• If $\mathcal{G} = \{\phi, \Omega\}$, Then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}X$

Random Walk revisited:

$S_n = S_0 + \sum_{i=1}^{n} X_i$, (X_i) i.i.d.

$d=1$:  $\mathbb{P}(X_i = 1) = p$, $\mathbb{P}(X_i = -1) = q$

If $p = q = \frac{1}{2}$, symmetric simple random walk.

More generally, many results under the assumption: (X_i) i.i.d., $\mathbb{E}X_i^2 < +\infty$

For now, assume $S_n \in \mathbb{Z}^d$

Lemma. $\boxed{\text{(Spatial Homogeniety)}}$

$$\mathbb{P}[S_n = j \mid S_0 = a] = \mathbb{P}[S_n = j-a \mid S_0 = 0] \quad \left(= \mathbb{P}\Big[\sum_{i=1}^{n} X_i = j-a\Big] \right)$$

Lemma. $\boxed{\text{(Time Homogeniety)}}$

$$\mathbb{P}[S_n = j \mid S_0 = a] = \mathbb{P}[S_{n+m} = j \mid S_m = a]$$

$$LHS = \mathbb{P}\Big[\sum_{i=1}^{n} X_i = j-a\Big] = \mathbb{P}\Big[\sum_{i=m+1}^{m+n} X_i = j-a\Big] = RHS$$

$\boxed{\text{Markov Property:}}$ $\mathbb{P}[S_{n+m} = j \mid S_0, S_1, \cdots, S_m]$
$= \mathbb{P}[S_{n+m} = j \mid S_m]$

" Position after m^{th} step does not depend on info. before m "

Recurrence & Transience

Let $T_y^0 = 0$, for $k \ge 1$, let $T_y^k = k^{th}$ hitting time of y ← a position

$$T_y^k \triangleq \inf\{t : t > T_y^{k-1} \ \&\ S_t = y\}$$

Def. y is called $\boxed{\text{recurrent}}$ if $\mathbb{P}_y[T_y^1 < +\infty] = 1$
$\boxed{\text{transient}}$ if $\mathbb{P}_y[T_y^1 < +\infty] < 1$

Rmk. ① If y is recurrent, by Markov. Prop.

$\forall k \in \mathbb{N}$, $\mathbb{P}[T_y^k < +\infty] = 1$

$\Rightarrow \mathbb{P}_y[S_n = y \text{ infinitely often}] = \mathbb{P}_y[\limsup_{n \to +\infty} \{S_n = y\}] = 1$ (ex.)

② If y is transient

Let $N_y = \sum_{n=1}^{+\infty} 1_{\{S_n = y\}}$, # of visits at y

one can sum up to k times hitting $y \Leftrightarrow$ the k-th time of hitting y can be reached.

$$\mathbb{E}N_y = \sum_k \mathbb{P}[N_y \ge k] = \sum_{k=1}^{+\infty} \mathbb{P}_y[T_y^k < +\infty]$$

$$\overset{\text{Markov Property}}{=} \sum_{k=1}^{+\infty} \mathbb{P}_y[T_y^1 < +\infty]^k$$

$$\overset{\text{by def. of transient}}{=} \frac{\mathbb{P}_y[T_y^1 < +\infty]}{1 - \mathbb{P}_y[T_y^1 < +\infty]} < +\infty$$

3/27 HTOP 17

Wednesday, March 27, 2024 11:17 AM

Recall : 1-d symmetric SRW starting at y

$\mathbb{E}N_y$ = mean # of returns at y

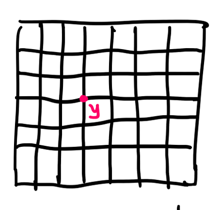
$$\begin{cases} \mathbb{E}N_x = \mathbb{P}(X_1=1)\cdot\mathbb{E}[N_x|X_1=1] + \mathbb{P}(X_1=-1)\cdot\mathbb{E}[N_x|X_1=-1] + \delta_y \\ \quad\;\; = \frac{1}{2}\mathbb{E}N_{x-1} + \frac{1}{2}\mathbb{E}N_{x+1} + \delta_y \quad \text{where } \delta_y(x) = \begin{cases} 1 & \text{if } y=x \\ 0 & \text{if } y\neq x \end{cases} \text{ (counting the start)} \\ \mathbb{E}N_0 = 0,\ \mathbb{E}N_N = 0 \quad \text{(bdy cond.)} \end{cases}$$

By solving the recurrence, $\mathbb{E}N_y \sim cN$

$$\Rightarrow \mathbb{E}N_y \to +\infty \text{ as } N\to+\infty \quad \text{(recurrent)}$$

$$\boxed{\text{Thm. } y \text{ is a recurrent state (or transient state) iff } \mathbb{E}N_y = +\infty \ (\text{or } \mathbb{E}N_y < +\infty)}$$

- SRW on $\mathbb{Z}^d$, $d\geq 2$



$[-N,N]\cap\mathbb{Z}^d$

$G(x) \triangleq \mathbb{E}N_x$ = mean # of returns at x before hitting the boundary

$$\begin{cases} G(x) = \frac{1}{2d}\sum_{z\sim x} G(z) + \delta_y \\ G(z) = 0, \quad \text{if } z\in\partial\square_N \end{cases}$$

Let $\Delta f(x) = \frac{1}{2d}\sum_{z\sim x}(f(z)-f(x))$

Then the recurrence can be written as :

$$\begin{cases} \Delta G(x) = -\delta_y(x) \\ G(z) = 0 \quad \text{if } z\in\partial\square_N \end{cases} \qquad \boxed{\text{discrete Green's funct.}}$$

Claim :

Set $y=0$. $\quad G(0) \sim \begin{cases} \log N, & d=2 \\ O(1), & d\geq 3 \end{cases}$ Rmk. recurrent in $d=2$ transcience in $d\geq 3$

Recall : Solve PDE by Fourier Transform

$$\Delta G(x) = -\delta_0(x) \quad \text{in } \mathbb{R}^d$$

$$\hat{G}(k) = \int_{\mathbb{R}^d} e^{ikx} G(x)\,dx$$

$$\widehat{\Delta G}(k) = (k_1^2+\cdots+k_d^2)\cdot\hat{G}(k)$$

$$= \int e^{ikx}\cdot\delta_0(x)\,dx = 1$$

$$\Rightarrow \hat{G}(k) = \frac{1}{|k|^2}$$

On $\mathbb{Z}^d$: $f_k(x) = e^{ikx}$ are eigenfunctions of Δ

with $\Delta f_k = \lambda_k f_k$, $\lambda_k = -\sum_{j=1}^{d}(1-\cos k_j)$, $k=(k_1,\cdots,k_d)$

When $d=1$, $\Delta f_k(x) = \frac{1}{2}\cdot f_k(x+1) + \frac{1}{2}f_k(x-1) - f_k(x)$

$$= f_k(x)\cdot\left(\tfrac{1}{2}e^{ik} + \tfrac{1}{2}e^{-ik} - 1\right)$$

$$= f_k(x)\cdot(\cos k - 1)$$

Fourier Transform :

$$\hat{G}(k) = \sum_{x\in\square_N} G(x)\cdot e^{ik\cdot x}$$

Inverse Fourier Transform :

$$G(x) = \frac{1}{|\square_N|}\sum_{k\in\square_N^*} e^{-ik\cdot x}\hat{G}(k)$$

$$\square_y^* \triangleq \left\{\frac{2\pi}{N}(n_1,\cdots,n_d),\ n_i\in[-N,N]\cap\mathbb{Z}\right\}$$

$$\Delta G(x) = \delta_0$$

Fourier Transfer : $\Delta\cdot\sum_{x\in\square_y} G(x)\cdot e^{ik\cdot x} = \lambda_k\cdot\hat{G}(k) = 1$

$$\hat{G}(k) = \frac{1}{\lambda_k}$$

using inverse : $G(0) = \frac{1}{|\square_N|}\sum_{k\in\square_N^*}\hat{G}(k)$

$$= \frac{1}{|\square_y|}\cdot\sum_{k\in\square_y^*}\frac{1}{\sum_{j=1}^{d}(1-\cos k_j)}$$

$$\xrightarrow{N\nearrow+\infty} \frac{1}{(2\pi)^d}\cdot\int_{[-\pi,\pi]^d}\underbrace{\frac{1}{\sum_{j=1}^{d}(1-\cos x_j)}}_{\frac{1}{|x|^2} \text{ for } x\approx 0}dx$$

In fact

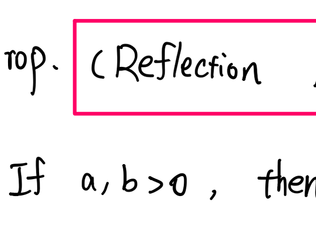
Hence, this converges for $d\geq 3$
diverges for $d=2$

#

- Random Walk : Counting the paths

SRW on $\mathbb{Z}$, starting at a

$\mathbb{P}(S_0=a, S_n=b) = \sum_r M_n^r(a,b)\,p^r q^{n-r}$ (reach b at time n)



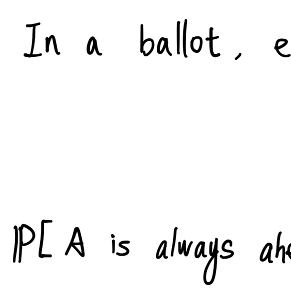
where $M_n^r(a,b) \triangleq$ # paths with $S_0=a$, $S_n=b$ that makes r right moves

As $r-(n-r) = b-a \Rightarrow r = \frac{1}{2}(n+b-a)$, we have $M_n^r(a,b) = C_n^{\frac{1}{2}(n+b-a)}$

$$\therefore \mathbb{P}(S_0=a, S_n=b) = C_n^{\frac{1}{2}(n+b-a)}\cdot p^r\cdot q^{n-r}$$

Prop. $\boxed{\text{(Reflection Principle)}}$

If $a,b>0$, then $N_0^n(a,b) = N^n(-a,b)$



$N_0^n(a,b)$: # paths : $a\to b$ that visits 0

$N^n(-a,b)$: # paths : $-a\to b$ taking n steps

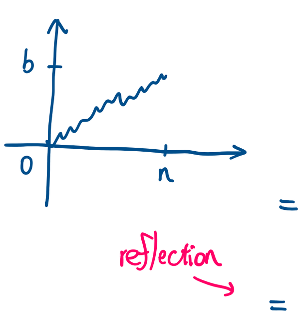
$T \triangleq \inf\{n : S_n = 0\}$

$\boxed{\text{· Bertrand's Ballot Theorem :}}$

In a ballot, eventually A has a votes
B has b votes , $a>b$

$$\mathbb{P}[A \text{ is always ahead of } B \mid a>b] = \frac{a-b}{a+b}$$

Indeed, $S_i \triangleq$ # votes for A at time i − # votes for B at time i



By the following thm.,

$$\mathbb{P}[A \text{ is always ahead of } B \mid a>b] = \frac{a-b}{a+b}\cdot 1 = \frac{a-b}{a+b}$$

Thm. Let S be a SRW on $\mathbb{Z}$, $S_0=0$, then $\mathbb{P}(S_n=b, S_i\neq 0, i=1,\cdots,n) = \frac{|b|}{n}\mathbb{P}(S_n=b)$

Pf.

Assume $b>0$.

paths $(0,0)\to(n,b)$ that does not visit 0

$$= N^{n-1}(1,b) - N_0^{n-1}(1,b)$$

$$\overset{\text{reflection}}{=} N^{n-1}(1,b) - N^{n-1}(-1,b)$$

$$= C_{n-1}^{\frac{1}{2}(n-1+b-1)} - C_{n-1}^{\frac{1}{2}(n-1+b+1)}$$

$$= \binom{n}{\frac{1}{2}(n+b)}\cdot\left(\frac{\frac{1}{2}(n+b)}{n} - \frac{\frac{1}{2}(n-b)}{n}\right)$$

$$= \frac{b}{n}\cdot\binom{n}{\frac{1}{2}(n+b)} = \frac{b}{n}N^n(0,b)$$

$$\Rightarrow \mathbb{P}(S_n=b,\ S_i\neq 0 \text{ for } i=1,\cdots,n) = \frac{b}{n}\mathbb{P}(S_n=b)$$

#

Cor. (Summing over b)

$$\mathbb{P}(S_i\neq 0, i=1,\cdots,n) = \frac{1}{n}\mathbb{E}|S_n| \quad \boxed{\overset{①}{\leq}\frac{1}{n}(\operatorname{Var}S_n)^{\frac{1}{2}} \overset{②}{=} \frac{1}{\sqrt{n}}}$$

For the case of symmetric SRW on $\mathbb{Z}$

Since $\mathbb{E}S_n = \sum_{k=1}^{n}\underbrace{\mathbb{E}X_k}_{0} = 0$,

then $\mathbb{E}|S_n| \overset{C\text{-}S}{\leq} \sqrt{\mathbb{E}S_n^2} = \sqrt{\operatorname{Var}S_n}$ ①

also, $\operatorname{Var}S_n = \operatorname{Var}\sum_{k=1}^{n}X_k$

$$= \mathbb{E}\Big(\sum_{k=1}^{n}X_k\Big)^2 - \Big(\sum_{k=1}^{n}\mathbb{E}X_k\Big)^2$$

$$= \sum_{k=1}^{n}\mathbb{E}X_k^2 + \sum_{i\neq j}\mathbb{E}X_iX_j - \sum_{k=1}^{n}(\mathbb{E}X_k)^2 - \sum_{i\neq j}\mathbb{E}X_i\cdot\mathbb{E}X_j$$

$$\overset{\text{i.i.d.}}{=} \sum_{k=1}^{n}\operatorname{Var}X_k = \sum_{k=1}^{n}\underbrace{\mathbb{E}X_k^2}_{1} = n \quad ②$$

Pf. $\mathbb{P}[S_i\neq 0, i=1,\cdots,n] = \sum_{b=-\infty}^{-1}\frac{|b|}{n}\mathbb{P}(S_n=b) + \sum_{b=1}^{+\infty}\frac{|b|}{n}\mathbb{P}(S_n=b)$

$$= 2\sum_{b=1}^{+\infty}\frac{b}{n}\mathbb{P}(S_n=b)$$

$$= \frac{2}{n}\mathbb{E}S_n$$

$$= \frac{1}{n}\mathbb{E}|S_n|$$

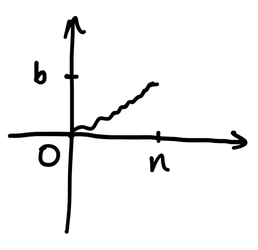
#

4/8 HTOP 18

Monday, April 8, 2024 11:20 AM

Final: May 8th, In-class

Recall: Ballot Thm.



$\mathbb{P}[S_n = b,\ S_i \neq 0,\ i=1,\cdots,n] = \frac{|b|}{n}\cdot \mathbb{P}[S_n = b]$

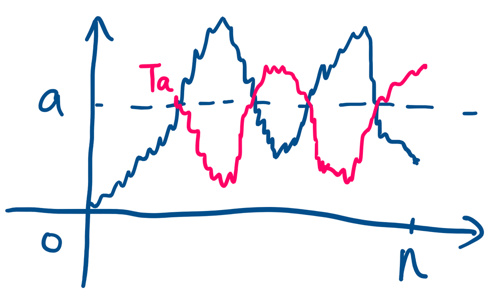
$\mathbb{P}[S_i \neq 0,\ i=1,\cdots,n] = \frac{1}{n}\mathbb{E}|S_n|$

Thm. Sps. $S_0 = 0$, $p=q=\frac{1}{2}$. Then $\mathbb{P}[\max_{1\le k\le n} S_k \ge a] = \mathbb{P}[S_n \ge a] + \mathbb{P}[S_n \ge a+1]$ $(a\ge 0)$

Pf. $LHS = \underbrace{\mathbb{P}[\max_{1\le k\le n} S_k \ge a\ ;\ S_n \ge a]}_{S_n \ge a} + \mathbb{P}[\max_{1\le k\le n} S_k \ge a\ ;\ S_n < a]$

Let $T_a = \inf\{n\ge 1 : S_n = a\}$

Define the reflected walk. $\tilde{S}_n = \begin{cases} S_n, & n \le T_a \\ \underbrace{2a - S_n}, & n > T_a \end{cases}$



has the same distri. of SRW starting at a

By Markov property, $\tilde{S}_n$ has the same distri. as the SRW.

$\therefore \mathbb{P}[\max_{1\le k\le n} S_k \ge a\ ;\ S_n < a] = \mathbb{P}[\tilde{S}_n > a] = \mathbb{P}[S_n > a] = \mathbb{P}[S_n \ge a+1]$

$a\ge 0$

#

Rmk. • We say that $(W_t)_{t\ge 0}$ is a Brownian motion if

① $W_0 = 0$

② For any $t>s>0$, $W_t - W_s \sim \mathcal{N}(0, t-s)$

③ $W_{t_2} - W_{t_1}$, $W_{t_3} - W_{t_2}$, $\cdots$, $W_{t_n} - W_{t_{n-1}}$ are indep.

④ $t \mapsto W_t$ is continuous

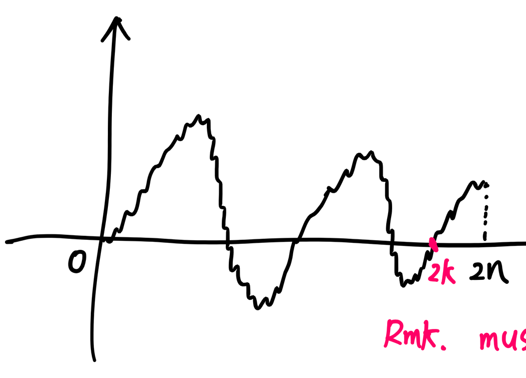
• For Brownian Motion,

$\mathbb{P}[\max_{t\in[0,T]} W_t \underset{(\ge)}{>} a] = 2\mathbb{P}[W_t \underset{(\ge)}{>} a]$

Thm. (Last visit to 0), 1-d SRW, $p=q=\frac{1}{2}$

$\mathbb{P}[\text{last visit to } 0 \text{ in } [0,2n] \text{ is } 2k]$, $k\in\{1,\cdots,n\}$

$= \mathbb{P}[S_{2k} = 0]\cdot \mathbb{P}[S_{2n-2k} = 0]$



Rmk. must be even steps

Pf. $LHS = \mathbb{P}[S_{2k} = 0]\cdot \mathbb{P}[S_i \neq 0,\ i = 2k+1,\cdots,2n \mid S_{2k} = 0]$

By Ballot Thm. $\mathbb{P}[S_i \neq 0,\ i=1,\cdots,2m] = \frac{1}{2m}\mathbb{E}|S_{2m}|$

$= \frac{1}{2m}\cdot 2\cdot \sum_{k=1}^{+\infty} 2k\cdot \mathbb{P}[S_{2m} = 2k]$

$= \frac{1}{m}\cdot \sum_{k=1}^{+\infty} 2k\cdot \binom{2m}{m+k}\cdot\left(\frac{1}{2}\right)^{2m}$

$= 2\sum_{k=1}^{+\infty}\left(\frac{1}{2}\right)^{2m}\cdot\left[\binom{2m-1}{m+k-1} - \binom{2m-1}{m+k}\right]$ By using: $\frac{2k}{2m}\binom{2m}{m+k} = \binom{2m-1}{m+k-1} - \binom{2m-1}{m+k}$

$= 2\cdot\left(\frac{1}{2}\right)^{2m}\cdot\binom{2m-1}{m}$

$= \mathbb{P}[S_{2m} = 0]$

By Markov Property (Time Homogeniety), take $m = n-k$, we complete the proof.

#

• $\boxed{\text{Stirling Formula}}$: $n! \overset{n\to+\infty}{\sim} \left(\frac{n}{e}\right)^n\cdot\sqrt{2\pi n}$

$\mathbb{P}[S_{2k} = 0] = \left(\frac{1}{2}\right)^{2k}\cdot\binom{2k}{k} \sim \frac{1}{\sqrt{\pi k}}$ as $k \nearrow +\infty$

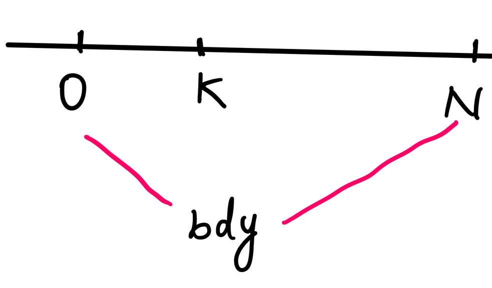
$\Rightarrow \mathbb{P}[\text{Last visit to } 0 \text{ in } [0,2n] \le 2n\cdot x] = \sum_{k\le nx}\mathbb{P}[S_{2k}=0]\,\mathbb{P}[S_{2n-2k}=0]$

$\overset{n\to+\infty}{\sim} \sum_{k\le nx}\frac{1}{\sqrt{\pi k}}\cdot\frac{1}{\sqrt{\pi(n-k)}}$

$\approx \int_0^x \frac{1}{\pi\sqrt{u(1-u)}}\,du = \frac{2}{\pi}\arcsin\sqrt{x}$, $x\in[0,1]$

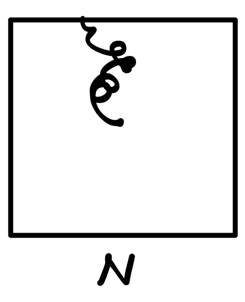
$\boxed{\text{Arcsine Law of SRW}}$

① 1-d SRW: $p=q=\frac{1}{2}$



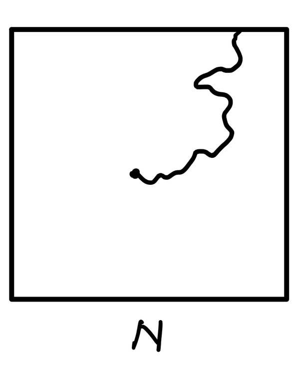
$\mathbb{E}[\#\text{ steps before hitting } 0, N] = k\cdot(N-k) \sim O(N^2)$ (ex.)

② SRW in $\mathbb{Z}^d$, $d\ge 2$



$\mathbb{E}[\#\text{ steps before hitting bdy}] = O(N^2)$

③ Loop erased RW



(d=2): $\mathbb{E}[\#\text{length of LERW}] = O(N^{\frac{5}{4}})$

(d=3): $\mathbb{E}[\#\text{length of LERW}] \approx N^{\alpha + O_N(1)}$

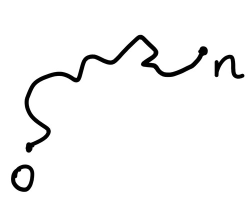
$\alpha = ?$ $\frac{5}{4} \le \alpha < 2$

Lawler–Schramm–Werner' 2002

Fields' 2006

Kozma' 2008

④ Self-awarding Walks

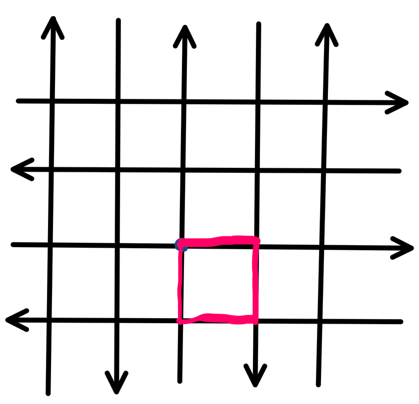


Unif measure

(d=2): Conjecture: $\sqrt{Var[SAW_n]} \approx O(N^{\frac{3}{4}})$

猜想

⑤ Manhatten Walks



p-scatter

Conjecture: $\mathbb{P}[\ \bigcirc\ \overset{\uparrow}{\underset{\downarrow}{L}}\ ] \le e^{-C_p L}$

$\forall p > 0$

where C_p is a constant depending on p

4/10 HTOP 19

Wednesday, April 10, 2024 11:17 AM

- Funct. of R.V.
- Generating funct., Characteristic funct.
  - ↳ Branching process
- Convergence of R.V.'s
- Law of large numbers / Central limit theorem.

§ Function of R.V.

eg. Let $X \sim \mathcal{N}(0,1)$, $Y = X^2$

Find the prob. density funct. of Y

Sol. $F_Y(y) = \mathbb{P}(Y \le y) = \begin{cases} \mathbb{P}(-\sqrt{y} \le X \le \sqrt{y}), & \text{if } y \ge 0 \\ 0, & \text{if } y < 0. \end{cases}$

Let $\Phi(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

Then $\mathbb{P}[-\sqrt{y} \le X \le \sqrt{y}] = \Phi(\sqrt{y}) - \Phi(-\sqrt{y})$

$= 2\Phi(\sqrt{y}) - 1$

$\therefore f_Y(y) = \Phi'(\sqrt{y}) \cdot \frac{1}{\sqrt{y}} = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \; y \ge 0$

#

Recall: If $T: \mathbb{R}^n \to T(\mathbb{R}^n) \subseteq \mathbb{R}^n$ is bijective

$(x_1, \cdots, x_n) \mapsto (y_1, \cdots, y_n)$, $A \subseteq \mathbb{R}^n$

Then $\int_A g(x_1, \cdots, x_n)\, dx_1 \cdots dx_n = \int_{T(A)} g\big(x_1(y_1, \cdots, y_n), \cdots, x_n(y_1, \cdots, y_n)\big) \cdot |J| \cdot dy_1 \cdots dy_n$

where $J = \det\left(\frac{\partial x_i}{\partial y_j}\right)_{i,j = 1, \cdots, n}$

Cor. If $(X_1, \cdots, X_n)$ has joint density function f. Then $(Y_1, \cdots, Y_n) = T(X_1, \cdots, X_n)$

has joint density funct. $f_{Y_1, \cdots, Y_n}(y_1, \cdots, y_n) = f\big(x_1(y_1, \cdots, y_n), \cdots, x_n(y_1, \cdots, y_n)\big) \cdot |J|$ if $(y_1, \cdots, y_n) \in Range(T)$

Pf. $\mathbb{P}\big((y_1, \cdots, y_n) \in T(A)\big) = \mathbb{P}\big((x_1, \cdots, x_n) \in A\big)$

$\int_{T(A)} f_{Y_1, \cdots, Y_n}(y_1, \cdots, y_n)\, dy_1 \cdots dy_n = \int_A f(x_1, \cdots, x_n)\, dx_1 \cdots dx_n$

$= \int_{T(A)} f\big(x_1(y_1, \cdots, y_n), \cdots, x_n(y_1, \cdots, y_n)\big)\, |J|\, dy_1 \cdots dy_n$

#

eg. Let X_1, X_2 have the joint density funct. f. $X_1 = aY_1 + bY_2$, $X_2 = cY_1 + dY_2$, $ad \ne bc$

Find $f_{Y_1, Y_2}(y_1, y_2)$

Sol. $|J| = \left\| \begin{matrix} \frac{\partial X_1}{\partial Y_1} & \frac{\partial X_1}{\partial Y_2} \\ \frac{\partial X_2}{\partial Y_1} & \frac{\partial X_2}{\partial Y_2} \end{matrix} \right\| = \left\| \begin{matrix} a & b \\ c & d \end{matrix} \right\| = |ad - bc|$

$f_{Y_1, Y_2}(y_1, y_2) = f(ay_1 + by_2, cy_1 + dy_2) \cdot |ad - bc|$

#

eg. Sps. X, Y have joint density funct. Show that $U = XY$

has density funct. $f_U(u) = \int_{-\infty}^{+\infty} f(x, \frac{u}{x}) \cdot \frac{1}{|x|} dx$

Sol. Let $\begin{cases} U = XY \\ V = X \end{cases}$, then $\begin{cases} X = V \\ Y = \frac{U}{V} \end{cases}$, then $|J| = \left|\det\begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{pmatrix}\right| = \left\| \begin{matrix} 0 & 1 \\ \frac{1}{v} & -\frac{u}{v^2} \end{matrix} \right\| = \frac{1}{|v|}$

Then $f_{U,V}(u, v) = f(v, \frac{u}{v}) \cdot \frac{1}{|v|}$, so that $f_U(u) = \int_{-\infty}^{+\infty} f(v, \frac{u}{v}) \cdot \frac{1}{|v|} dv$

#

eg. Let X_1, X_2 be indep. $Exp(\lambda)$. Find the joint density funct. of $Y_1 = X_1 + X_2$, $Y_2 = \frac{X_1}{X_2}$

Sol. $\begin{cases} Y_1 = X_1 + X_2 \\ Y_2 = \frac{X_1}{X_2} \end{cases}$, then $\begin{cases} X_1 = \frac{Y_1 Y_2}{1 + Y_2} \\ X_2 = \frac{Y_1}{1 + Y_2} \end{cases}$. $|J| = \left\| \begin{matrix} \frac{y_2}{1+y_2} & \frac{y_1}{(1+y_2)^2} \\ \frac{1}{1+y_2} & -\frac{y_1}{(1+y_2)^2} \end{matrix} \right\| = \left| \frac{y_1(1+y_2)}{(1+y_2)^3} \right| = \frac{|y_1|}{(1+y_2)^2}$

$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}\left(\frac{y_1 y_2}{1 + y_2}, \frac{y_1}{1 + y_2}\right) \cdot \frac{|y_1|}{(1+y_2)^2}$

$= \underbrace{\lambda^2 e^{-\lambda y_1} \cdot |y_1|}_{f_{Y_1}} \cdot \underbrace{\frac{1}{(1+y_2)^2}}_{f_{Y_2}}$

since it can be factorized into terms involving only y_1 & y_2

Also, we can verify by directly calculating from $Y_1 = X_1 + X_2$, $Y_2 = \frac{X_1}{X_2}$

#

(ex.) Let $X_1, X_2 \sim \mathcal{N}(0,1)$ indep. Then $X_1 + X_2$ and $X_1 - X_2$ are indep.

eg. Let X, Y be indep. $\mathcal{N}(0,1)$. Find the joint density funct. of $R = \sqrt{X^2 + Y^2}$

$\Theta = \tan^{-1}\left(\frac{Y}{X}\right)$

Sol. $\begin{cases} X = R\cos\Theta \\ Y = R\sin\Theta \end{cases}$, $|J| = \left\| \begin{matrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{matrix} \right\| = r$

$f_{R,\Theta} = f_{X,Y}(r\cos\theta, r\sin\theta)\, r$ 　 $f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2 + y^2}{2}}$ (check)

$= \frac{1}{2\pi} e^{-\frac{r^2}{2}} r$

$\Rightarrow \Theta \sim Unif[0, 2\pi)$ & $f_R(r) = re^{-\frac{r^2}{2}}$ $\boxed{\text{Rayleigh distribution}}$

#

(ex.) Show that R^2, Θ are indep. and $R^2 \sim Exp(\frac{1}{2})$

Sol. Indep. is shown by separating variables.

Let $L = R^2$, $f_R(r) = f_L(l) \cdot \left|\frac{dl}{dr}\right| = f_L(l) \cdot 2r \Rightarrow f_L(l) = \frac{1}{2r} f_R(r) = \frac{1}{2\sqrt{l}} f_R(\sqrt{l}) = \frac{1}{2\sqrt{l}} \cdot \sqrt{l} \cdot e^{-\frac{1}{2}l} = \frac{1}{2} e^{-\frac{1}{2}l}$

$\therefore R^2 \sim Exp(\frac{1}{2})$.

#

- Application: Use Uniform R.V. to simulate std. normals.

Let U_1, U_2 be $Unif[0,1]$, indep.

- $-2\log U_1 \sim Exp(\frac{1}{2})$: as $\mathbb{P}(-2\log U_1 > x) = \mathbb{P}(U_1 < e^{-\frac{x}{2}}) = e^{-\frac{x}{2}}$ if $x > 0$

  - For $f(x) = \lambda e^{-\lambda x}$ $(x > 0)$ $\sim Exp(\lambda)$

    $\int_x^{+\infty} f(t) dt = -e^{-\lambda t}\Big|_{t=x}^{t=+\infty} = e^{-\lambda x}$

- $2\pi U_2 \sim Unif[0, 2\pi]$

Take $\begin{cases} R^2 = -2\log U_1 \\ \Theta = 2\pi U_2 \end{cases}$, i.e. $\begin{cases} X = \sqrt{-2\log U_1} \cos(2\pi U_2) \\ Y = \sqrt{-2\log U_1} \sin(2\pi U_2) \end{cases}$ & $\overset{\text{how? ①}}{\boxed{X, Y \text{ are indep. } \mathcal{N}(0,1)}}$ $\boxed{\text{Box-Muller Transform}}$

Answer to ①:

Let $L = R^2$

$X = \sqrt{L}\cos\Theta$

$Y = \sqrt{L}\sin\Theta$

$|J| = \left\| \begin{matrix} \frac{1}{2\sqrt{l}}\cos\theta & -\sqrt{l}\sin\theta \\ \frac{1}{2\sqrt{l}}\sin\theta & \sqrt{l}\cos\theta \end{matrix} \right\| = \frac{1}{2}$

$\therefore f_{L,\Theta}(l, \theta) = f_{X,Y}(\sqrt{l}\cos\theta, \sqrt{l}\sin\theta) \cdot \frac{1}{2}$

$\| $

$\frac{1}{2} e^{-\frac{1}{2}l} \cdot \frac{1}{2\pi}$

$\Rightarrow f_{X,Y}(\sqrt{l}\cos\theta, \sqrt{l}\sin\theta) = \frac{1}{2\pi} e^{-\frac{1}{2}l}$

$\Rightarrow f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2 + y^2)} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$

$\Rightarrow X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$, indep.

eg. X, Y are indep. $\mathcal{N}(0,1)$, for $\sigma_1, \sigma_2 > 0$, $\rho \in (-1, 1)$

Let $\begin{cases} U = \sigma_1 X \sim \mathcal{N}(0, \sigma_1^2) \\ V = \sigma_2 \rho X + \sigma_2\sqrt{1-\rho^2}\, Y \sim \mathcal{N}(0, \sigma_2^2) \end{cases}$

(ex.) Show that the joint density funct. $f_{U,V}(u, v) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(u,v)}$

where $Q(u, v) = \frac{1}{1-\rho^2}\left(\left(\frac{u}{\sigma_1}\right)^2 - 2\rho \cdot \frac{u}{\sigma_1} \cdot \frac{v}{\sigma_2} + \left(\frac{v}{\sigma_2}\right)^2\right)$] Bivariate normal (Lec 14)

eg. Let $\begin{cases} U \sim \mathcal{N}(0, \sigma_1^2) \\ V \sim \mathcal{N}(0, \sigma_2^2) \end{cases}$ be bivariate normal; Compute $\underset{\sigma_1\sigma_2\rho}{\mathbb{E}UV}$, $\underset{\frac{\sigma_2}{\sigma_1}\rho U}{\mathbb{E}[V|U]}$, $\underset{\sigma_2^2(1-\rho^2)}{Var[V|U]}$

$\mathbb{E}UV = \sigma_1\sigma_2\rho \underbrace{\mathbb{E}X^2}_{\underbrace{VarX}_{1} - \underbrace{(\mathbb{E}X)^2}_{0}} + \sigma_1\sigma_2\sqrt{1-\rho^2} \underbrace{\mathbb{E}XY}_{\underbrace{\mathbb{E}X}_{0} \cdot \underbrace{\mathbb{E}Y}_{0}} = \sigma_1\sigma_2\rho$

Given $U = u$, $V = \sigma_2\rho \cdot \frac{u}{\sigma_1} + \sigma_2\sqrt{1-\rho^2}\, Y \sim \mathcal{N}\left(\frac{\sigma_2}{\sigma_1}\rho u, \sigma_2^2(1-\rho^2)\right)$

Thus, $\mathbb{E}[V|U=u] = \frac{\sigma_2}{\sigma_1}\rho u$

$\therefore \mathbb{E}[V|U] = \frac{\sigma_2}{\sigma_1}\rho U$

$Var[V|U] = \sigma_2^2(1-\rho^2)$

#

4/15 HTOP 20

Monday, April 15, 2024 11:18 AM

§ Generating Functions

- Given a seq. $(a_n)_{n\in\mathbb{N}}$, define the $\boxed{\text{generating funct.}}$: $G_a(s) \triangleq \sum_{n\ge 0} a_n s^n$

  Reconstruct: $a_n = \frac{1}{n!} G_a^{(n)}(0)$ eg. $a_n = \binom{N}{n}$, $G_a(s) = \sum_n \binom{N}{n} s^n = (1+s)^N$

eg. Let $a_n = e^{in\theta}$, which forms an orthonormal basis of $L^2[0, 2\pi)$

$$G_a(s) = \sum_{n=0}^{+\infty} e^{in\theta} \cdot s^n = \frac{1}{1-e^{i\theta}\cdot s}$$

- Convolution of seq.: Given $(a_i), (b_i)$.

def: $c_n \triangleq a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0$

Then $G_c(s) = G_a(s) \cdot G_b(s)$

pf.
$$G_c(s) = \sum_{n\ge 0} c_n s^n = \sum_{n=0}^{+\infty}\sum_{k=0}^{n} a_k b_{n-k} s^n$$
$$= \sum_{k=0}^{+\infty}\sum_{n=k}^{+\infty} a_k s^k \cdot b_{n-k} s^{n-k}$$
$$= \sum_{k=0}^{+\infty} a_k s^k \underbrace{\sum_{n=k}^{+\infty} b_{n-k} s^{n-k}}_{\sum_{n=0}^{+\infty} b_n s^n}$$
$$= G_a(s)\cdot G_b(s)$$

#

[Discrete Case]

Def. Let X be a discrete R.V. taking value in $\mathbb{N}$.

The $\boxed{\text{prob. generating funct.}}$: $G_X(s) \triangleq \mathbb{E}(s^X) = \sum_{n=0}^{+\infty} \mathbb{P}(X=n) s^n$

Cor. If X, Y are indep., then $G_{X+Y}(s) = G_X(s)\cdot G_Y(s)$ (proved by convolution)

eg. Let $X \sim Poi(\lambda)$, $Y \sim Poi(\mu)$. X, Y are indep. Show that $X+Y \sim Poi(\lambda+\mu)$. (Reci 4)

pf. $G_X(s) = \sum_{n\ge 0} \mathbb{P}(X=n) s^n = \sum_{n\ge 0} \frac{\lambda^n}{n!} e^{-\lambda} s^n = e^{\lambda s} e^{-\lambda} = e^{\lambda(s-1)}$

$G_Y(s) = e^{\mu(s-1)}$.

$\therefore G_{X+Y}(s) \overset{\text{Cor.}}{=} G_X(s)\cdot G_Y(s) = e^{(\lambda+\mu)(s-1)} = G_{Poi(\lambda+\mu)}(s)$

$\therefore X+Y \sim Poi(\lambda+\mu)$

#

[Continuous Case]

Def. The $\boxed{\text{prob. generating funct.}}$ for X is $G_X(s) \triangleq \mathbb{E}(s^X) = \int s^x \cdot f(x)\,dx = \mathcal{L}[f](s)$ (Laplace Transformation)

Rmk. ① There is a radius of conv. R s.t. $G_X(s)$ converges absolutely for $|s|<R$, and diverges for $|s|>R$

In fact, $R \ge 1$ b.c. $G_X(1) = 1$

② For $|s| < R$, we can differentiable / integrate term by term

③ If $G_a(s) = G_b(s)$ for $s \in (-\delta, \delta)$ for some $\delta > 0$. Then $G_a \equiv G_b$ and $a_n = \frac{1}{n!} G^{(n)}(0) = b_n$

④ Abel's Thm.

- If $a_i \ge 0$ and $G_a(s)$ converges absolutely for $|s|<1$

  Then $\lim_{s\nearrow 1} G_a(s) = G_a(1) = \sum_{n\ge 0} a_n$

- Generating funct. determines all moments (thus the distribution)

lem. If X has generating funct. $G(s)$, then

① $G'(1) = \mathbb{E}X$

② $G^{(k)}(1) = \mathbb{E}\left[X(X-1)\cdots(X-k+1)\right]$

pf. We know $R \ge 1$

For every $|s|<1$, $G^{(k)}(s) = \mathbb{E}\left[X(X-1)\cdots(X-k+1)\, s^{X-k}\right]$

Apply Abel's Thm. to send $s \nearrow 1$

#

- Moments: $\mathbb{E}X = G_X'(1)$

$\mathbb{E}X^2 = \mathbb{E}X(X-1) + \mathbb{E}X = G''(1) + G'(1)$

$\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = G''(1) + G'(1) - [G'(1)]^2$

eg. Coin flips, Bernolli(p). A wins if m^{th} head occurs before n^{th} tail

$\mathbb{P}[\text{A wins}] \triangleq P_{m,n}$ (Lec 5)

Sol. (Pascal)
$$\begin{cases} P_{m,n} = p\cdot P_{m-1,n} + q\cdot P_{m,n-1} \\ P_{m,0} = 0 \\ P_{0,n} = q^n \end{cases}$$

Generating function: $G(x,y) \triangleq \sum_{\substack{m\ge 0\\ n\ge 0}} P_{m,n} x^m y^n$

$$\therefore G(x,y) = \sum_{\substack{m\ge 1\\ n\ge 0}} p\cdot P_{m-1,n} x^m y^n + \sum_{\substack{m\ge 0\\ n\ge 1}} q\cdot P_{m,n-1} x^m y^n + \sum_{n\ge 0} \underbrace{P_{0,n}}_{q^n} y^n + \sum_{m\ge 0} \underbrace{P_{m,0}}_{0} x^m$$
$$= \sum_{\substack{m\ge 1\\ n\ge 0}} px\cdot P_{m-1,n} x^{m-1} y^n + \sum_{\substack{m\ge 0\\ n\ge 1}} qy\cdot P_{m,n-1} x^m y^{n-1} + \frac{1}{1-qy}$$
$$= (px+qy)\, G(x,y) + \frac{1}{1-qy} \Rightarrow G(x,y) = \frac{1}{(1-qy)(1-px-qy)}$$

Then we can reconstruct $P_{m,n}$ by Taylor-expansion.

#

- $\boxed{\text{Moment Generating Function}}$: $M_X(t) \triangleq \mathbb{E}e^{tX} = G_X(e^t)$

- If $e^t < R$, then $M_X(t) = \mathbb{E}\left(\sum_{n=0}^{+\infty} \frac{(tX)^n}{n!}\right) = \sum_{n=0}^{+\infty} \frac{t^n}{n!} \mathbb{E}X^n$, $\mathbb{E}X^n = M_X^{(n)}(0)$

  radius of conv.

  way easier to compute moments than $G_X(s)$

eg. ① $X \sim Poi(\lambda)$. $G_X(s) = e^{\lambda(s-1)}$

$M_X(t) = G_X(e^t) = e^{\lambda(e^t-1)}$

② $X \sim \mathcal{N}(0,1)$. $G_X(s) = s^{\frac{1}{2}\ln s}$, $M_X(t) = e^{\frac{1}{2}t^2}$

$Y \sim \mathcal{N}(\mu, \sigma^2)$. $G_Y(s) = s^{\mu}\cdot s^{\frac{1}{2}\sigma^2 \ln s}$, $M_Y(t) = \mathbb{E}e^{(\sigma X+\mu)t} = e^{\mu t} e^{\frac{1}{2}\sigma^2 t^2}$

③ $X \sim \text{Bernoulli}(p)$: $G_X(s) = \mathbb{E}s^X = (1-p) + ps$, $M_X(t) = (1-p) + pe^t$

④ $X \sim \text{Geometric}(p)$: $G_X(s) = \sum_{n\ge 1} (1-p)^{n-1} p s^n = \frac{ps}{1-(1-p)s}$, $M_X(t) = \frac{pe^t}{1-(1-p)e^t}$

⑤ $X \sim \text{Binomial}(n,p)$. $X = Y_1 + \cdots + Y_n$, $Y_i \sim \text{Bernoulli}(p)$, Y_i i.i.d.

$G_X(s) = G_{Y_1}(s)\cdots G_{Y_n}(s) = G_{Y_1}^n(s) = (1-p+ps)^n$. $M_X(t) = (1-p+pe^t)^n$

- $\boxed{\text{Random Sum Formula}}$

$S_N = \sum_{i=1}^{N} X_i$, (X_i) are i.i.d. with generating funct. G_X

N is indep. of X, having generating funct. G_N

Then the generating funct. of S_N is $G_N(G_X(s))$

pf.
$$G_{S_N}(s) = \mathbb{E}\left[s^{S_N}\right] \overset{\text{Towering}}{=} \mathbb{E}\left[\mathbb{E}\left[s^{S_N} \mid N\right]\right]$$
$$= \sum_n \mathbb{P}(N=n)\cdot \underbrace{\mathbb{E}\left[s^{S_N} \mid N=n\right]}_{\mathbb{E}\left[s^{X_1+\cdots+X_n}\right] \overset{\text{indep.}}{=} \left(\mathbb{E}\left[s^{X_1}\right]\right)^n}$$
$$= \sum_n \mathbb{P}(N=n)\left(G_X(s)\right)^n$$
$$\overset{\text{def.}}{=} G_N(G_X(s))$$

#

4/17 HTOP 21

Wednesday, April 17, 2024 11:18 AM

Recall: Generating funct: $G_X(s) = \mathbb{E}s^X$

Random sum formula: $S_N \triangleq \sum_{i=1}^{N} X_i$, X_i i.i.d. — G_X

N indep. of (X_i) — G_N

$$G_{S_N}(x) = G_N(G_X(s))$$

eg. $N \sim Poi(\lambda)$, X_i i.i.d. Bernoulli (p)

$$S_N = X_1 + \cdots + X_N$$

Sol. $G_N(s) = e^{\lambda(s-1)}$, $G_X(s) = ps + (1-p)$

$\Rightarrow G_{S_N}(s) = G_N(G_X(s)) = e^{\lambda p(s-1)} = G_{Poi(\lambda p)}(s) \Rightarrow S_N \sim Poi(\lambda p)$

#

Def. The $\boxed{\text{joint generating funct.}}$ for X_1 & X_2 is defined by

$$G_{X_1,X_2}(s_1, s_2) \triangleq \mathbb{E}s_1^{X_1} s_2^{X_2}$$

Thm. X_1, X_2 are indep. $\Leftrightarrow G_{X_1,X_2}(s_1,s_2) = G_{X_1}(s_1) \cdot G_{X_2}(s_2)$, for s_1, s_2 in some neighbhd $(-\delta, \delta)$

pf. $\Rightarrow$): Recall that if X_1, X_2 are indep.

Then $\mathbb{E}f(X_1)g(X_2) = \mathbb{E}f(X_1) \cdot \mathbb{E}g(X_2)$, $\forall f, g$ m'ble

Take $f(x) = s_1^x$, $g(x) = s_2^x \Rightarrow G_{X_1,X_2}(s_1,s_2) = G_{X_1}(s_1) \cdot G_{X_2}(s_2)$

$\Leftarrow$): (only for discrete R.V.'s)

$$G_{X_1,X_2}(s_1,s_2) = \sum_{i,j} \mathbb{P}(X_1 = i, X_2 = j)\, s_1^i s_2^j$$

$$G_{X_1}(s) \cdot G_{X_2}(s) = \sum_i \mathbb{P}(X_1 = i)\, s_1^i \cdot \sum_j \mathbb{P}(X_2 = j)\, s_2^j$$

Coeff of $s_1^i s_2^j$: $\mathbb{P}(X_1 = i, X_2 = j) = \mathbb{P}(X_1 = i) \cdot \mathbb{P}(X_2 = j)$

$\Rightarrow X_1, X_2$ are indep.

#

eg. X, Y are indep. $N(u, \sigma^2)$. Find the joint density & distri. of $X+Y$ and $X-Y$.

Sol. $\boxed{\text{Joint moment generating funct.}}$: $\mathbb{E}\left[e^{t(X+Y)} e^{s(X-Y)}\right] = \mathbb{E}e^{(t+s)X} \cdot \mathbb{E}e^{(t-s)Y}$

$[X \sim N(u, \sigma^2),\ \mathbb{E}e^{tX} = e^{ut + \frac{1}{2}\sigma^2 t^2}]$

$$= e^{u(t+s) + \frac{1}{2}\sigma^2(t+s)^2} \cdot e^{u(t-s) + \frac{1}{2}(t-s)^2\sigma^2}$$

$$= e^{2ut + \sigma^2 t^2} \cdot e^{\sigma^2 s^2}$$

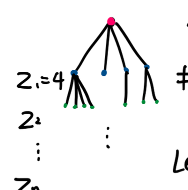
Thus, $X+Y \sim N(2u, 2\sigma^2)$, $X-Y \sim N(0, 2\sigma^2)$, indep. Then $f_{X+Y,X-Y}(t,k) = \frac{1}{2\pi \cdot 2\sigma^2} e^{-\frac{(t-2u)^2 + k^2}{2\sigma^2}}$

$$F_{X+Y,X-Y}(p,q) = \int_{-\infty}^{p} \frac{1}{\sqrt{2\pi \cdot 2\sigma^2}} e^{-\frac{(t-2u)^2}{2\sigma^2}} dt + \int_{-\infty}^{p} \frac{1}{\sqrt{2\pi \cdot 2\sigma^2}} e^{-\frac{k^2}{2\sigma^2}} dk$$

#

§ Branching Process (Galton-Watson Process)

Motivation: evolution of population / family names / spread of disease



$Z_0 = 1$

$Z_1 = 4$

Z_2

$\vdots$

Z_n

\# of children for each individual (X_i) is i.i.d., each has generating funct. G

Let $G_n(s) = \mathbb{E}s^{Z_n}$ where Z_n is the R.V. of the # of children at the n^{th} generation

lem. $G_n(s) = G_{n-1}(G(s)) = \underbrace{G(G(\cdots G(s))\cdots)}_{n}$

pf. $Z_n = X_1 + X_2 + \cdots + X_{Z_{n-1}}$ by random sum formula:

$G_n(s) = G_{Z_{n-1}}(G_X(s)) = G_{n-1}(G(s))$ since $G_X(s) = G_{Z_1}(s)$.

#

In practice, one can relate moments of Z_n to moments of Z_1

lem. Let $u = \mathbb{E}Z_1$, $\sigma^2 = Var\,Z_1$. Then $\mathbb{E}Z_n = u^n$ and $Var\,Z_n = \begin{cases} n\sigma^2, & u = 1 \\ \dfrac{\sigma^2(u^n - 1)u^{n-1}}{u-1}, & u \neq 1 \end{cases}$

pf. $G_n(s) = G_{n-1}(G(s))$

Diff. $G_n'(1) = G_{n-1}'(\underbrace{G(1)}_{1}) \cdot \underbrace{G'(1)}_{\mathbb{E}Z_1 = u} \Rightarrow \mathbb{E}Z_n = \mathbb{E}Z_{n-1} \cdot u \Rightarrow \mathbb{E}Z_n = u^n$

Diff twice: $G_n''(1) = G_{n-1}''(G(1)) \cdot (G'(1))^2 + G_{n-1}'(G(1)) \cdot \underbrace{G''(1)}$

$= G_{n-1}''(1) \cdot u^2 + u^{n-1}(\sigma^2 + u^2 - u)$

$\mathbb{E}[Z_1(Z_1 - 1)]$

$= \mathbb{E}Z_1^2 - \mathbb{E}Z_1$

$= Var\,Z_1 + (\mathbb{E}Z_1)^2 - \mathbb{E}Z_1$

$= \sigma^2 - u^2 - u$

Solve the recursion.

#

eg. $\boxed{\text{Geometric Branching:}}$

Assume $\mathbb{P}(Z_1 = k) = p^k q$, $q = 1 - p$.

$$G(s) = \mathbb{E}s^{Z_1} = \frac{q}{1 - ps}$$

ex. Show that

$$G_n(s) = \begin{cases} \dfrac{n - (n-1)s}{n+1 - ns}, & p = q = \frac{1}{2} \\ \dfrac{q(p^n - q^n - ps(p^{n-1} - q^{n-1}))}{p^{n+1} - q^{n+1} - ps(p^n - q^n)}, & p \neq q \end{cases}$$

- Question: Extinction / Non-extinction ?

$\mathbb{P}(Z_n = 0) = G_n(0) = \begin{cases} \dfrac{n}{n+1}, & p = q = \frac{1}{2} \\ \dfrac{q \cdot (p^n - q^n)}{p^{n+1} - q^{n+1}}, & p \neq q \end{cases} \Rightarrow \mathbb{P}[\text{Extinction}] = \begin{cases} 1, & p \leq \frac{1}{2} \\ \frac{q}{p}, & p > \frac{1}{2} \end{cases}$

since $G_n(s) = \sum_s s^i \mathbb{P}(X = i)$

As $n \uparrow +\infty$, $\{Z_n = 0\} \uparrow \bigcup_{n=1}^{+\infty} \{Z_n = 0\} = \{\text{Extinction}\}$

Since $\mathbb{E}Z_1 = \frac{p}{q}$, then the result is: $\begin{cases} \mathbb{E}Z_1 \leq 1 \Rightarrow \text{extinction} \\ \mathbb{E}Z_1 > 1 \Rightarrow \text{positive prob. of infinite growth} \end{cases}$

Thm. (General Case)

If ① $u = \mathbb{E}Z_1 < 1$, then $\mathbb{P}[\text{Extinction}] = 1$

② $u = \mathbb{E}Z_1 > 1$, then $\mathbb{P}[\text{Extinction}] < 1$

③ $u = \mathbb{E}Z_1 = 1$, then $\sigma^2 = Var\,Z_1 > 0$, $\mathbb{P}[\text{Extinction}] = 1$

→ Moreover, $\eta = \mathbb{P}[\text{Extinction}]$, then η is the smallest non-negative solution to $\eta = G(\eta)$

pf. Let $\eta_n = \mathbb{P}(Z_n = 0)$, then $\eta_n = G_n(0) = G(G_{n-1}(0)) = G(\eta_{n-1})$

take $n \uparrow +\infty$, notice that $\eta_n \uparrow \eta$, we have $\eta = G(\eta)$

Claim: If $\varphi \geq 0$ is a sol. to $\varphi = G(\varphi)$, then $\eta \leq \varphi$

pf. Indeed, $G(s) = \mathbb{E}s^{Z_1}$ is increasing in $[0, +\infty)$ since $Z_1 \geq 0$

$\eta_1 = G(0) \leq G(\varphi) = \varphi$

$\eta_2 = G(\eta_1) \leq G(\varphi) = \varphi$

Iterate, $\eta_n = G(\eta_{n-1}) \leq G(\varphi) = \varphi$

sending $n \uparrow +\infty$

#

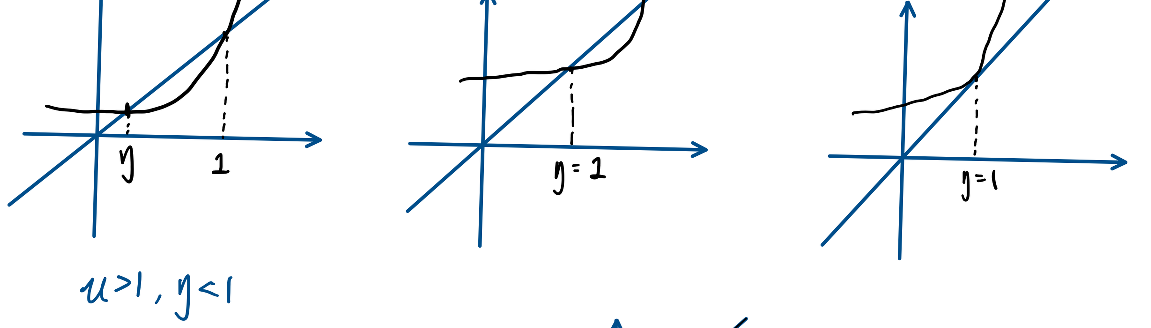
Now to prove ①, ②, ③ situations

Since $G(s)$ is increasing in $[0, +\infty)$

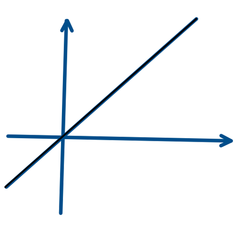
$G'(1) = u$

$G''(s) = \mathbb{E}[Z_1(Z_1 - 1)s^{Z_1 - 2}] \geq 0$

Z_1 takes value in $\mathbb{N}$



$u > 1$, $\eta < 1$

for ③: if $\sigma^2 = 0$, then 

#

4/19 HTOP (Reci 11)

Friday, April 19, 2024 11:18 AM

ex. $30. [$1] [$2] [$5] [$10]

Sol. (Using generating funct.)

$$G(z) = (1+z+z^2+\cdots)(1+z^2+z^4+\cdots)(1+z^5+z^{10}+\cdots)(1+z^{10}+z^{20}+\cdots)$$

$$= \frac{1}{1-z}\cdot\frac{1}{1-z^2}\cdot\frac{1}{1-z^5}\cdot\frac{1}{1-z^{10}} \quad \text{for } |z|<1$$

Compute $\frac{1}{30!}\frac{d^{(30)}G(z)}{dz^{(30)}}\Big|_{z=0}$

#

ex. Using generating funct. to show:

$$\sum_{\substack{j,k\\ j+k=l}}\binom{m+j-1}{j}\binom{n+k-1}{k} = \binom{m+n+l-1}{l}$$

Sol. Binomial $(1+z)^n = \sum_k \binom{n}{k} z^k$, $n\in\mathbb{N}$

$$\frac{1}{(1+z)^n} = 1 - nz + \frac{(-n)(-n-1)}{2!}z^2 + \frac{(-n)(-n-1)(-n-2)}{3!}z^3+\cdots$$

$$= \sum_k (-1)^k z^k \cdot \frac{(n+k-1)!}{(n-1)!\,k!} = \sum_k (-1)^k\binom{n+k-1}{k} z^k$$

#

ex. R.V. $U: \Omega \to [0,1]$

$$U(\omega) = (0.\xi_1(\omega)\xi_2(\omega)\cdots\xi_n(\omega)\cdots)_2$$

Show that (ξ_i) i.i.d. Bernoulli$(\frac{1}{2})$ $\Leftrightarrow$ U is uniform in $[0,1]$

Sol: (Using moment generating funct.)

Let $X = (0.\xi_1\xi_2\cdots)_2 = \sum_{i\ge1}\frac{\xi_i}{2^i}$ where (ξ_i) i.i.d. Bernoulli$(\frac{1}{2})$

$$\mathbb{E}[e^{tX}] = \mathbb{E}\left[e^{t\sum_{i\ge1}\frac{\xi_i}{2^i}}\right] \overset{\text{i.i.d.}}{=} \prod_{i\ge1}^{+\infty}\mathbb{E}\left[e^{\frac{t}{2^i}\xi_i}\right]$$

$$\overset{\text{MGF of Bernoulli}(\frac{1}{2})}{=} \prod_{i\ge1}^{+\infty}\left(\frac{1}{2}+\frac{1}{2}e^{\frac{t}{2^i}}\right)$$

Since $\prod_{i\ge1}^{n}\left(\frac{1}{2}+\frac{1}{2}\cdot e^{\frac{t}{2^i}}\right) = \frac{1}{2^n}\cdot\frac{1}{1-e^{\frac{t}{2^n}}}\cdot\prod_{i=1}^{n}(1+e^{\frac{t}{2^i}})(1-e^{\frac{t}{2^n}})$

$$= \frac{1}{2^n}\cdot\frac{1}{1-e^{\frac{t}{2^n}}}\cdot(1-e^t)$$

$$\xrightarrow{n\to\infty} \frac{1-e^t}{t}$$

Then $\mathbb{E}[e^{tX}] = \frac{1-e^t}{t} = \int_0^1 e^{tu}\,du = \mathbb{E}[e^{tU}]$

Thus $X \overset{(d)}{=} U$.

#

Def. Let $(X_n)_{n\ge1}$, Y be R.V.'s, then we say X_n $\boxed{\text{converges in distribution}}$ to Y

$X_n \xrightarrow{D} Y$ iff $\mathbb{P}(X_n\le x) \longrightarrow \mathbb{P}(Y\le x)$ for all x s.t. F_Y is conti. at x.

$\Leftrightarrow F_{X_n}(x) \longrightarrow F_Y(x)$

Thm. If $M_{X_n}(t)\to M_Y(t)$ for all $t\in(-\delta,\delta)$

Then $X_n \xrightarrow{D} Y$.

Recall: Thm. (Chernoff Bound) (Reci 6)

· Let (X_i) i.i.d. are Bernoulli(p), $X = \sum_{i=1}^n X_i$

Then $\mathbb{P}[X \ge (1+\delta)np] \le \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{np}$

Thm. $\boxed{\text{(Hoeffding Ineq.)}}$

· Let (X_i) i.i.d. $X_i\in[a_i,b_i]$, $u=\sum_{i=1}^n \mathbb{E}X_i$

Then $\mathbb{P}\left(\left|\sum_{i=1}^n X_i - u\right|\ge t\right) \le c\,e^{-\frac{2t^2}{\sum_{i=1}^n(b_i-a_i)^2}}$

Lem. Let X be a R.V., $\mathbb{E}X=0$, $X\in[a,b]$

Then $\mathbb{E}[e^{tX}] \le e^{\frac{1}{8}t^2(b-a)^2}$

Essential Idea:

C-S Ineq: $\mathbb{P}(X\ge a) \le \frac{1}{a}\mathbb{E}(X)$

Let f increasing $\mathbb{R}^+\to\mathbb{R}^+$ $\Rightarrow \mathbb{P}(f(X)\ge a) \le \frac{1}{a}\mathbb{E}[f(X)]$

$\Rightarrow \mathbb{P}(X \ge f^{-1}(a)) \le \frac{1}{a}\cdot\mathbb{E}[f(X)]$

let $b=f^{-1}(a)$ $\Rightarrow \mathbb{P}(X\ge b) \le \frac{1}{f(b)}\mathbb{E}[f(X)]$

let $f_c(x)=e^{cx}$ $\Rightarrow \mathbb{P}(X\ge b) \le \frac{1}{e^{cb}}\mathbb{E}[e^{cX}] \triangleq g(c)$

for a given b, we can optimize c to find a good bound.

4/21 HTOP 22

Sunday, April 21, 2024 11:19 AM

Recall: Moment generating funct. $M(t) = \mathbb{E}e^{tX}$ (Laplace transform) determines the distribution

However, $M(t) < +\infty$ only for r.v.'s with $f(x)$ decays exponentially as $x \to +\infty$

eg. $X \sim$ Cauchy (1): $f(x) = \frac{1}{\pi(1+x^2)}$, $\mathbb{E}X = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{x}{1+x^2}\,dx = +\infty$

$\therefore M(t)$ not defined for $\forall t \neq 0$. so does generating function $G(s)$

Rmk. There are examples that $\mathbb{E}X < \infty$ but MGF does not exist.

· Characteristic Function:

The char. funct. of X is defined by: $\phi(t) \triangleq \mathbb{E}e^{itX} = \int e^{itx} f_X(x)\,dx = \mathbb{E}\cos tX + \mathbb{E}\,i\sin tX$

eg. $X \sim$ Cauchy (1). $\phi(t) = \int_{-\infty}^{+\infty} e^{itx} \underbrace{\frac{1}{\pi(1+x^2)}}_{f(x)}\,dx = e^{-|t|}$, $t \in \mathbb{R}$

Sol.

$$\int_{-R}^{R} f(x)\,dx + \int_{A_R} f(z)\,dz = \int_{C_R} f(z)\,dz = 2\pi i \operatorname*{Res}_{z=i} f(z) = 2\pi i \cdot \lim_{z\to i} (z-i)f(z) = e^{-t} \quad ①$$

Since $\left|\int_{A_R} f(z)\,dz\right| \leq \frac{1}{\pi}\cdot\int_{A_R} \frac{|e^{itz}|}{|1+z^2|}\,dz \overset{t\geq 0}{\leq} \frac{1}{\pi}\cdot\frac{1}{R^2-1}\cdot\pi R = \frac{R}{R^2-1} \to 0$ as $R \to \infty$

Thus sending $R \to \infty$ in ①: $\phi(t) = \int_{-\infty}^{+\infty} f(x)\,dx = e^{-|t|}$, $\forall t \in \mathbb{R}$

\#

Thm. The char. funct. ϕ satisfies:

① $\phi(0) = 1$, $|\phi(t)| \leq 1$, $\forall t \in \mathbb{R}$ (trivial)

② ϕ is uniformly conti. in $\mathbb{R}$

③ ϕ is positive semidefinite: $\forall t_1, \cdots, t_n \in \mathbb{R}$, $z_1, \cdots, z_n \in \mathbb{C}$

$$\sum_{i,j} \phi(t_i - t_j)\cdot z_i \overline{z_j} \geq 0$$

pf. of ③. $\sum_{i,j} \mathbb{E}e^{i(t_i-t_j)X}\cdot z_i\overline{z_j} = \sum_{i,j}\mathbb{E}\left[e^{it_iX}\cdot \overline{e^{it_jX}}\right] z_i\overline{z_j}$

$$= \mathbb{E}\left[\sum_{i,j} e^{it_iX}\cdot z_i \cdot \overline{e^{it_jX} z_j}\right] = \mathbb{E}\left|\sum_i e^{it_iX} z_i\right|^2 \geq 0$$

\#

Recall: Monotone Convergence: If $f_n \geq 0$, and $f_n \uparrow f$ a.e., then $\lim \int f_n\,du = \int f\,du$

Fatou's lemma: If $f_n \geq 0$, then

ⓐ $\int \liminf_{n\to\infty} f_n\,du \leq \liminf_{n\to\infty} \int f_n\,du$

ⓑ If $\sup f_n \leq f$, and $\int f\,du < +\infty$, then

$\int \limsup_{n\to\infty} f_n\,du \geq \limsup_{n\to\infty}\int f_n\,du$

[Prob. Version]:

If $X_n \geq 0$, then $\mathbb{E}\liminf X_n \leq \liminf \mathbb{E}X_n$

pf. of Fatou's (a):

$\liminf f_n = \sup_{m\geq 1} \overbrace{\inf_{n\geq m} f_n}^{g_m}$, then $g_m \uparrow \liminf f_n$

By MCT, $\int \liminf_{n\to\infty} f_n\,du = \int \sup_{m\geq 1} g_m\,du = \lim_{m\to\infty}\int g_m\,du$ ⊛

On the other hand, $\forall n \geq m$, $\int g_m\,du \leq \int f_n\,du \Rightarrow \int g_m\,du \leq \inf_{n\geq m}\int f_n\,du$

sending $n \to +\infty$, $\lim_{m\to\infty}\int g_m\,du \leq \liminf_{n\to\infty}\int f_n\,du$

$\|$ ⊛

$\int \liminf_{n\to\infty} f_n\,du$

pf. of Fatou's (b): (ex.)

Hint: apply (a) to $f - f_n \geq 0$

\#

Rmk. ① equality of Fatou's lemma may not be attained

For example, let $f_n = 1_{[n,n+1)}$, then $0 = \int \liminf f_n\,du < \liminf \int f_n\,du = 1$

② If $\{f_n\}$ are not non-negative, Fatou's lemma may fail

For example, let $f_n = -1_{[n,n+1)}$, then $0 = \int \liminf f_n\,du > \liminf \int f_n\,du = -1$

Thm. (Dominated Conv.)

Let $(\Omega, \mathcal{F}, u)$ be a measure space, $\{f_n\}$, f are m'ble funct.'s

- $f_n \to f$ a.e.
- There is a m'ble funct. g s.t. $|f_n| \leq g$ for $\forall n \geq 1$, and $\int g\,du < +\infty$

then f is integrable, $\lim_{n\to+\infty}\int f_n\,du = \int f\,du$

pf. Since $f_n + g \geq 0$, by Fatou: $\int f + g \leq \int \liminf (f_n + g) = \liminf \int f_n + \int g$

$\Rightarrow \int f \leq \liminf \int f_n$

Since $g - f_n \geq 0$. by Fatou, $\int g - f = \int \liminf (g - f_n) \leq \int g - \liminf \int f_n$

$\Rightarrow \int f \geq \liminf \int f_n$

$\therefore \int f = \liminf \int f_n$

\#

Cor. (Bounded Conv.) If $u(\Omega) < +\infty$, $f_n \to f$ a.e., and $\exists K < +\infty$ s.t. $|f_n| \leq K$ a.e.

Then $\lim \int f_n = \int f$

pf. of ②:

$|\phi(t+h) - \phi(t)| = \left|\mathbb{E}e^{i(t+h)X} - \mathbb{E}e^{itX}\right|$

$= \left|\mathbb{E}e^{itX}\cdot(e^{ihX} - 1)\right|$

$\leq \mathbb{E}\left|e^{ihX} - 1\right|$ for $t \in \mathbb{R}$ since X is real

By bounded conv., $\lim_{h\to 0}\mathbb{E}\left|e^{ihX} - 1\right| = 0$. Thus ϕ is uniformly conti. in $\mathbb{R}$.

\#

Conversely,

· Bochner's thm. If ϕ satisfies ①, ②, ③.

Then there exists a unique prob. measure $\mathbb{P}$ s.t. $\phi(t) = \int e^{itX}\,d\mathbb{P}$

$\phi(t) = \mathbb{E}e^{itX} = 1 + it\,\mathbb{E}X - \frac{1}{2}t^2\mathbb{E}X^2 - \frac{1}{3!}it^3\mathbb{E}X^3 + \cdots$

Thm. ① If $\phi^{(k)}(0)$ exists, then $\mathbb{E}|X|^k < +\infty$ if k even

$\mathbb{E}|X|^{k-1} < +\infty$ if k odd

$\left(\mathbb{E}|X|^{2j+1}\right)^2 \leq \mathbb{E}|X|^{2j}\cdot\mathbb{E}|X|^{2j+2}$

② If $\mathbb{E}|X|^k < +\infty$, then $\phi(t) = \sum_{j=0}^{k}\frac{(it)^j}{j!}\mathbb{E}X^j + o(t^k)$

Thm. If X, Y are indep., then $\phi_{X+Y}(t) = \phi_X(t)\cdot\phi_Y(t)$

$\Leftrightarrow \mathbb{E}\left[e^{it(X+Y)}\right] = \mathbb{E}\left[e^{itX}\right]\cdot\mathbb{E}\left[e^{itY}\right]$

lem. If $Y = aX + b$, $a, b \in \mathbb{R}$, $\phi_Y(t) = e^{itb}\phi_X(at)$

Def. The joint charact. funct. of X, Y: $\phi(s,t) = \mathbb{E}\left[e^{isX + itY}\right]$

Thm. X, Y are indep. iff $\phi(s,t) = \phi_X(s)\cdot\phi_Y(t)$

pf. $\Leftarrow$ want to show $\mathbb{P}(X \leq a, Y \leq b) = \mathbb{P}(X \leq a)\cdot\mathbb{P}(Y \leq b)$. Follows from Fourier Inversion Thm.

4/22 HTOP 23

Monday, April 22, 2024 11:16 AM

$M(t) = \mathbb{E}e^{tX}$, $\phi(t) = \mathbb{E}e^{itX}$

Q1 : Does $\phi(t) = M(it)$ always ?

Q2 : Does moments : $m_k = \mathbb{E}X^k$, $\forall k \in \mathbb{N}$, uniquely determine the distribution ?

A : Yes. If $M(t)$ is finite in some $t \in (-\delta, \delta)$ nbhd

Thm. [Analytic Extension of $M(t)$]

For $\forall a > 0$, T.F.A.E. (the followings are equivalent)

① $|M(t)| < +\infty$ for $|t| < a$

② $\phi(z) \triangleq \mathbb{E}e^{zX}$ is analytic in the strip $|\operatorname{Re} z| < a$ (Cauchy R.V. is not true)

③ The moments m_k , $\forall k \in \mathbb{N}$ exists, and $\lim_{k \to +\infty} \left(\frac{m_k}{k!}\right)^{\frac{1}{k}} \lesssim \frac{1}{a}$

In this case, $M(t)$ can be extended to an analytic funct. in $|\operatorname{Re} z| < a$ & $\phi(t) = M(it)$

· Common Charact. funct. [Compare with Lec. 20 to see $M_X(t)$ & $G_X(s)$]

① delta measure ($X = a$) : $\phi(t) = \mathbb{E}e^{itX} = \mathbb{E}e^{iat}$

② Bernoulli (p) : $\phi(t) = \mathbb{E}e^{itX} = (1-p) + pe^{it}$

③ Binomial (n, p) : $X = Y_1 + \cdots + Y_n$, $Y_i \sim$ Bernolli (p) , $\mathbb{E}e^{itX} = (\mathbb{E}e^{itY_1})^n = (1-p+pe^{it})^n$

④ Poisson (λ) : $f(k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $\phi(t) = \mathbb{E}e^{itX} = \sum_k e^{itk} \cdot \frac{\lambda^k}{k!}e^{-\lambda} = e^{\lambda e^{it} - \lambda}$

⑤ Exp (λ) : $f(x) = \lambda e^{-\lambda x} 1_{x \geq 0}$, $\phi(t) = \int_0^{+\infty} e^{itx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - it}$

⑥ $\mathcal{N}(0, 1)$: $\phi(t) = \mathbb{E}e^{itX} = \int_{\mathbb{R}} e^{itx} e^{-\frac{1}{2}x^2} dx$

$M(t) = \mathbb{E}e^{tX} = e^{\frac{1}{2}t^2}$, by analytic extension, $\phi(t) = e^{-\frac{1}{2}t^2}$

$\mathcal{N}(u, 6^2)$: $Y \overset{(d)}{=} u + 6X$ where $X \sim \mathcal{N}(0,1)$, $\phi_Y(t) = e^{iut} \cdot \mathbb{E}e^{i6tX} = e^{iut - \frac{1}{2}6^2t^2}$

⑦ Multivariate Normal :

$(X_1, \cdots, X_n) \sim \mathcal{N}(0, V)$ iff $\operatorname{Cov}(X_i, X_j) = v_{ij}$

joint density : $f(x_1, \cdots, x_n) = \frac{1}{\sqrt{(2\pi)^n \cdot \det V}} e^{-\frac{1}{2}\vec{X}^T V^{-1} \vec{X}}$ where $\vec{X} = (X_1, \cdots, X_n)$

For $t_1, \cdots, t_n \in \mathbb{R}$: joint charact. funct. $\mathbb{E}\left[e^{i(t_1X_1 + \cdots + t_nX_n)}\right] = \frac{1}{\sqrt{(2\pi)^n \cdot \det V}} \int e^{i(t_1X_1 + \cdots + t_nX_n)} e^{-\frac{1}{2}\vec{X}^T V^{-1}\vec{X}} d\vec{X} = e^{-\frac{1}{2}\vec{t}^T V \vec{t}}$

$\exists$ orthogonal matrix B s.t. $BV^{-1}B^T = \Lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$

Let $\vec{Y} = B\vec{X} \Leftrightarrow \vec{X} = B^T\vec{Y}$.

Then $\vec{X}^T V^{-1} \vec{X} = \vec{Y}^T B V^{-1} B^T \vec{Y} = \vec{Y}^T \Lambda \vec{Y} = \sum_{i=1}^n \lambda_i Y_i^2 = \cdots \overset{(ex.)}{=} e^{-\frac{1}{2}\vec{t}^T V \vec{t}}$ where $\vec{t} = (t_1, \cdots, t_n)$

· Alternatively. Let $\vec{Y} = (Y_1, \cdots, Y_n) \triangleq B\vec{X}$. Then $Y \sim \mathcal{N}(0, \Lambda)$. $Y_1, \cdots, Y_n$ are indep. normal

$\Rightarrow t_1X_1 + \cdots + t_nX_n = \vec{t}^T\vec{X} = \vec{t}^T B^T \vec{Y}$ is a linear transform of indep. Normal & is therefore normal

Mean : $\mathbb{E}[t_1X_1 + \cdots + t_nX_n] = 0$. Variance : $\operatorname{Var}(t_1X_1 + \cdots + t_nX_n) = \sum_{i,j} t_i t_j \operatorname{Cov}(X_i, X_j) = \sum_{i,j} t_i V_{ij} t_j = \vec{t}^T V \vec{t}$

Then $\mathbb{E}\left[e^{i(t_1X_1 + \cdots + t_nX_n)}\right] = e^{-\frac{1}{2}\operatorname{Var}(t_1X_1 + \cdots + t_nX_n)} = e^{-\frac{1}{2}\vec{t}^T V \vec{t}}$

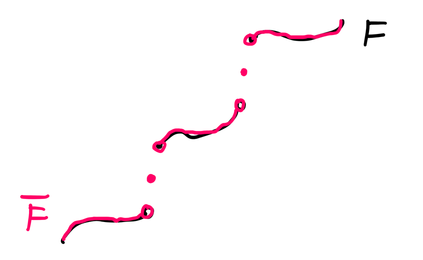
#

· Inversion and Continuity Theorem .

Recall : $\phi(t) = \frac{1}{2\pi}\int e^{itx} f(x)\, dx$, then $f(x) = \frac{1}{2\pi}\int e^{-itx}\phi(t)\, dt$ at every point x s.t. f is differentiable

Thm. (Inversion)

Let X have the distri. F and charact. funct. ϕ.



Let $\bar{F}(x) = \frac{1}{2}(F(x) + \lim_{y \nearrow x} F(y))$

Then for $\forall a < b$, $\bar{F}(b) - \bar{F}(a) = \lim_{N \to \infty} \int_{-N}^{N} \frac{e^{-iat} - e^{-ibt}}{2\pi i t} \phi(t)\, dt$

Cor. If $\phi_X(t) = \phi_Y(t)$, then $F_X = F_Y$

pf. Apply Inversion Thm. with $a = -\infty$: $\bar{F}_X(b) = \bar{F}_Y(b)$

$\forall x \in \mathbb{R}$, We use the right continuity of F_X and F_Y , take $b_n \searrow x$ s.t. F_X , F_Y are conti. at b_n

Take $\bar{F}_X(b_n) = F_X(b_n) = F_Y(b_n) = \bar{F}_Y(b_n) \Rightarrow F_X(x) = F_Y(x)$

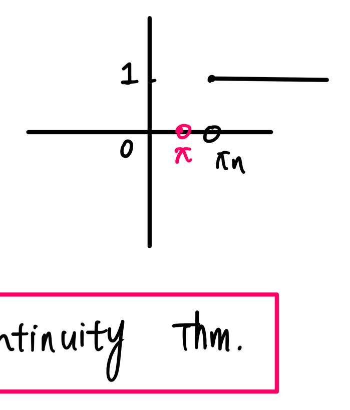
#

Def. We say that the r.v.'s $(X_n)_{n \geq 1}$ conv. in distri. to X : $X_n \xrightarrow{D} X$

iff $F_{X_n}(x) \longrightarrow F_X(x)$ at every point x where F_X is conti.

Rmk. Continuity assumption at x is to deal with discrete r.v. . Let $X_n = x_n \searrow x = X$

Then $F_{X_n}(x) = 0$ but $F_X(x) = 1$



· Continuity Thm.

Sps. $(X_n)_{n \geq 1}$ are r.v.'s with corresponding charact. funct. $(\phi_n)_{n \geq 1}$

ⓐ If $X_n \xrightarrow{D} X$, then $\phi_n(t) \to \phi(t)$ for $\forall t$

ⓑ Conversely , if $\phi_n(t) \to \phi(t)$ exists for $\forall t$, and $\phi(t)$ is conti. at $t=0$

Then ϕ is the charact. funct. of some r.v. X , and we have $X_n \xrightarrow{D} X$

eg. [Weak Law of Large Numbers 1]

Let $X_1, X_2, \cdots, X_n$ be a seq. of i.i.d. r.v.'s with finite mean u.

Then $S_n = X_1 + \cdots + X_n$ satisfies $\frac{1}{n}S_n \xrightarrow{\boxed{D}} u$

eg. [Strong Law of Large Numbers]

Let $X_1, \cdots, X_n : \Omega \to \mathbb{R}$ be i.i.d. with finite mean u , and $\mathbb{E}|X_1| < +\infty$, then $\frac{1}{n}S_n \longrightarrow u$ a.s.

pf. (Weak Law)

Let ϕ_n be the charact. funct. of $\frac{1}{n}S_n$

$\phi_n(t) = \mathbb{E}e^{it \cdot \frac{1}{n}(X_1 + \cdots + X_n)} \overset{\text{i.i.d.}}{=} \left(\mathbb{E}e^{i\frac{t}{n}X_1}\right)^n \overset{\mathbb{E}X_1 < +\infty}{=} \left(1 + i\frac{t}{n}\overset{u}{\mathbb{E}X_1} + o\left(\frac{t}{n}\right)\right)^n$

$\therefore \lim_{n \to +\infty} \phi_n(t) = \lim_{n \to +\infty} \left(1 + i\frac{t}{n}u + o\left(\frac{t}{n}\right)\right)^n = e^{itu} \underset{\text{Continuity thm.}}{\Rightarrow} \frac{1}{n}S_n \xrightarrow{D} u$

#

4/24 HTOP 24
Wednesday, April 24, 2024 11:20 AM

§ Convergence of Random Variables

Def. $(X_n)_{n\geq 1}$, X are r.v. in $(\Omega, \mathcal{F}, \mathbb{P})$

a. $X_n \to X$ $\boxed{\text{almost surely (a.s.)}}$ iff $\mathbb{P}(\{w : \lim_{n\to+\infty} X_n(w) = X(w)\}) = 1$ (a.e.)

b. $X_n \xrightarrow{\mathbb{P}} X$ $\boxed{\text{in probability}}$ iff for $\forall \varepsilon > 0$, $\mathbb{P}(|X_n - X| > \varepsilon) \longrightarrow 0$ as $n \nearrow +\infty$ (in measure)

c. $X_n \to X$ $\boxed{\text{in } p^{th} \text{ moment}}$ iff $\mathbb{E}|X_n|^p < +\infty$, $\forall n \in \mathbb{N}$ and $\mathbb{E}|X_n - X|^p \longrightarrow 0$ as $n \nearrow +\infty$ (in L^p sense)

d. $X_n \xrightarrow{\mathcal{D}} X$ $\boxed{\text{in distribution}}$ iff $F_{X_n}(x) \to F_X(x)$ at every point where F_X is conti.

Recall [Chebyshev Ineq.]: Let f be a m'ble funct. on $(\Omega, \mathcal{F})$,

Then $\forall \varepsilon > 0$, $\forall p \in \mathbb{N}$, $u(|f| > \varepsilon) \leq \frac{1}{\varepsilon^p} \int |f|^p \, du$

pf. of ①: $\forall \varepsilon > 0$, $\mathbb{P}(|X_n - X| > \varepsilon) \leq \frac{1}{\varepsilon^p} \mathbb{E}|X_n - X|^p \to 0$

#

Relation Map:

$$\begin{tikzcd}
X_n \to X \text{ a.s.} \arrow[d, shift left, "\leftarrow \text{Egorov}"] \arrow[drr, Rightarrow] \arrow[dd, dashed, bend right=60, "\text{+ Dominated / Monotone / Bounded Conv. } (*)"'] & & & \\
X_n \to X \text{ a.u.} \arrow[u, Rightarrow, shift left] \arrow[rr, Rightarrow] \arrow[d, dashed, shift left, "+ (*)"] & & X_n \xrightarrow{\mathbb{P}} X \arrow[r, Rightarrow] \arrow[ull, dashed, bend right, "\text{subseq.}"'] \arrow[ll, dashed, bend right, "\text{subseq.}"'] & X_n \xrightarrow{\mathcal{D}} X \\
X_n \to X \text{ in } L^p \arrow[u, dashed, shift left, "\text{subseq.}"] \arrow[urr, Rightarrow, "\text{Chebyshev}"'] \arrow[uu, dashed, bend left=40, "\text{subseq.}"] & & &
\end{tikzcd}$$

Recall: $\limsup_{n\to+\infty} A_n = \bigcap_{m\geq 1} \bigcup_{n\geq m} A_n$; $\liminf_{n\to+\infty} A_m = \bigcup_{m\geq 1} \bigcap_{n\geq m} A_n$

$\forall \varepsilon > 0$, $\forall n \geq 1$. Let $B_n(\varepsilon) = \{w : |X_n(w) - X(w)| \leq \varepsilon\}$

Then $\{X_n \to X\} = \{w \in \Omega : \forall \varepsilon > 0, \exists N \in \mathbb{N}^+ \text{ s.t. } \forall n \geq N, |X_n(w) - X(w)| \leq \varepsilon\}$

$$= \bigcap_{\varepsilon>0} \bigcup_{N\geq 1} \bigcap_{n\geq N} B_n(\varepsilon) = \bigcap_{k\geq 1} \bigcup_{N\geq 1} \bigcap_{n\geq N} B_n(\tfrac{1}{k}) = \bigcap_{k\geq 1} \liminf_{n\to+\infty} B_n(\tfrac{1}{k})$$

Lem. $X_n \to X$ a.s. $\overset{\text{conti. of prob.}}{\Longleftrightarrow}$ $\lim_{k\to+\infty} \mathbb{P}(\liminf_{n\to+\infty} B_n(\frac{1}{k})) = 1$

$\Leftrightarrow \lim_{k\to+\infty} \mathbb{P}(\limsup_{n\to+\infty} B_n^c(\frac{1}{k})) = 0 \Leftrightarrow \lim_{\varepsilon \searrow 0} \mathbb{P}(\limsup_{n\to+\infty} B_n^c(\varepsilon)) = 0$

· $\boxed{\text{Borel - Cantelli Thm.}}$

$(\Omega, \mathcal{F}, \mathbb{P})$ is a prob. space. $\{A_i\} \subseteq \mathcal{F}$

infinitely often

① if $\sum_{i=1}^{+\infty} \mathbb{P}(A_i) < +\infty$, then $\mathbb{P}(A_i, \text{ i.o.}) = 0 \Leftrightarrow \mathbb{P}(\limsup_{i\to+\infty} A_i) = 0$

② if $\sum_{i=1}^{+\infty} \mathbb{P}(A_i) = +\infty$ and A_i are indep.. Then $\mathbb{P}(A_i, \text{ i.o.}) = 1 \Leftrightarrow \mathbb{P}(\limsup_{i\to+\infty} A_i) = 1$

Rmk. · ② can be extended to (A_i) having pairwise indep.

· The thm. says certain event of infinite seq. occurs with prob. 0 or 1

· Special Case of the $\boxed{\text{Kolmogorov 0-1 Law}}$

$(X_n)_{n\geq 1}$, $\boxed{\text{filtration}}$ (increasing seq. of σ-algebra): $\sigma(X_1)$, $\sigma(X_1, X_2)$, ...

$\sigma(X_1, X_2) = \sigma(\sigma(X_1) \cup \sigma(X_2))$

Tail σ-algebra: $\mathcal{J} \triangleq \bigcap_{n=1}^{+\infty} \sigma(X_i, i \geq n)$

If (X_i) are indep., then elements of $\mathcal{J}$ has prob. either 0 or 1

eg. Coin flips ∞-many times. $A = \{\text{patterns } \underbrace{HH\cdots H}_{1000} \text{ occurs } \infty\text{-many times}\}$

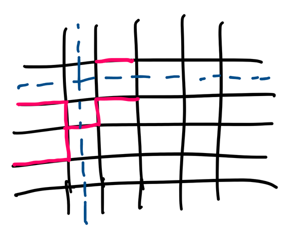
Sol. $A \in \mathcal{J}$, by 0-1 Law, $\underbrace{\mathbb{P}(A) = 0}_{\text{impossible}}$ or $\mathbb{P}(A) = 1$

We can also see from Borel - Cantelli Thm.

#

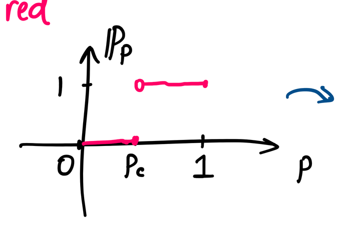
eg. Percolation (p)

连通到 infinity 的 connected path

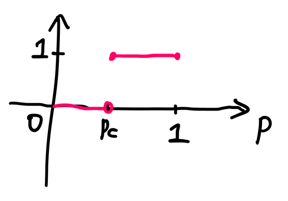


$\mathbb{P}_p$ (infinite open red cluster) = 0 or 1

Conjuncture:



$\rightarrow$ proved for $\mathbb{Z}^d$ where $d=2$ & $d \geq 19$.



pf. of Borel - Cantelli Thm.

①: $\mathbb{P}[\limsup_{i\to+\infty} A_i] = \mathbb{P}(\bigcap_{m\geq 1} \bigcup_{n\geq m} A_n) \overset{\text{conti. of prob.}}{=} \lim_{m\to+\infty} \mathbb{P}(\bigcup_{n\geq m} A_n) \leq \lim_{m\to+\infty} \sum_{n\geq m} \mathbb{P}(A_n) \to 0$

since $\sum_{n=1}^{+\infty} \mathbb{P}(A_n) < +\infty$

②: $\mathbb{P}[(\limsup_{i\to+\infty} A_i)^c] = \mathbb{P}[\bigcup_{m\geq 1} \bigcap_{n\geq m} A_n^c] = \lim_{m\to+\infty} \mathbb{P}(\bigcap_{n\geq m} A_n^c) \overset{\text{indep.}}{=} \lim_{m\to+\infty} \prod_{n\geq m} [1 - \mathbb{P}(A_n)]$

$1 - x \leq e^{-x}$ $\leq \lim_{m\to+\infty} \prod_{n\geq m} e^{-\mathbb{P}(A_n)} = \lim_{m\to+\infty} e^{-\sum_{n\geq m} \mathbb{P}(A_n)} = 0$ b.c. $\sum_{n=1}^{+\infty} \mathbb{P}(A_n) = +\infty$.

#

pf. of the Borel - Cantelli ② for extending (A_i) to be pairwise indep.

Let $S_n = \sum_{i=1}^{n} \mathbb{1}_{A_i}$. Therefore $\mathbb{P}[A_i, \text{ i.o.}] = \mathbb{P}[S_n = +\infty]$

We have $\mathbb{E}S_n = \sum_{i=1}^{n} \mathbb{P}(A_i) \longrightarrow +\infty$ as $n \to +\infty$ ②

$\text{Var } S_n = \text{Var}(\sum_{i=1}^{n} \mathbb{1}_{A_i}) \overset{①}{=} \sum_{i,j=1}^{n} \text{Cov}(\mathbb{1}_{A_i}, \mathbb{1}_{A_j}) \underset{\text{pairwise indep.}}{=} \sum_{i=1}^{n} \text{Var}(\mathbb{1}_{A_i}) = \sum_{i=1}^{n} [\mathbb{P}(A_i) - \mathbb{P}(A_i)^2] \overset{③}{\leq} \mathbb{E}S_n$ ③

$\mathbb{P}(S_n < \frac{1}{2}\mathbb{E}S_n) \leq \mathbb{P}(|S_n - \mathbb{E}S_n| > \frac{1}{2}\mathbb{E}S_n) \underset{\text{Markov Ineq. (Reci 6)}}{\leq} \frac{4}{(\mathbb{E}S_n)^2} \text{Var}(S_n) \overset{③}{\leq} \frac{4}{(\mathbb{E}S_n)^2} \cdot \mathbb{E}S_n = \frac{4}{\mathbb{E}S_n}$ ④

Thus for any fixed $k < +\infty$, $\exists N \in \mathbb{N}^+$ s.t. for all $n \geq N$,

$\mathbb{P}(S_n < k) \overset{②}{\leq} \mathbb{P}(S_n < \frac{1}{2}\mathbb{E}S_n) \overset{④}{\leq} \frac{4}{\mathbb{E}S_n} \overset{②}{\leq} \varepsilon$

Thus we have $\mathbb{P}(\sum_{i=1}^{+\infty} \mathbb{1}_{A_i} < +\infty) = 0$.

i.e. $\mathbb{P}(A_i, \text{ i.o.}) = \mathbb{P}(S_n = +\infty) = 1$.

#

① is always true since:

157

The answer to your question is "Sometimes, but not in general".

To see this let $X_1, \ldots, X_n$ be random variables (with finite variances). Then,

$$\text{var}\left(\sum_{i=1}^n X_i\right) = E\left(\left[\sum_{i=1}^n X_i\right]^2\right) - \left[E\left(\sum_{i=1}^n X_i\right)\right]^2$$

Now note that $(\sum_{i=1}^n a_i)^2 = \sum_{i=1}^n \sum_{j=1}^n a_i a_j$, which is clear if you think about what you're doing when you calculate $(a_1 + \ldots + a_n) \cdot (a_1 + \ldots + a_n)$ by hand. Therefore,

$$E\left(\left[\sum_{i=1}^n X_i\right]^2\right) = E\left(\sum_{i=1}^n \sum_{j=1}^n X_i X_j\right) = \sum_{i=1}^n \sum_{j=1}^n E(X_i X_j)$$

similarly,

$$\left[E\left(\sum_{i=1}^n X_i\right)\right]^2 = \left[\sum_{i=1}^n E(X_i)\right]^2 = \sum_{i=1}^n \sum_{j=1}^n E(X_i)E(X_j)$$

so

$$\text{var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \sum_{j=1}^n (E(X_i X_j) - E(X_i)E(X_j)) = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(X_i, X_j)$$

by the definition of covariance.

4/26 HTOP (Reci 12)

Friday, April 26, 2024 11:20 AM

Problem 1:

· Use charact. funct. to prove the $\boxed{\text{Central Limit Theorem}}$:

Let (X_i) be i.i.d., $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 = 1$.

Assume $\mathbb{E}X_i^3 < +\infty$. Show that $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$

(Weak law of large numbers: $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathcal{D}} 0$)

pf. $\phi_n(t) = \mathbb{E}\left[e^{i\cdot\frac{t}{\sqrt{n}}\cdot\sum_{i=1}^{n}X_i}\right]$

$$= \left(\mathbb{E}\left[e^{i\cdot\frac{t}{\sqrt{n}}\cdot X_1}\right]\right)^n$$

$$= \Big[1 + i\cdot\frac{t}{\sqrt{n}}\cdot\underbrace{\mathbb{E}X_1}_{0} - \frac{1}{2}\cdot\frac{t^2}{n}\cdot\underbrace{\mathbb{E}X_1^2}_{1} + o\Big(\frac{t^3}{n^{\frac{3}{2}}}\Big)\Big]^n$$

this is the main contributor to the difference between $\frac{1}{\sqrt{n}}S_n$ and $\mathcal{N}(0,1)$, and there should be a bound there which is called: Berry - Essean Bound

$\xrightarrow{n\to+\infty} e^{-\frac{t^2}{2}}$ which is the charact. funct. of $\mathcal{N}(0,1)$

#

Recall: · $f_n \to f$ a.e. iff $\mu(\{x : \lim_{n\to+\infty} f_n(x) \neq f(x)\}) = 0$

Def. (f_n) converges to f $\boxed{\text{almost uniformly}}$ iff $\forall\varepsilon>0$, $\exists N\in\mathcal{F}$ with $\mu(N)<\varepsilon$

s.t. $f_n \to f$ uniformly on N^c (uniformly a.e.)

We know: $f_n \to f$ a.e. iff $\forall\varepsilon>0$, $\mu\left(\limsup_{n\to+\infty}\{|f_n-f|>\varepsilon\}\right) = 0 = \mu\left(\bigcap_{m=1}^{+\infty}\bigcup_{n\geq m}\{|f_n-f|>\varepsilon\}\right)$

Compare these two

Prob. 2: $f_n \to f$ almost uniformly iff $\forall\varepsilon>0$, $\lim_{m\to+\infty}\mu\left(\bigcup_{n\geq m}\{|f_n-f|>\varepsilon\}\right) = 0$

pf. $\Rightarrow$): If $f_n\to f$ a.u., then $\forall\delta>0$, $\exists N\in\mathcal{F}$ s.t. $\mu(N)<\delta$ and $f_n\to f$ unif. on N^c

Thus $\forall\varepsilon>0$, $\exists m\in\mathbb{N}$ s.t. $\forall n\geq m$, $|f_n-f|\leq\varepsilon$ on N^c

For such m: $\bigcup_{n\geq m}\{|f_n-f|>\varepsilon\} \subseteq N$

$\Rightarrow \mu\left(\bigcup_{n\geq m}\{|f_n-f|>\varepsilon\}\right) \leq \mu(N) < \delta \Rightarrow \lim_{m\to+\infty}\mu\left(\bigcup_{n\geq m}\{|f_n-f|>\varepsilon\}\right)\leq\delta$

sending $\delta\searrow 0$ to conclude.

$\Leftarrow$): Sps. for $\forall k\geq 1$, $\lim_{m\to+\infty}\mu\left(\bigcup_{n\geq m}\{|f_n-f|>\frac{1}{k}\}\right) = 0$

$\forall\delta>0$, we can find $m = m(\delta,k)$ s.t. $\mu\left(\bigcup_{n\geq m}\{|f_n-f|>\frac{1}{k}\}\right)\leq\frac{\delta}{2^k}$

Let $N = \bigcup_{k=1}^{+\infty}\bigcup_{n\geq m(k,\delta)}\{|f_n-f|>\frac{1}{k}\}$

(ex.) Check that $\mu(N)<\delta$, and $f_n\to f$ unif. on N^c.

#

Littlewood 3 principle:

· m'ble set $\approx$ finite sum of intervals

(Lusin) · m'ble funct. $\approx$ conti. funct.

(Egorov) · Conv. seq. of (f_n) $\approx$ uniformly conv. (f_n)

Recall: · $f_n\to f$ in measure/prob. iff $\forall\varepsilon>0$, $\mu(|f_n-f|>\varepsilon)\longrightarrow 0$ as $n\to+\infty$

Cor. · If $f_n\to f$ a.u., then $f_n\to f$ in measure/prob.

· If $\mu(\Omega)<+\infty$, then $f_n\to f$ a.e. $\Rightarrow$ $f_n\to f$ in measure/prob.

Recall: [Borel - Cantelli] (extended)

ⓐ If $\sum_{i=1}^{+\infty}\mathbb{P}(A_i)<+\infty$, then $\mathbb{P}(\limsup A_i) = 0$

ⓑ $\sum_{i=1}^{+\infty}\mathbb{P}(A_i) = +\infty$, and (A_i) are pairwise indep., then $\mathbb{P}(\limsup A_i) = 1$

· $\boxed{\text{Sufficient criteria for a.s. convergence:}}$

(1) If $\forall\varepsilon>0$, $\sum_{n=1}^{+\infty}\mathbb{P}(|X_n-X|>\varepsilon)<+\infty$, then $X_n\to X$ a.s.

(2) If (X_n-X) are pairwise indep. and $\exists\,\varepsilon_k\searrow 0$, $\sum_{i=1}^{+\infty}\mathbb{P}(|X_n-X|>\varepsilon_k) = +\infty$,

then $X_n\not\to X$ a.s.

(Show that $\frac{1}{n}\sum X_i \longrightarrow \mathbb{E}X = \frac{1}{2}$ a.s.) (ex.)

eg. (X_n) i.i.d. Unif $[0,1]$, $Y_n = \min\{X_1,\cdots,X_n\}$, then $Y_n\longrightarrow 0$ a.s.

pf. $\mathbb{P}(|Y_n|>\varepsilon) = \mathbb{P}(X_1>\varepsilon, X_2>\varepsilon,\cdots,X_n>\varepsilon)$

$$= (1-\varepsilon)^n$$

Then $\sum_{n=1}^{+\infty}\mathbb{P}(|Y_n|>\varepsilon)<+\infty$

$\Rightarrow Y_n\longrightarrow 0$ by (1)

#

4/29 HTOP 25

Monday, April 29, 2024 11:20 AM

Recall: Borel – Cantelli:

- If $\sum_{i=1}^{+\infty} \mathbb{P}(A_i) < +\infty$, then $\mathbb{P}(A_i \text{ i.o.}) = \mathbb{P}(\limsup A_i) = 0$
- If $\sum_{i=1}^{+\infty} \mathbb{P}(A_i) = +\infty$, and (A_i) pairwise indep., then $\mathbb{P}(A_i \text{ i.o.}) = \mathbb{P}(\limsup A_i) = 1$

Application: If $\forall \varepsilon > 0$, $\mathbb{P}[|X_i - X| > \varepsilon] < +\infty$, then $X_i \to X$ a.s.

eg. $\boxed{\text{(Extreme Values)}}$

(X_i) i.i.d. $\mathrm{Exp}(1)$: $f(x) = e^{-x} 1_{\{x>0\}}$

Let $M_n \triangleq \max_{i=1,\cdots,n} X_i$,

Then $\frac{M_n}{\log n} \to 1$ a.s., also $\limsup \frac{X_n}{\log n} \to 1$ a.s.

pf. $\mathbb{P}(X_n > r\log n) = \int_{r\log n}^{+\infty} e^{-x}\,dx = e^{-r\log n} = n^{-r}$

Since $\sum_{n\geq 1} n^{-r} = \begin{cases} < +\infty, & r > 1 \\ = +\infty, & r \leq 1 \end{cases}$, by Borel Cantelli,

$\mathbb{P}[X_n > r\log n \text{ i.o.}] = \begin{cases} 0 & \text{if } r > 1 \\ 1 & \text{if } r \leq 1 \end{cases}$

take $r=1$

$\Longrightarrow \mathbb{P}[X_n > \log n \text{ i.o.}] = 1 \Rightarrow \limsup \frac{X_n}{\log n} \geq 1 \Rightarrow \limsup \frac{M_n}{\log n} \geq 1$

Claim: for $\forall \varepsilon > 0$, $\limsup \frac{M_n}{\log n} \leq 1+\varepsilon$ a.s.

Since $\mathbb{P}[X_n > (1+\varepsilon)\log n, \text{ i.o.}] = 0 \Rightarrow \exists N_\varepsilon < +\infty$ s.t. $\forall n > N_\varepsilon$, $X_n \leq (1+\varepsilon)\log n$

$\Rightarrow \frac{\max\{X_n : n = N_\varepsilon+1, \cdots, n\}}{\log n} \leq 1+\varepsilon$, Also, $\frac{\max\{X_n : n=1,\cdots,N_\varepsilon\}}{\log n} \xrightarrow{n\to+\infty} 0$

$\Rightarrow \limsup \frac{M_n}{\log n} \leq 1+\varepsilon$ a.s.

#

Sending $\varepsilon \searrow 0$, we get $\limsup \frac{M_n}{\log n} = 1$.

Claim: $\forall \varepsilon > 0$, $\liminf \frac{M_n}{\log n} \geq 1-\varepsilon$ a.s.

$\mathbb{P}(M_n < (1-\varepsilon)\log n) = \mathbb{P}[X_1 < (1-\varepsilon)\log n]^n$

$= \left(1 - \frac{1}{n^{1-\varepsilon}}\right)^n \leq 2 \cdot \left(\frac{1}{e}\right)^{n^\varepsilon}$

Since $\sum_{n\geq 1} p_n \leq 2\sum_{n\geq 1}\left(\frac{1}{e}\right)^{n^\varepsilon} < +\infty$,

by B-C: $\mathbb{P}[M_n < (1-\varepsilon)\log n \text{ i.o.}] = 0$

$\Rightarrow$ with prob. 1, $\liminf \frac{M_n}{\log n} \geq 1-\varepsilon$.

#

$\liminf \frac{M_n}{\log n} \leq \limsup \frac{M_n}{\log n} \leq 1$.

$\therefore \liminf \frac{M_n}{\log n} = 1$.

Thus, $\frac{M_n}{\log n} \to 1$ a.s.

Also, $\limsup \frac{X_n}{\log n} \leq \limsup \frac{M_n}{\log n} \leq 1$,

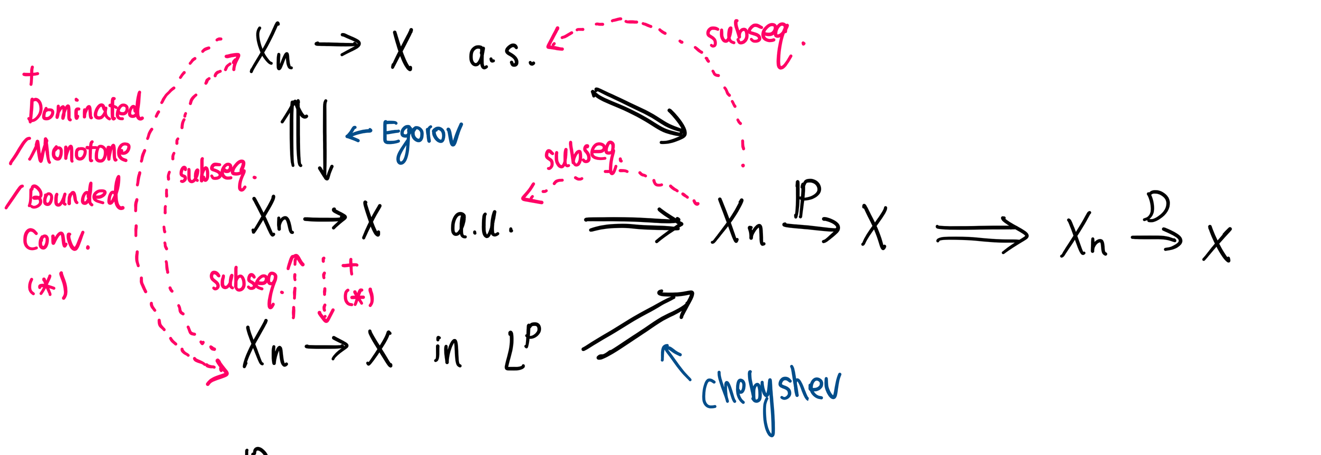
& $\limsup \frac{X_n}{\log n} \geq 1$

$\Rightarrow \limsup \frac{X_n}{\log n} \to 1$ a.s.

#

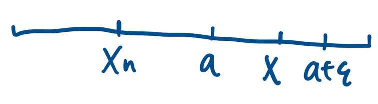
Recall:
- $X_n \to X$ a.s. $\Leftrightarrow \forall \varepsilon > 0$, $\mathbb{P}\left(\bigcap_{k=1}^{+\infty}\bigcup_{i=k}^{+\infty}\{|X_i - X| > \varepsilon\}\right) = 0$
- $X_n \to X$ a.u. $\Leftrightarrow \forall \varepsilon > 0$, $\mathbb{P}\left(\bigcup_{i=k}^{+\infty}\{|X_i - X| > \varepsilon\}\right) \xrightarrow{k\to+\infty} 0$
- $X_n \xrightarrow{\mathbb{P}} X \Leftrightarrow \forall \varepsilon > 0$, $\mathbb{P}(\{|X_i - X| > \varepsilon\}) \xrightarrow{i\to+\infty} 0$

Relation Map:



Thm. $X_n \xrightarrow{\mathbb{P}} X \Rightarrow X_n \xrightarrow{\mathcal{D}} X$.

pf. Note that $\forall a \in \mathbb{R}$, $\forall \varepsilon > 0$, $\{X_n \leq a\} \subseteq \{X \leq a+\varepsilon\} \cup \{|X_n - X| > \varepsilon\}$



$\Rightarrow \mathbb{P}(X_n \leq a) \leq \mathbb{P}(X \leq a+\varepsilon) + \underbrace{\mathbb{P}(|X_n - X| > \varepsilon)}_{\to 0}$

Similarly, $\{X \leq a-\varepsilon\} \subseteq \mathbb{P}(X_n \leq a) + \underbrace{\mathbb{P}(|X_n - X| > \varepsilon)}_{\to 0}$

By using $X_n \xrightarrow{\mathbb{P}} X$, taking $n \nearrow +\infty$, we have

$\limsup \mathbb{P}(X_n \leq a) \leq \mathbb{P}(X \leq a+\varepsilon) \leftarrow F_X(a+\varepsilon)$

$\liminf \mathbb{P}(X_n \leq a) \geq \mathbb{P}(X \leq a-\varepsilon) \leftarrow F_X(a-\varepsilon)$

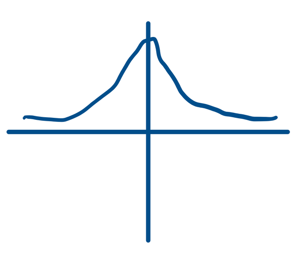
If a is a continuity point of F_X, send $\varepsilon \searrow 0$

to conclude $\lim \mathbb{P}(X_n \leq a) = \mathbb{P}(X \leq a)$

#

(Lec 22 only talked about Cauchy(1))

eg. $\boxed{X_n \sim \mathrm{Cauchy}(n),}$ $f_n(x) = \frac{n}{(1+n^2x^2)\pi}$.



Then $X_n \xrightarrow{\mathcal{D}} 0$.

Sol. In fact, $X_n \xrightarrow{\mathbb{P}} 0$: $\forall \varepsilon > 0$, $\mathbb{P}(|X_n| > \varepsilon) = 2\int_\varepsilon^{+\infty} \frac{n}{(1+n^2x^2)\pi}\,dx$

$= 2\left(1 - \frac{2}{\pi}\arctan(n\varepsilon)\right) \to 0$ as $n \to +\infty$

#

eg. $\boxed{\text{[Weak Law of Large Numbers 2]}}$ (Lec 23)

Let $X_1, \cdots, X_n$ be i.i.d. r.v.'s with finite mean u

and finite variance $\mathbb{E}X_i^2 < +\infty$. Then $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{\boxed{\mathbb{P}}} u$

pf. In fact, $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{L^2} u$: $\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n (X_i - u)\right]^2 = \frac{1}{n^2} \cdot \sum_{i,j=1}^n \overbrace{\mathbb{E}(X_i - u)(X_j - u)}^{\mathrm{Cov}(X_i, X_j)}$

$= \frac{1}{n^2} \cdot \sum_{i=1}^n \mathrm{Var}\, X_i \leq \frac{C}{n} \to 0$ as $n \to +\infty$

#

Prop. $X_n \xrightarrow{\mathbb{P}} X \Leftrightarrow$ for any subseq. $(X_{n'})$, there exists a further subseq. (X_{n_k}) s.t. $X_{n_k} \to X$ $\boxed{\text{a.u.}}$

$\Leftrightarrow$ for any subseq. $(X_{n'})$, there exists a further subseq. (X_{n_k}) s.t. $X_n \to X$ $\boxed{\text{a.s.}}$

pf. $\Leftarrow$): (ex.) By Contraction

$\Rightarrow$): for any subseq. $(X_{n'})$, $X_{n'} \xrightarrow{\mathbb{P}} X$

Therefore $\forall k \geq 1$, $\mathbb{P}(|X_{n'} - X| > \frac{1}{k}) \to 0$ as $n' \to +\infty$

Choose n_k $(> n_{k-1})$ s.t. $\mathbb{P}(|X_{n_k} - X| > \frac{1}{k}) \leq \frac{1}{2^k}$.

$\Rightarrow \mathbb{P}\left(\bigcup_{k=m}^{+\infty}\{|X_n - X| > \frac{1}{k}\}\right) \leq \frac{1}{2^{m-1}}$

for m sufficiently large

For $\forall \varepsilon > 0$, $\mathbb{P}\left(\bigcup_{k=m}^{+\infty}\{|X_{n_k} - X| > \varepsilon\}\right) \leq \mathbb{P}\left(\bigcup_{k=m}^{+\infty}\{|X_{n_k} - X| > \frac{1}{k}\}\right) \leq \frac{1}{2^{m-1}} \to 0$ as $m \to +\infty$

$\Rightarrow X_{n_k} \to X$ a.u.

#

5/3 HTOP (Reci 13)

Friday, May 3, 2024 11:29 AM

Examinable Content

- Joint distribution ; covariance $\Rightarrow$ Recover marginal / cond. density (Lec 13 & 14)
- Conditional distri. & Cond. expectation (Lec 15 & 16)
  - Towering (Lec 15)
  - Important examples (bivariate normal) (Lec 16)
- Functions of R.V.'s
  - examples related to exponential & bivariate normal (Lec 19)
  - Box - Muller (Lec 19)
- Random Walks (Lec 16 ~ 18)
  - recurrence / transience , Markov Prop. (Lec 16)
  - Condition on the 1^{st} step $\Rightarrow$ recurrence relations / PDE s (Lec 17)
  - path counting , reflection principle , Ballot Thm. (Lec 18)
- Generating funct. / Moment generating funct. (Lec 20 ~ 21)
  - solve recurrence (Lec 20)
  - relation of moments (Taylor expansion) (Lec 20)
  - sum of indep. R.V.'s / Random Sum Formula (Lec 20)
  - Branching process (Lec 21)
  - Joint generating funct. / moment generating funct. (Lec 21)
- Characteristic funct. (Lec 22 ~ 23)
  - Compute charact. funct. $<$ analytic extension (Lec 23) ; complex integration / residue (Lec 22)
- GF / MGF / CF 's applications to conv. in distri. (Lec 20 , 21 , 22 , 23)
  $\downarrow$ Cauchy(1)
- Convergence of R.V.'s (Lec 22 , 24 ~ 25)
  - Different modes of conv. and their relations (Lec 24 ~ 25)
  - Conv. Thm's : Fatou ; Dominated Convergence ; Bounded convergence (Lec 22)
  - Criteria of a.s. / a.u. conv. , Egorov Thm. (Reci 12)
  - Borel - Cantelli ; sufficient criteria for a.s. conv. ; important examples (Reci 12)
  - Weak Law of Large Numbers ($\mathbb{P}$. conv. & Distri. conv. , 2 proofs) ; Central Limit Thm. (1 pf. by charact. funct.) (Lec 23 , 25 , Reci 12)
  - Chebyshev Ineq. ($L^p \Rightarrow$ conv. in $\mathbb{P}$.) ; estimate tail prob. ; Chernoff Bound $\longleftrightarrow$ Combine with MGF estimates (Lec 24)
    $\downarrow$ Lec 24 in the proof of extended Borel - Cantelli
  - Conv. in $\mathbb{P}$ $\Rightarrow$ Subsequence conv. a.s. (Lec 25)

eg. (X_n) i.i.d. coin flips , $\mathbb{P}(H) = \mathbb{P}(T) = \frac{1}{2}$.

Let L_n be the length of the longest run of heads at time n.

for example , HHTTTTHHT $\underbrace{\text{HHHH}}_{L_n = 4}$TTHT | n

$l_j = 4$, $l_j = 1$, $l_j = 0$

Show that $\frac{L_n}{\log_2 n} \to 1$ a.s.

pf. Let l_j = the length of run of heads at time j.

$\mathbb{P}(l_j = k) = (\frac{1}{2})^{k+1}$ and $L_n = \max_{j \leq n} l_j$

Upper Bound : for $\forall \varepsilon > 0$,

$$\mathbb{P}[l_n > (1+\varepsilon)\log_2 n] = \sum_{k > (1+\varepsilon)\log_2 n} (\tfrac{1}{2})^{k+1} \leq (\tfrac{1}{2}) \cdot (\tfrac{1}{2})^{(1+\varepsilon)\log_2 n} = \tfrac{1}{2} \cdot n^{-(1+\varepsilon)} \leq n^{-(1+\varepsilon)}$$

Since $\sum_n \mathbb{P}[l_n > (1+\varepsilon)\log_2 n] < +\infty \Rightarrow \mathbb{P}[l_n > (1+\varepsilon)\log_2 n \text{ , i.o.}] = 0$

$\Rightarrow \exists N_\varepsilon$ s.t. for $\forall n > N_\varepsilon$, $l_n \leq (1+\varepsilon)\log_2 n \Rightarrow \limsup \frac{L_n}{\log_2 n} \leq 1+\varepsilon$ a.s.

Lower Bound : We want for $\forall \varepsilon > 0$, $\liminf \frac{L_n}{\log_2 n} \geq 1-\varepsilon$

H $\cdots$ H

0 $\underbrace{\quad}_{(1-\varepsilon)\log_2 n}$ n

$\mathbb{P}(\text{all H in each block}) = (\frac{1}{2})^{(1-\varepsilon)\log_2 n} = n^{-(1-\varepsilon)}$

$$\mathbb{P}(L_n < (1-\varepsilon)\log_2 n) \leq \mathbb{P}(\text{all blocks fail to have all H})$$

$$= \left(1 - \frac{1}{n^{1-\varepsilon}}\right)^{\frac{n}{(1-\varepsilon)\log_2 n}}$$

$$= \Bigg[\underbrace{\left(1 - \frac{1}{n^{1-\varepsilon}}\right)^{n^{1-\varepsilon}}}_{e^{-1} < 1}\Bigg]^{\underbrace{\frac{n^{\varepsilon}}{(1-\varepsilon)\log_2 n}}_{\xrightarrow{n\to\infty} \infty}}$$

$$\longrightarrow 0 \quad \text{as} \quad n \to \infty$$

By Borel - Cantelli , $\mathbb{P}(L_n < (1-\varepsilon)\log_2 n \text{ , i.o.}) = 0$

$\Rightarrow \exists N_\varepsilon$ s.t. for $\forall n > N_\varepsilon$, $\frac{L_n}{\log_2 n} \geq 1-\varepsilon$.

#

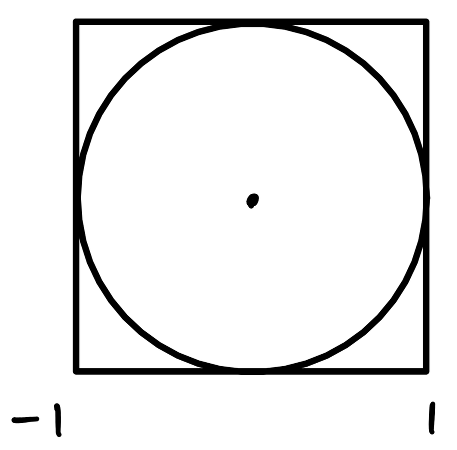
5/6 HTOP 26

Monday, May 6, 2024 12:01 PM

- Law of Large Numbers

Sample mean $\xrightarrow{n \nearrow +\infty}$ Theoretical mean

eg. (Monte Carlo Simulation): Numerically simulate π



+ generate $U_i, V_i \sim \text{Unif}[-1,1]$, indep., $i=1,\cdots,n$
- If $U_i^2 + V_i^2 \le 1$, set $X_i = 1$

otherwise, $X_i = 0$

Note that $\mathbb{E}X_i = \mathbb{P}(X_i = 1) = \frac{\pi}{4}$

+ Sample mean: $\underbrace{\overline{X_n}}_{\text{good approximator of } \frac{\pi}{4}} \triangleq \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathbb{P}} \mathbb{E}X_i = \frac{\pi}{4}$

- WLLN: Let (X_n) be a seq. of uncorrelated r.v.'s $\mathbb{E}[X_i X_j] = \mathbb{E}X_i \cdot \mathbb{E}X_j$, with the same distri. s.t. $\mathbb{E}X_i = u$, $\mathbb{E}X_i^2 < +\infty$. Then $\overline{X_n} = \frac{1}{n}\sum_i X_i \xrightarrow{\mathbb{P}} u$

Def. Given r.v.'s (X_n), a confidence interval for the theoretical mean u with confidence level $\beta\%$ is an interval of length 2ε s.t. $\mathbb{P}[u \in (\overline{X_n} - \varepsilon, \overline{X_n} + \varepsilon)] \ge \frac{\beta}{100}$

Rmk. The confidence interval is defined upon n, ε, β.

eg. What is the smallest n s.t. one obtains a 4-digit accuracy of π with confidence level 99%

Sol. $\varepsilon = \frac{1}{4000}$, $\beta = 99$

We look for n s.t. $\mathbb{P}\left(|\overline{X_n} - \frac{\pi}{4}| > \frac{1}{4000}\right) \le 1\%$ (*)

By Chebyshev $\mathbb{P}\left(|\overline{X_n} - \frac{\pi}{4}| > \frac{1}{4000}\right) \le 1\%$

Other methods: [large deviation / Chernoff]

$$\mathbb{P}\left(|\overline{X_n} - \frac{\pi}{4}| > \frac{1}{4000}\right) \le 4000^2 \cdot \text{Var}\,\overline{X_n}$$
$$= 4000^2 \cdot \frac{1}{n^2} \cdot \sum_{i=1}^{n} \text{Var}\,X_i$$
$$\le \frac{4000^2}{n^2} \cdot \left(\frac{1}{4} n\right)$$
$$= \frac{4 \times 10^8}{n}$$

Bernoulli$(n, \frac{\pi}{4})$

$\text{Var}\,X_i = \frac{\pi}{4}\left(1 - \frac{\pi}{4}\right) \le \frac{1}{4}.$

If $n \ge 4 \times 10^8$, then (*)

\#

eg. Coupon Collector: $\Omega = \{1, \cdots, n\}$, X_i i.i.d. unif. $\{1, \cdots, n\}$

Denote by $T_i = \inf\{n : |\{X_1, \cdots, X_n\}| = i\}$ the 1st time to collect i different coupons

We showed, $\frac{\mathbb{E}T_n}{n\log n} \to 1$

Claim: $\frac{T_n}{n\log n} \xrightarrow{\mathbb{P}} 1$

pf. Indeed, $T_n = \sum_{i=1}^{n}(T_i - T_{i-1})$, $(T_i - T_{i-1})$ indep. $\sim \text{Geo}\left(1 - \frac{i-1}{n}\right)$

We already showed: $\mathbb{E}T_n = n\log n + o(n)$

$$\text{Var}\,T_n = \sum_{i=1}^{n} \text{Var}(T_i - T_{i-1})$$

$$= \sum_{i=1}^{n} \frac{\frac{i-1}{n}}{\left(1 - \frac{i-1}{n}\right)^2} \le \sum_{i=1}^{n} \frac{1}{\left(1 - \frac{i-1}{n}\right)^2} = n^2 \cdot \sum_{m=1}^{n} \frac{1}{m^2} \le C n^2 \quad \text{where we can pick } C = \sum_{m=1}^{+\infty} \frac{1}{m^2} = \frac{\pi^2}{6}$$

$Y \sim \text{Geo}(p)$: $\text{Var}\,Y = \frac{1-p}{p^2}$.

In particular, $\frac{\text{Var}\,T_n}{(\mathbb{E}T_n)^2} \le \frac{C n^2}{(n\log n)^2 + o(n)} = \frac{C}{\log^2 n + o(\frac{1}{n})} \to 0$ as $n \to +\infty$

Then by Extension of WLLN: $\frac{T_n - \mathbb{E}T_n}{n\log n} \xrightarrow{\mathbb{P}} 0 \Rightarrow \frac{T_n}{n\log n} \xrightarrow{\mathbb{P}} 1$

- Extension of the Weak Law of Large Numbers:

$T_n = \sum_{i=1}^{n} X_i$, for every seq. (a_n) s.t. $\frac{\text{Var}\,T_n}{a_n^2} \to 0$ as $n \nearrow +\infty$.

We have $\frac{T_n - \mathbb{E}T_n}{a_n} \xrightarrow{\mathbb{P}} 0$ (ex.)