LECTURE 1

LOGISTICS

(1) TAs: Debsurya De, Yufei Zhan

(2) Survey form on Canvas

(3) OH survey: this week

(4) HW 1: next week (Prob. Th. I review)

Themes of PTI / probability you've seen?

- Measure theory: (Kolmogorov 1933)
  cf. Bertrand random chord paradox (1889)
  clarifies fuzzy notions of "random", "uniformly random", etc.

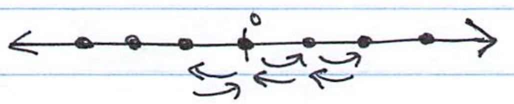
- Limit theorems: sequences of r.v. (S_n) "converge" (LLN, random series) or sequences of $\mathbb{P}[A_n]$ (CLT, LDP).

- Independence: (= product measures) fundamental assumption leading to all main limit theorems (though can be loosened)

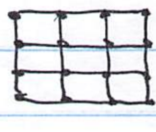
- Key example: sums of iid: $X_i \overset{iid}{\sim} \rho$, $S_n := \sum_{i=1}^{n} X_i$.

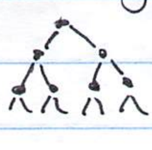
This class: generalizations of sum-of-iid model, sampling of more advanced topics.

- $X_i = \begin{cases} +1 & \text{w/p } 1/2 \\ -1 & \text{w/p } 1/2 \end{cases} \longrightarrow (S_n)$ describes random walk on $\mathbb{Z}$



Q: Random walk on graph G?

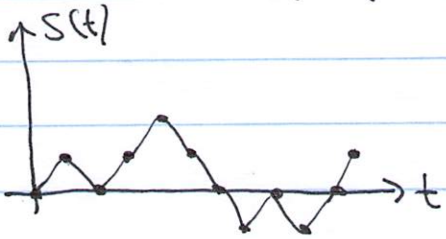




$\longrightarrow$ Markov chains.

- X_i = winnings from bet of \$1.
  Adaptive strategy: bet depending on history $(S_0, S_1, \dots, S_{n-1})$
  $\longrightarrow$ Martingales.
  Surprisingly related: analysis of $f(X_1, \dots, X_n)$ beyond summation.

- Can view $(S_0, S_1, \dots, S_n, \dots)$ as $S: \mathbb{Z}_{\geq 0} \to \mathbb{R}$



interpolate $\to$ <u>random function</u>.

"Zoom out" appropriately $\to$ "converge" to <u>Brownian motion</u>
... for <u>any</u> (nice) p (Donsker: "dynamical CLT")

Tools for random functions $\mathbb{R} \to \mathbb{R}$ = <u>stochastic calculus</u>.
(Rk: functions $\mathbb{R}^d \to \mathbb{R}$, manifold $\to \mathbb{R}$ = <u>random fields</u>.)

Review of Prob Th I ideas:

<u>Rk</u>: Truncation argument to upgrade to $\mathbb{E}|X_i| < \infty$.

① <u>(Weak) LLN</u>: If $\mathbb{E}X_i^2 < \infty$, $\frac{1}{n} S_n \xrightarrow{\mathbb{P}} \mathbb{E}X_1 =: \mu$.
i.e., $\forall \varepsilon > 0$, $\mathbb{P}\left[\left|\frac{1}{n} S_n - \mu\right| > \varepsilon\right] \to 0$.

Pf: $$= \mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right| > \varepsilon\right]$$
$$\leq \frac{1}{\varepsilon^2} \operatorname{Var}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right] \qquad \text{(Chebyshev)}$$
$$= O\left(\frac{1}{n}\right) \to 0.$$

"Strong" version: $\frac{1}{n} S_n \xrightarrow{a.s.} \mu$, i.e. $\mathbb{P}\left[\lim \frac{1}{n} S_n = \mu\right] = 1$.
Fundamentally different: weak about $(\mathrm{Law}(S_n))$, strong about $\mathrm{Law}((S_n))$.
Main tool: <u>Borel–Cantelli</u>.

2 Characteristic functions : $\phi_X(t) := \mathbb{E}\exp(itX)$

Rk: If X has density $f(x)$, $= \int_{-\infty}^{\infty} \exp(itx) f(x)\,dx$
$=$ Fourier transform of f.

Main properties:

- $\text{Law}(X) = \text{Law}(Y) \iff \phi_X = \phi_Y$
- $\phi_{X_n}(t) \longrightarrow \phi_Y(t)\ \forall t \in \mathbb{R} \Rightarrow X_n \Rightarrow Y$ — "conv. in distribution"
  (Clearer: $\text{Law}(X_n) \xrightarrow{\text{weak}} \text{Law}(Y)$, i.e. $\mathbb{E}f(X_n) \longrightarrow \mathbb{E}f(Y)\ \forall$ bdd, cts. f.) — Continuity Thm.
- X and Y independent $\Longrightarrow \phi_{X+Y} = \phi_X \phi_Y$
  (Fourier transform effect on convolution / translation.)

Point: computational tool for conv. in distribution.

Ex: Alternative approach to weak LLN:

$$\phi_{\frac{1}{n}S_n}(t) = \phi_{\frac{1}{n}X_1 + \cdots + \frac{1}{n}X_n}(t)$$
$$= \phi_{\frac{1}{n}X_1}(t) \cdots \phi_{\frac{1}{n}X_n}(t)$$
$$= \phi_{\frac{1}{n}X_1}(t)^n$$
$$= \left(\mathbb{E}\exp\left(it\frac{X_1}{n}\right)\right)^n$$
$$= \left(1 + \frac{it\,\mathbb{E}X_1}{n} + O\left(\frac{1}{n^2}\right)\right)^n \longrightarrow \exp(it\mu) = \phi_\mu(t)$$

deterministic.

Def: dist. fn.
$F_X(t) := \mathbb{P}[X \leq t]$

$X_n \Rightarrow Y$ iff $F_{X_n}(t) \to F_Y(t)$ when F_Y cts. at t.

3 Lyapunov Proof of CLT:

Statement: if (X_i) iid, $\mathbb{E}X_i = \mu$, $\mathrm{Var}\,X_i = \sigma^2$, $\mathbb{E}|X_i|^3 < \infty$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu) \Longrightarrow \mathcal{N}(0, \sigma^2).$$

↳ can remove with small modification.

Pf: WLOG $\mu = 0$, so LHS $= \frac{1}{\sqrt{n}} S_n$.

(same steps)

$$\phi_{\frac{1}{\sqrt{n}}S_n}(t) = \cdots = \phi_{\frac{1}{\sqrt{n}}X_1}(t)^n$$

$$= \left(\mathbb{E}\exp\left(it\frac{X_1}{\sqrt{n}}\right)\right)^n = \left(1 + \cancelto{0}{\frac{it}{\sqrt{n}}\mathbb{E}X_1} - \frac{t^2}{2n}\sigma^2 + O\left(\frac{1}{n^{3/2}}\right)\right)^n$$

$$\longrightarrow \exp\left(-\frac{\sigma^2 t^2}{2}\right) = \phi_{\mathcal{N}(0,\sigma^2)}(t).$$

Easy proof, but rather magical / mysterious — not a probabilistic explanation of why CLT holds.

Next week: more "hands-on" proof technique (Lindeberg exchange method) ⇝ ways to handle X_i non-iid and to get quantitative versions of weak convergence, i.e.

$$\left| \mathbb{E}f\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu)\right) - \underset{g \sim \mathcal{N}(0,\sigma^2)}{\mathbb{E}} f(g) \right| \leq ???$$