THE CENTRAL LIMIT THEOREM

- Consider a simple random walk

$$S_n = X_1 + \cdots + X_n \qquad \text{where } X_i \sim \text{Rademacher iid}.$$

SLLN $\Rightarrow$

$$\frac{S_n}{n} \xrightarrow{a.s.} 0 \qquad \text{as } n \to \infty.$$

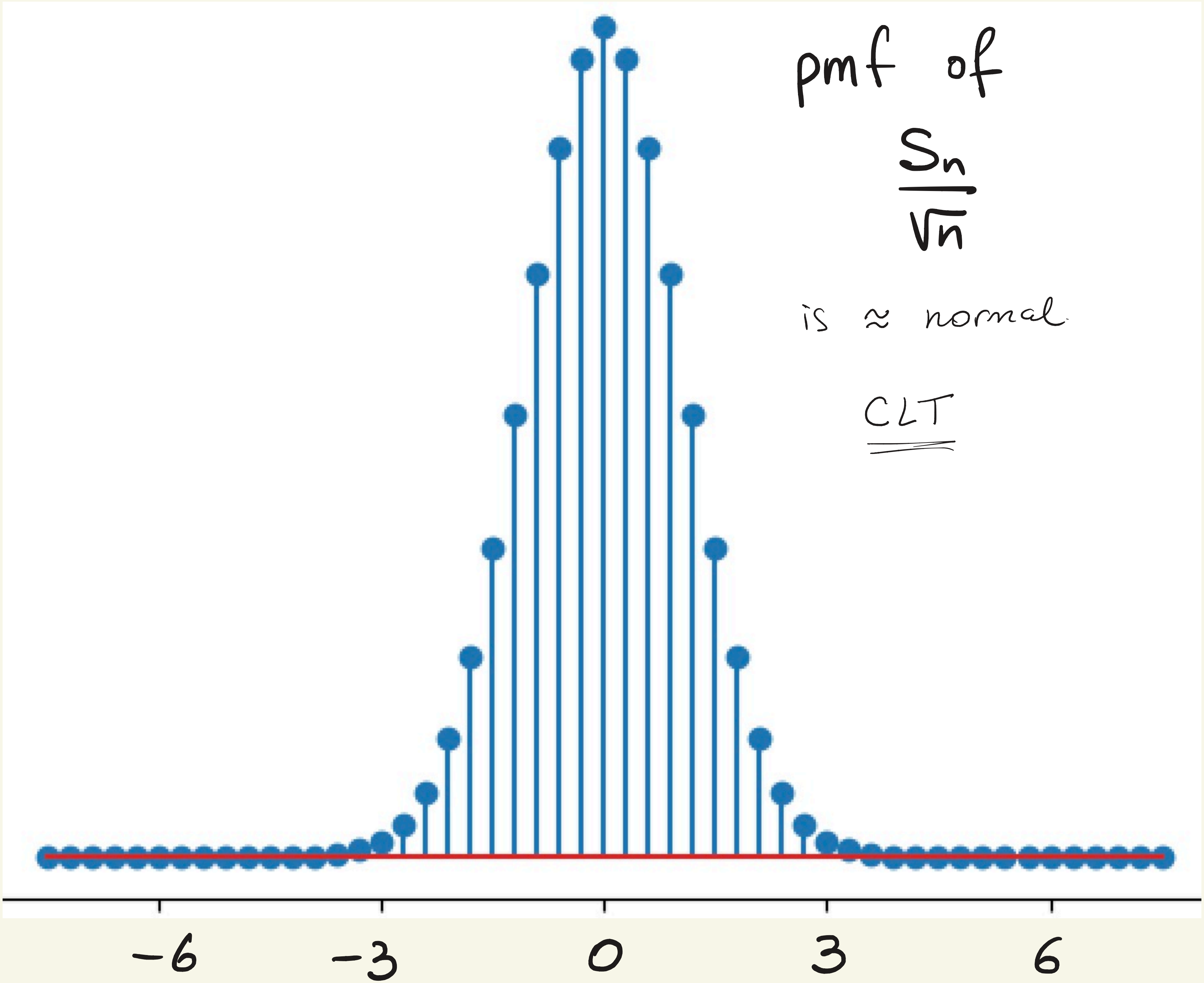
- But $\mathrm{Var}(S_n) = n$, so it makes more sense to normalize by the standard deviation of S_n, which is $\sqrt{n}$:

$$\frac{S_n}{\sqrt{n}} \to ?$$

- ? can't be 0 anymore, or any constant $\left(\mathrm{Var}\left(\frac{S_n}{\sqrt{n}}\right) = 1 \right)$.

pmf of $\frac{S_n}{\sqrt{n}}$:

Made with Goodnotes



pmf of $\frac{S_n}{\sqrt{n}}$ is $\approx$ normal

CLT

A PHYSICS PROOF OF THE CLT

- pmf: $f(x,n) := \mathbb{P}\{S_n = x\}$.

- $f(x,n+1) = \mathbb{P}\{S_{n+1} = x\} = \mathbb{P}\{S_n = x+1,\ X_n = -1\} + \mathbb{P}\{S_n = x-1,\ X_n = 1\}$

$$= \mathbb{P}\{S_n = x+1\}\cdot \mathbb{P}\{X_n = -1\} + \mathbb{P}\{S_n = x-1\}\cdot \mathbb{P}\{X_n = 1\} \quad \text{(independence)}$$

$$= f(x+1,n)\cdot \tfrac{1}{2} + f(x-1,n)\cdot \tfrac{1}{2}$$

- Subtract $f(x,n)$ from both sides, rearrange the terms $\Rightarrow$

$$\underbrace{f(x,n+1) - f(x,n)}_{\approx\ \frac{\partial f}{\partial n}(x,n)\ \text{Heuristically}} = \underbrace{\frac{f(x+1,n) - f(x,n)}{2}}_{\approx\ \frac{1}{2}\frac{\partial f}{\partial x}(x,n)} - \underbrace{\frac{f(x,n) - f(x-1,n)}{2}}_{\approx\ \frac{1}{2}\frac{\partial f}{\partial x}(x-1,n)}$$

$$\frac{1}{2}\frac{\partial f}{\partial x}(x,n) - \frac{1}{2}\frac{\partial f}{\partial x}(x-1,n) \approx \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x,n)$$

$$\Rightarrow \boxed{\frac{\partial f}{\partial n} = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}}$$

Heat (a.k.a. diffusion) PDE
with initial condition $f(x,0) = \delta(x)$

- Solve: the (fundamental) solution is

$$(*) \quad \boxed{f(x,n) = \frac{1}{\sqrt{2\pi n}} \exp\left(-\frac{x^2}{2n}\right)} = \text{pdf}(x) \text{ of } \boxed{N(0,\sqrt{n})}$$

$$\Rightarrow \boxed{S_n \approx N(0,\sqrt{n})} \leftarrow \text{CLT}$$

- Remarks
  1. A random walk models a diffusion (of 1 molecule)
  2. To make this rigorous, what does it even mean that $S_n \to N(0,1)$?
     Not $\xrightarrow{\mathbb{P}}$ or $\xrightarrow{a.s.}$;

     could be on different prob. spaces!

Made with Goodnotes

CONVERGENCE IN DISTRIBUTION

Def A sequence of r.v's $X_1, X_2, \dots$ converges in distribution (a.k.a. "weakly") to a r.v. X if the cdf's satisfy

$$F_n(x) \to F(x) \quad \forall \text{ continuity point } x \text{ of } F$$

$$\iff \quad P\{X_n \le x\} \to P\{X \le x\}. \quad \forall x \text{ such that } P\{X = x\} = 0$$

Denoted $X_n \xrightarrow{d} X$.

Remarks 1. All r.v's $X, X_1, X_2, \dots$ may be defined on different prob. spaces — unlike in a.s. convergence and conv. in probability.

most important mode of conv. in prob. theory

2. Since F is bounded & monotone, the set of discontinuities is finite or countable.

Examples (a) Let X be $\forall$ r.v. Then $X_n := X + \frac{1}{n} \xrightarrow{d} X$

$$F_n(x) = P\{X + \tfrac{1}{n} \le x\} = P\{X \le x - \tfrac{1}{n}\} = F(x - \tfrac{1}{n}) \to F(x)$$

if F is continuous at x.

This example shows why only the continuity pts are considered in the def.

(b) Consider a coin that comes up H with prob. p.

$X_p :=$ # coin flips until the first H. (= waiting time for 1^{st} success)

$\sim \text{Geom}(p)$ Geometric distribution

$$P\{X_p > k\} = P\{\underbrace{TT\dots T}_{k}\} = (1-p)^k, \quad k = 1, 2, \dots$$

Geometric distribution
$X \sim \text{Geom}(p)$

$$\Rightarrow P\{pX_p > x\} = (1-p)^{x/p} \to e^{-x} \quad \text{as } p \to 0.$$

Exponential distribution:
$Z \sim \text{Exp}(1)$ if $P\{Z > x\} = e^{-x}$ $\forall x > 0$

Hence, we proved:

Prop A limit thm for geometric distribution

If $X_p \sim \text{Geom}(p)$ then $pX_p \xrightarrow{d} Z$ as $p \to 0$ where $Z \sim \text{Exp}$.

Made with Goodnotes

Relation to other modes of convergence?

Prop $X_n \xrightarrow{P} X \quad \Rightarrow \quad X_n \xrightarrow{d} X$

Proof Fix a continuity pt x of $F \Rightarrow$

- $$F_n(x) = \mathbb{P}\{X_n \le x\} \overset{\forall \varepsilon>0}{\le} \underbrace{\mathbb{P}\{X_n \le x \,\&\, |X_n - X| \le \varepsilon\}}_{\overset{\wedge\!\!|\,\Delta}{\mathbb{P}\{X \le x+\varepsilon\} = F(x+\varepsilon)}} + \underbrace{\mathbb{P}\{|X_n - X| > \varepsilon\}}_{\downarrow \text{assumption} \atop 0}$$

$$\Rightarrow \limsup_n F_n(x) \le F(x+\varepsilon) \longrightarrow F(x) \text{ as } \varepsilon \downarrow 0 \quad (\text{continuity assm.})$$

$$\Rightarrow \limsup_n F_n(x) \le F(x).$$

- Similarly, the lower bd:

$$F(x-\varepsilon) = \mathbb{P}\{X \le x-\varepsilon\} \le \underbrace{\mathbb{P}\{X \le x-\varepsilon \,\&\, |X_n - X| \le \varepsilon\}}_{\overset{\wedge\!\!|\,\Delta}{\mathbb{P}\{X_n \le x\} = F_n(x)}} + \underbrace{\mathbb{P}\{|X_n - X| > \varepsilon\}}_{\downarrow \text{assm.} \atop 0}$$

$$\Rightarrow \liminf_n F_n(x) \ge F(x-\varepsilon) \to F(x) \text{ as } \varepsilon \downarrow 0 \quad (\text{continuity assm.})$$

$$\Rightarrow \liminf_n F_n(x) \ge F(x). \quad \square$$

Remark The converse is false, because X_n may be defined on different prob. spaces.

- A more concrete example:

$$X_n \sim \mathrm{Ber}(1/2) \text{ iid } \Rightarrow X_n \overset{d}{=} X_1 \quad \text{but} \quad X_n \overset{P}{\not\to} X_1$$

$$\left(\mathbb{P}\{|X_n - X_1| > \tfrac{1}{2}\} = \tfrac{1}{2} \not\to 0 \right)$$

However:

Made with Goodnotes

THM (Skorokhod Representation Theorem)

Assume $X_n \xrightarrow{d} X$. Then $\exists$ r.v's $Y, Y_1, Y_2, \dots$ defined on the same probability space, and such that:

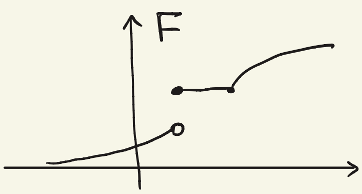
$$Y \overset{d}{=} X, \quad Y_n \overset{d}{=} X_n, \quad Y_n \xrightarrow{a.s.} Y.$$

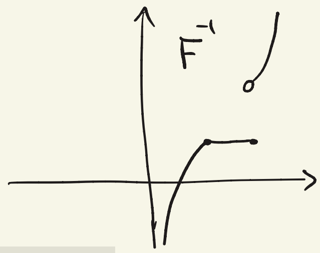
The proof uses the same method as we used to generate a r.v. from the uniform distribution: $X \overset{d}{=} \underset{\text{cdf of } X}{F}{}^{-1}(\overset{\text{Unif}[0,1]}{U})$ $(*)$

Def The generalized inverse of a nondecreasing function $F: \mathbb{R} \to \mathbb{R}$ is

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

- If F is one-to-one, F^{-1} is the usual inverse.





Lem If F is nondecreasing and right-continuous, then

$$x \ge F^{-1}(y) \iff F(x) \ge y$$

$\ulcorner$ $(\Leftarrow)$ by def of F

$(\Rightarrow)$ N↑ $F(x) < y \Rightarrow x \notin \{x' : F(x') \ge y\}$

$F \uparrow \Rightarrow x$ must be on the left of this set

F is right-continuous $\Rightarrow$ the set is closed

$\Rightarrow x < \inf\{x' : F(x') \ge y\} = F^{-1}(y)$ $\lrcorner$

Proof of $(*)$: $P\{F^{-1}(U) \le x\} \overset{\text{Lem}}{=} P\{U \le F(x)\} = F(x)$

Made with Goodnotes

Lem (convergence of inverse functions)

Let $F, F_1, F_2, \dots$ be nondecreasing & right-continuous.
If $F_n(x) \to F(x)$ for all continuity points x of F,
then $F_n^{-1}(y) \to F^{-1}(y)$ for all continuity points y of F^{-1}.

- **U.B.** : Fix a continuity pt y of F^{-1}, $\varepsilon > 0$,

$$x := F^{-1}(y+\varepsilon).$$

Fix $\forall \delta > 0$ s.t. $x+\delta$ is a continuity pt. of F (dense $\Rightarrow$ δ can be $\forall$ small)

Assm $\Rightarrow$ $F_n(x+\delta) \to F(x+\delta) \overset{\text{Lem}}{\geq} y+\varepsilon$

$\Rightarrow$ $F_n(x+\delta) \geq y$ for large enough n.

Lem $\Rightarrow$ $x+\delta \geq F_n^{-1}(y)$

$\Rightarrow$ $\limsup_n F_n^{-1}(y) \leq x+\delta = F^{-1}(y+\varepsilon)+\delta$

Since $\varepsilon, \delta > 0$ can be $\forall$ small,

$$\limsup_n F_n^{-1}(y) \leq F^{-1}(y).$$

- **LB** is similar: $x := F^{-1}(y-\varepsilon)$, $x-\delta$ is a cont. pt. of F

Assm $\Rightarrow$ $F_n(x-\delta) \to F(x-\delta) \overset{\text{Lem}}{<} y-\varepsilon$ $\Rightarrow$ $F_n(x-\delta) < y$ for large enough n

Lem $\Rightarrow$ $x_0 - \delta < F_n^{-1}(y)$ $\Rightarrow$ $\liminf_n F_n^{-1}(y) \geq x-\delta = F^{-1}(y-\varepsilon)-\delta$ $\overset{\forall \varepsilon,\delta}{\Rightarrow}$ $\liminf_n F_n^{-1}(y) \leq F^{-1}(y)$.

Made with Goodnotes

Proof of Skorokhod Representation Thm

Let F, F_n be cdf's of X, X_n. Set

$$Y := F^{-1}(U), \quad Y_n := F_n^{-1}(U) \quad \text{where } U \sim \text{Unif}[0,1]$$

- We already showed that $Y \overset{d}{=} X$, $Y_n \overset{d}{=} X_n$ (p. 117).

- $X_n \xrightarrow{d} X \overset{\text{def}}{\Rightarrow} F_n(x) \to F(x)$ $\forall$ continuity pts x of F

  $\overset{\text{lem}}{\Rightarrow} F_n^{-1}(u) \to F^{-1}(u)$ $\forall$ cont. pts u of F^{-1}

  The set of discontinuities is at most countable $\Rightarrow$

  U is a continuity pt a.s.

  $$\Rightarrow F_n^{-1}(U) \xrightarrow{a.s.} F^{-1}(U) \overset{\text{def}}{\Rightarrow} Y_n \xrightarrow{a.s.} Y. \qquad \square.$$

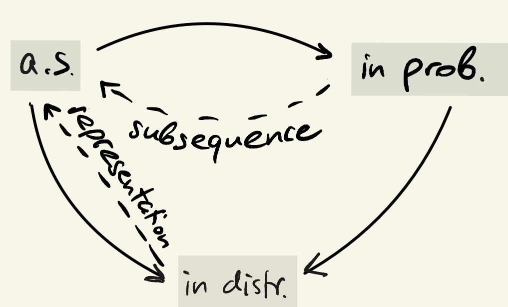
- **Remark** It is not true that the joint distr's are the same:

$$(X_1, X_2, \dots) \overset{d}{=} (Y_1, Y_2, \dots) \qquad (*)$$

(Example: iid $X_n \sim \text{Ber}(1/2)$ satisfy $X_n \overset{d}{=} X_1$ but $(*)$ would imply that $Y_n \sim \text{Ber}(1/2)$ are iid, and Y_n don't converge a.s.)

- Remark: "a.s." $\to$ "everywhere" (by redefining $F_n^{-1}, F \equiv 0$ at discontinuities)

- Summary of the modes of convergence:



Made with Goodnotes