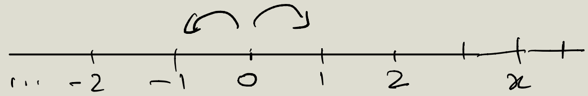
TOWARD THE CENTRAL LIMIT THEOREM ("CLT")

Ex A simple random walk:



$S_n = X_1 + \cdots + X_n$ where $X_i \sim$ Rademacher, independent.

Approximate the pmf $P\{S_n = x\} =: f(x,n)$?

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

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

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

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

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

Heuristically:

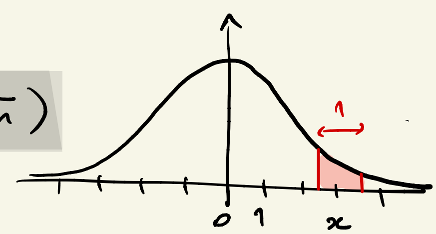
$$\frac{\partial f}{\partial n}(x,n) \qquad \frac{1}{2}\frac{\partial f}{\partial x}(x,n) - \frac{1}{2}\frac{\partial f}{\partial x}(x-1,n) \approx \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x,n)$$

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

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

- Solve: the (fundamental) solution is

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



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

- Remarks
  1. A random walk models a diffusion (of 1 molecule)
  2. A rigorous proof of $(*)$: express $f(x,n)$ in terms of binomial coefficients; use Taylor's approximation $\Rightarrow$ "De Moivre–Laplace Thm"

  ☹ This argument is limited to binomial distr.

  For general X_n, what version of $(*)$ do we expect? what does it even mean that $S_n \to N(0,1)$? Not $\xrightarrow{P}$ or $\xrightarrow{a.s.}$ ← could be on different prob. spaces!

WEAK CONVERGENCE

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

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

Denoted $X_n \xrightarrow{w} X$

cdf's: $F_n(x) = P\{X_n \le x\}$, $F(x) = P\{X \le x\}$.

Remark F is bounded and monotone $\Rightarrow$ has at most countable # of discontinuities.

Examples

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

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

if F is continuous at x.

This example shows why we require convergence only at points of continuity of F.

(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

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

Substitute $k = x/p \Rightarrow$

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

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

Hence, we proved:

A limit thm for geometric distribution

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

Relation of weak convergence to other modes of convergence?

Prop $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{w} X$

Proof $\forall$ point of continuity x of F:

$$F_n(x) = P\{X_n \le x\} \overset{\Delta}{\le} \underbrace{P\{|X_n - X| > \varepsilon\}}_{\downarrow \text{(assumption)}\ 0} + \underbrace{P\{X \le x+\varepsilon\}}_{= F(x+\varepsilon)}$$

$\Rightarrow \limsup_n F_n(x) \le F(x+\varepsilon) \longrightarrow F(x)$ as $\varepsilon \downarrow 0$ (continuity)

$\Rightarrow \limsup_n F_n(x) \le F(x)$. Similarly, $\liminf_n F_n(x) \ge F(x)$ (DIY) $\square$.

Remark The converse is false: X_n can be on different prob. spaces.

Example: let $X, X_1, X_2, \dots$ be iid r.v's. Then

$X_n \overset{distr}{=} X \Rightarrow X_n \xrightarrow{w} X$ trivially, but $X_n \overset{P}{\nrightarrow} X$ (why?)

However:

THM (Almost sure representation) [Skorokhod]

If $Y_n \xrightarrow{w} Y$ then $\exists$ r.v's $X, X_1, X_2 \dots$ all on the same prob. space and such that $X \overset{dist}{=} Y$, $X_n \overset{dist}{=} Y_n$ $\forall n$, and

$X_n \to X$ everywhere ($\Rightarrow$ a.s.)

Proof The construction is the same as in the proof that $\forall$ r.v. can be realized on $\Omega = [0,1]$ ($\mathcal{F} = \mathcal{B}$, P = Lebesgue) (p.17)

Let $F, F_1, F_2, \dots$ be cdf's of $Y, Y_1, Y_2, \dots$

① Assuming that F, F_n are continuous and strictly increasing:

$$X(x) := F^{-1}(x), \quad X_n(x) := F_n^{-1}(x), \quad x \in [0,1]$$

Then cdf of X is F, cdf of X_n is F_n

(indeed, $P\{X \le a\} = \mu\{x \in [0,1] : \underbrace{F^{-1}(x) \le a}_{\Updownarrow\ x \le F(a)}\} = F(a)$)

$Y_n \xrightarrow{w} Y \Rightarrow F_n \to F$ pointwise. Assumption $\Rightarrow F^{-1}, F_n^{-1}$ are continuous, strictly increasing

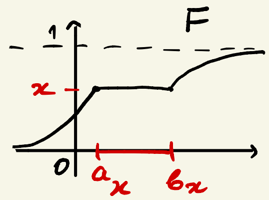
$\Rightarrow F_n^{-1} \to F^{-1}$ pointwise (check!) $\Rightarrow X_n \to X$ everywhere. $\square$

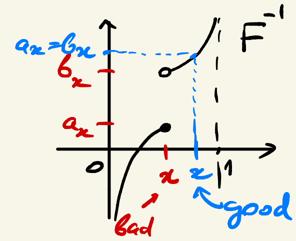
HW.

② General case:

$X(x) := F^{-1}(x) := \sup\{y : F(y) < x\}$

$X_n(x) := F_n^{-1}(x) := \sup\{y : F_n(y) < x\}$





Decompose $\Omega = [0,1] = \Omega_{\smile} \sqcup \Omega_{\frown}$

where $a_x := \sup\{y : F(y) < x\}$, $b_x := \inf\{y : F(y) > x\}$,

$\Omega_{\smile} = \{x : a_x = b_x\}$.

- **$\Omega_{\frown}$ is countable** since the intervals (a_x, b_x) are disjoint (each (a_x, b_x) contains a different rational #)

- **F is strictly increasing on $\Omega_{\smile}$**, i.e. $\forall x \in \Omega_{\smile}$:

(*) $y > F^{-1}(x) \Rightarrow F(y) > x$ and $y < F^{-1}(x) \Rightarrow F(y) < x$

(indeed: $F^{-1}(x) = a_x = b_x$, def of x; $F^{-1}(x) = a_x$, def of a_x)

- **$\limsup F_n^{-1}(x) \le F^{-1}(x) \quad \forall x \in \Omega_{\smile}$**

Choose $\forall\, y > F^{-1}(x)$. (*) $\Rightarrow F(y) > x$.

If y is a point of continuity of F, then by assumption of thm,

$F_n(y) \to F(y) \longrightarrow F_n(y) > x$ for $\forall$ suff. large n

$\Rightarrow F_n^{-1}(y) \le y$ (by def. of F_n^{-1})

Summarizing:

$\limsup_n F_n^{-1}(x) \le y \quad \forall\, y > F^{-1}(x)$ that is a pt of continuity of F.

Since the points of continuity are everywhere dense (pts of discontinuity are countable)

$\Rightarrow$ qed.

- Similarly, $\liminf F_n^{-1}(x) \ge F^{-1}(x) \quad \forall x \in \Omega_{\smile}$ (DIY)

$\Rightarrow F_n^{-1}(x) \to F^{-1}(x) \quad \forall x \in \Omega_{\smile}$

- Recall $\Omega_{\frown}$ is countable $\Rightarrow$ null.

Redefine $X_n(x) = X(x) := 0$ for $x \in \Omega_{\frown}$

$\Rightarrow F_n^{-1}(x) \to F^{-1}(x) \quad \forall x \in \mathbb{R}$. QED

Remark It is NOT true that joint distr. is the same, i.e. that

$$(X, X_1, X_2, \dots) \overset{dist}{=} (Y, Y_1, Y_2, \dots)$$

e.g. if $Y, Y_1, Y_2, \dots$ are i.i.d. (indeed, $Y_n \xrightarrow{w} Y$ trivially but $X_n \overset{a.s}{\not\to}$ anywhere)

Summary of convergence of r.v's: